Studies of Galaxy Evolution
Using Stacking Techniques

Jacinta Delhaize B.Sc.
School of Physics, International Centre for Radio Astronomy Research
University of Western Australia

This thesis is presented for the degree of Doctor of Philosophy at the
University of Western Australia

2014
This thesis is dedicated to Professor Steven Rawlings (1961 - 2012). Steve, it feels like there was so much more I was supposed to learn from you. You were a brilliant supervisor, an inspiring mentor and a caring friend. I know you’re up there guiding photons my way. I hope this thesis does justice to your memory.
Acknowledgements

Completing a thesis is by no means a one-person feat. There are so many people whom I would like to sincerely thank for helping me through my PhD. First and foremost, I would like to express my eternal gratitude to my excellent supervisors Professor Lister Staveley-Smith, Associate-Professor Martin Meyer and Professor Steven Rawlings. Thank you for your generosity in sharing your wealth of knowledge, your precious time and your pearls of wisdom with me. Thank you for all your patience and dedication, for inspiring me and for challenging me.

My heart-felt thanks also to Dr Matt Jarvis, Dr Martin Hardcastle and Dr Tom Mauch for adopting me as a student when I was in need. It was a privilege to work with you and I am forever grateful.

A huge thanks to those who assisted with editing during the writing-up process: Associate-Professor Chris Power, Associate-Professor Minh Huynh, Associate-Professor Tobias Westmeier, Dr Chris Willott, Dr Daniel Cunnama, Dr Christopher Perrella, Alyssa Nixon-Lloyd and Corin Delhaize.

I would like to thank everyone at the International Centre for Radio Astronomy Research (ICRAR). I was fortunate to have been at ICRAR from its birth and to help shape the culture and traditions of the institute. ICRAR has been a fantastic working environment and has provided me with many opportunities to interact with the national and international research community. I wish ICRAR and WA radio astronomy all the best for the future - I’ll be back one day!

Thanks to Professor Peter Quinn, the ‘Big Boss’, for hooking me into astronomy and being a supportive mentor ever since the early years, and to Dr Renu Sharma for her genuine concern for student welfare. Thanks to all the many ICRAR astronomers, ICT staff and administration staff who have given me assistance and advice throughout my PhD.

A special thanks to all the ICRAR students. It was a privilege to share the postgraduate experience with you all and I am proud of how supportive, enthusiastic and friendly the ICRAR student community has become. I hope the Plot of the Week and afternoon tea traditions continue on into the future. Particular shout-outs go to my office mates Morag Scrimgeour, Stefan Westerlund and Jurek Malarecki for putting up with my frequent outbursts of elation or exasperation. Thanks also to Dr Toby Potter and Stefan Westerlund for patiently assisting me when my computer was not co-operating and to Laura Hoppmann for being my H1 astronomy accomplice.

I would also like to thank everyone at the UWA School of Physics, particularly Associate-Professor Paul Abbott for being a supportive Graduate Research Coordinator, Professor Ian McArthur for being a truly awesome Head of School and Jegatheva (Jay Jay) Jegathesan for being a charismatic School Manager.
I gratefully acknowledge funding support from ICRAR and from a UWA PhD Completion Scholarship.

UWA and ICRAR have also given me the opportunity to share my passion for astronomy and science through public education and outreach. An enormous thank you to ICRAR Outreach’s Pete Wheeler and Kirsten Gottschalk and to everyone from SPICE. These opportunities have vastly enriched my PhD experience, provided me with a worthy outlet and helped remind me why I love what I do.

Thanks to the University of Oxford sub-department of Astrophysics and the University of Hertfordshire Centre for Astrophysics Research for hosting me during my research exchange trips. In addition to those mentioned above, I’d like to give a special thanks to Dr David Bonfield, Dr Dan Smith, Jasmeer Virdee and Vanessa Ferraro-Wood. Thanks also to those who took me under their wing when I was alone in foreign countries - Dr Danny Price and Lindall Kidd, Dr Griffin Foster, Dr Jack Hickish, Dr Charles Copley, Moses Mogotsi, Dr Daniel Cumnara, Dr Richard Armstrong, Glenn and Fiona Pereira and my extended UK family. Apologies if I missed anyone here.

Thank you to the Astronomical Society of Australia (ASA) for providing many opportunities for collaboration and student learning and development. A particular thanks to Associate-Professor Andrew Hopkins for all your kind encouragement and advice, Professor Brian Boyle for your constructive comments and to all the ASA PhD students around Australia. I look forward to working with you in the future.

I would like to express my gratitude to the Galaxy and Mass Assembly (GAMA) collaboration. Being part of GAMA has opened up many new avenues for me, both in terms of data access and professional development. There are a few members who I would like to particularly thank for their help - Aaron Robotham, Lee Kelvin, Mehmet Alpaslan, Amanda Bauer and GAMA’s almighty leader Simon Driver.

Last but not least, thank you to all my amazing friends and family for your constant support over the years. Mum, Dad, Corin, Simon, Belinda, Amanda, David and Peran, thank you for everything. Sarah Wiseman, Lisa Laurendet, Pahnia Redgrave, Christiane Lynn, Tana Carley and Jonathon Carley thank you for inspiring me and giving me the encouragement I needed. You are all wonderful people. Special mentions go to Alyssa Nixon-Lloyd and Christopher Perrella for all those pep-talks and endless moral support. You were my guiding wisdom and motivation. And of course, thank you to Luke Bishop for all your love and patience and helping me deal with normal life on top of writing a thesis - and for waiving your consulting fee for your graphics and design assistance! I could never have done it without you.
Abstract

Large samples of galaxies studied over a range of redshifts are fundamental to understanding the various components and processes involved in galaxy evolution. However, large multiwavelength galaxy surveys are often restricted by the sensitivity or field-of-view of currently available telescopes. This thesis demonstrates how the innovative ‘stacking’ analysis technique can be used to overcome observational limits to allow studies of galaxy evolution over larger volumes than possible with direct detection methods.

Stacking is the process of combining the weak signals of many individual galaxies so as to increase the signal-to-noise ratio and allow a strong statistical detection. This thesis explores, develops and implements this technique to investigate (i) how the neutral atomic hydrogen gas (H\textsc{i}) component of galaxies varies with cosmic time, and (ii) how radio-loud active galaxies influence the interstellar medium and star formation of their host galaxies.

21\,cm observations taken with the Parkes 64\,m radio telescope are combined with the Two-Degree Field Galaxy Redshift Survey (2dFGRS) to trace H\textsc{i} in galaxies out to a redshift of $z = 0.13$. Stacking 21\,cm data at the positions of 15,093 2dFGRS galaxies at $z < 0.04$ results in a strong $31\sigma$ detection. Stacking 3,277 2dFGRS galaxies at $0.04 < z < 0.13$, each undetected at 21\,cm, also results in a clear $12\sigma$ detection. Using these statistical signals, the cosmic H\textsc{i} mass density ($\Omega_{\text{H}\textsc{i}}$) is found to be $(2.82^{+0.30}_{-0.59}) \times 10^{-4} \,h^{-1}$ and $(3.19^{+0.43}_{-0.59}) \times 10^{-4} \,h^{-1}$ over the lower and higher redshift ranges, respectively. This indicates no observed evolution in $\Omega_{\text{H}\textsc{i}}$ over the past $\sim 1 \,h^{-1}\text{Gyr}$, which is consistent with the results of direct detection methods. These measurements are more robust to the influence of cosmic variance due to the larger volume examined via stacking.

21\,cm Parkes observations are also collected over two 48\,deg$^2$ fields overlapping with the Galaxy and Mass Assembly (GAMA) redshift survey. 990 GAMA galaxies are examined at $z < 0.04$ and 4,898 at $0.04 < z < 0.11$. Severe radio frequency interference prevents analysis at higher redshifts. The well-characterised completeness of GAMA, along with the available multiwavelength information in this field, are used to minimise bias and develop more sophisticated methods of deriving $\Omega_{\text{H}\textsc{i}}$ through stacking techniques. H\textsc{i} mass-to-light ratios are found to be $\langle M_{\text{H}\textsc{i}}/L_r \rangle = 0.77 \pm 0.21 \,M_\odot/L_\odot$ at $z < 0.04$ and $0.16 \pm 0.02 \,M_\odot/L_\odot$ at $z > 0.04$. H\textsc{i}-to-stellar-mass ratios are found to be $\langle M_{\text{H}\textsc{i}}/M \rangle = 2.19 \pm 0.19$ and $0.15 \pm 0.01$.
over the lower and higher redshift ranges, respectively. Again, no compelling evidence for evolution in $\Omega_{H_1}$ is found over $0 < z < 0.1$.

Both the 2dFGRS and GAMA studies reveal that the stacking technique can be used to accurately study H$\text{I}$ at high redshifts with dramatically shorter integration times than normally required. No intrinsic limitation in the achievable signal-to-noise of the stacked detection is found, provided non-Gaussian noise components (such as radio frequency interference) are carefully excised. This indicates that high-redshift stacking analyses should be possible with the future generation of radio telescopes.

The stacking technique is applied to far-infrared imaging data provided by the Herschel Space Observatory Astrophysical Terahertz Large Area Survey (H-ATLAS) to examine the influence of radio-loud active galactic nuclei (AGN) activity on star formation. A complete sample of 250 radio galaxies with 1.4 GHz flux densities $> 10$ mJy is compiled from the NRAO VLA Sky Survey and the Faint Images of the Radio Sky at Twenty Centimeters source catalogues. Highly-sensitive, newly-available optical and near-infrared surveys are used to identify host galaxy positions and derive photometric redshifts for a deeper sample of galaxies than previously possible. Stacking the 250 $\mu$m H-ATLAS data at these source positions results in statistically-significant detections of dust in the galaxies out to $z = 5$. The far-infrared luminosity, and therefore the star formation rate, is found to increase with both redshift and 1.4 GHz luminosity over $0.05 < z < 5.0$. This continues trends seen in lower-redshift populations and is not simply a result of Malmquist bias since stacking overcomes the sensitivity limits of the data.

Average star formation rates as high as $200 M_\odot$ yr$^{-1}$ are detected, along with individual star formation rates of $> 500 M_\odot$ yr$^{-1}$ in $z > 2$ galaxies. These are likely to be radio-loud AGN hosting extreme star bursts. The availability of further sensitive, wide-field near-infrared/optical data will allow the construction of a high-redshift radio-quiet comparison sample. This will help to identify whether the observed far-infrared evolution is a general trend of the galaxy population as a whole, or whether the radio-loud activity is directly influencing the star formation.
Declaration

This thesis is composed of a series of published and in preparation papers, in compliance with the University of Western Australia’s regulations for PhD thesis submission. The co-authors and affiliations are listed at the start of each relevant chapter. I confirm that I have permission from the co-authors to reproduce the papers here. I declare that this thesis is my own composition, with contributions from other co-authors as stated below.

Publications arising from this thesis are:

   
   • I am the primary author on this paper. The co-authors and I shared Parkes observing responsibilities for this project, with assistance from those mentioned in the acknowledgements. The data reduction, analysis and composition of the paper is all my own work, with supervision and useful discussion provided by the co-authors. The 2dFGRS and HIPASS data products are publicly available and use of these has been explicitly stated in the paper.

2. **GAMA: Studying H I properties of galaxies to z = 0.1 with spectral stacking.** In preparation. Chapter 4 of this thesis.

   • I am the primary author on this paper. The co-authors and I shared Parkes observing responsibilities for this project, with assistance from those mentioned in the acknowledgements. The data reduction, analysis and composition of the paper is all my own work, with supervision and useful discussion provided by the co-authors. The use of GAMA team data products are explicitly stated in the paper.

• I am the primary author on this paper. The Herschel-ATLAS data stacking pipeline was developed by Hardcastle. The cross-matched catalogue between NVSS and FIRST radio sources was created by Virdee, which I extended slightly to cover all sources relevant to my work. Photometric redshifts were generated by Bonfield using the LEPHARE software, although I provided the input photometry. The use of Herschel-ATLAS and GAMA team data products, along with other required survey data, has been explicitly stated in the paper. All other catalogue compilation, data analysis and composition of the paper is my own work, with supervision and useful discussion provided by the co-authors.

Signed:

Candidate: ..........................................................

Supervisor: ..........................................................
# Contents

1 Introduction .............................................. 1
  1.1 Thesis overview ..................................... 1
  1.2 Galaxy Evolution ...................................... 3
    1.2.1 Overview of evolutionary processes .......... 3
    1.2.2 Observations of galaxy evolution and the role of stacking  .... 6
  1.3 Neutral hydrogen gas ................................... 12
    1.3.1 The role of $\text{H} \text{i}$ in galaxy evolution .... 12
    1.3.2 Tracing $\text{H} \text{i}$ ............................. 15
      1.3.2.1 Simulations .................................. 15
      1.3.2.2 Lyman $\alpha$ observations .................. 18
      1.3.2.3 $21\text{cm}$ emission line .................... 19
    1.3.3 $21\text{cm}$ surveys .............................. 21
      1.3.3.1 Surveys of the local Universe ............... 21
      1.3.3.2 Deep surveys .................................. 23
    1.3.4 Spectral stacking of the $21\text{cm}$ line .......... 27
    1.3.5 This work ........................................ 31
  1.4 Active galactic nuclei and radio galaxies .......... 33
    1.4.1 Overview of AGN and radio galaxies .......... 33
    1.4.2 The role of AGN in galaxy evolution .......... 36
    1.4.3 Observations of star formation in radio galaxies .... 41
    1.4.4 This work ........................................ 45

2 Stacking techniques and data quality considerations 47
  2.1 Overview of the stacking technique ................. 47
    2.1.1 Spectral stacking ................................ 48
    2.1.2 Image plane stacking ............................. 52
  2.2 RFI in $\text{H} \text{i}$ data ........................... 54
    2.2.1 Characteristics of RFI in Parkes data and mitigation strategies 54
    2.2.2 Impact of RFI on stacking analyses .............. 60
  2.3 Redshift error ....................................... 63

3 Detection of $\text{H} \text{i}$ in distant galaxies using spectral stacking 68
  Foreword and context ................................... 68
  Authors and affiliations ................................ 69
  3.1 Abstract ........................................... 69
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Introduction</td>
<td>70</td>
</tr>
<tr>
<td>3.3</td>
<td>The data</td>
<td>73</td>
</tr>
<tr>
<td>3.3.1</td>
<td>The 2dFGRS optical catalogue</td>
<td>73</td>
</tr>
<tr>
<td>3.3.2</td>
<td>H I observations and data</td>
<td>74</td>
</tr>
<tr>
<td>3.3.2.1</td>
<td>South Galactic Pole observations</td>
<td>74</td>
</tr>
<tr>
<td>3.3.2.2</td>
<td>HIPASS observations</td>
<td>75</td>
</tr>
<tr>
<td>3.3.2.3</td>
<td>Telescope scanning mode</td>
<td>75</td>
</tr>
<tr>
<td>3.3.2.4</td>
<td>Data reduction and interference mitigation</td>
<td>76</td>
</tr>
<tr>
<td>3.3.2.5</td>
<td>Gridding</td>
<td>77</td>
</tr>
<tr>
<td>3.4</td>
<td>Analysis</td>
<td>78</td>
</tr>
<tr>
<td>3.4.1</td>
<td>SGP source selection</td>
<td>78</td>
</tr>
<tr>
<td>3.4.2</td>
<td>HIPASS source selection</td>
<td>80</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Stacking process</td>
<td>80</td>
</tr>
<tr>
<td>3.4.3.1</td>
<td>Extracting H I spectra</td>
<td>80</td>
</tr>
<tr>
<td>3.4.3.2</td>
<td>Stacking spectra</td>
<td>82</td>
</tr>
<tr>
<td>3.4.3.3</td>
<td>Average H I mass</td>
<td>84</td>
</tr>
<tr>
<td>3.5</td>
<td>Results</td>
<td>85</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Noise behaviour</td>
<td>85</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Confusion</td>
<td>86</td>
</tr>
<tr>
<td>3.5.2.1</td>
<td>Line width estimation with the Tully-Fisher relation</td>
<td>87</td>
</tr>
<tr>
<td>3.5.2.2</td>
<td>Identification of confused sources</td>
<td>88</td>
</tr>
<tr>
<td>3.5.2.3</td>
<td>Luminosity correction</td>
<td>89</td>
</tr>
<tr>
<td>3.5.3</td>
<td>H I mass density</td>
<td>90</td>
</tr>
<tr>
<td>3.5.4</td>
<td>Cosmic variance</td>
<td>95</td>
</tr>
<tr>
<td>3.5.5</td>
<td>Binning</td>
<td>96</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>4</td>
<td>GAMA: Studying H I properties of galaxies to ( z = 0.1 ) with spectral stacking</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Foreword and context</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Authors and affiliations</td>
<td>106</td>
</tr>
<tr>
<td>4.1</td>
<td>Abstract</td>
<td>106</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction</td>
<td>107</td>
</tr>
<tr>
<td>4.3</td>
<td>The data</td>
<td>110</td>
</tr>
<tr>
<td>4.3.1</td>
<td>GAMA</td>
<td>110</td>
</tr>
<tr>
<td>4.3.2</td>
<td>21 cm observations</td>
<td>111</td>
</tr>
<tr>
<td>4.4</td>
<td>Analysis</td>
<td>114</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Source selection</td>
<td>114</td>
</tr>
<tr>
<td>4.4.1.1</td>
<td>Selection based on redshift and position</td>
<td>114</td>
</tr>
<tr>
<td>4.4.1.2</td>
<td>Completeness and sample biases</td>
<td>115</td>
</tr>
<tr>
<td>4.4.1.3</td>
<td>Summary of source selection</td>
<td>118</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Stacking H I spectra</td>
<td>120</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Correction for confusion</td>
<td>125</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Cosmic H I mass density</td>
<td>130</td>
</tr>
<tr>
<td>4.5</td>
<td>Discussion of results</td>
<td>135</td>
</tr>
</tbody>
</table>
5 Herschel-ATLAS/GAMA: Infrared properties of bright radio galaxies in the science demonstration phase field 139
5.1 Abstract ........................................... 140
5.2 Introduction ........................................ 141
5.3 The data ............................................ 146
  5.3.1 1.4GHz radio data ............................... 146
  5.3.2 Near-infrared counterparts ....................... 148
  5.3.3 Optical counterparts ............................. 149
  5.3.4 Obtaining redshift information .................. 149
    5.3.4.1 Spectroscopic redshifts ...................... 149
    5.3.4.2 Photometric redshifts ....................... 150
    5.3.4.3 Photometric redshift quality ............... 152
  5.3.5 Far-infrared counterparts and photometry ...... 155
  5.3.6 Quasar identification .......................... 156
5.4 Analysis ........................................... 159
  5.4.1 FIR luminosity and $K$-correction ............... 159
  5.4.2 Far-infrared stacking .......................... 161
  5.4.3 Temperature fitting within bins ................ 167
  5.4.4 Examining possible biases ........................ 169
5.5 Discussion of results ................................ 171
  5.5.1 Comparison of star formation rates ............. 171
  5.5.2 Comparison to the low-redshift population ...... 174
5.6 Conclusions ....................................... 177

6 Summary and future work 180
  6.1 Thesis summary ................................... 180
  6.2 Future work ..................................... 185

References 189

A List of Acronyms 204

B Luminosity and stellar mass density functions and the completeness extrapolation factor 208
  B.1 Luminosity density ............................... 208
  B.2 Stellar mass density .............................. 210

C Radio galaxy catalogue 212

D Images of radio galaxies 222
List of Figures

1.1 A flow chart of galaxy formation and evolution processes ............. 4
1.2 The spectral energy distribution of a typical galaxy .................. 7
1.3 The position of various survey fields on the sky ....................... 9
1.4 The distribution of 2dFGRS galaxies over the sky, as a function of redshift. ......................................................... 10
1.5 A compilation of observational measurements and fits to the star-formation rate density history ................................. 11
1.6 The evolution of the cosmic $\mathrm{H\,i}$ mass density, stellar mass density and star formation rate density ............................. 14
1.7 Comparison of predicted evolution of cold gas density with redshift 16
1.8 The double-horned 21 cm line of the galaxy HIPASSJ0344-44. .... 20
1.9 The HIPASS $\mathrm{H\,i}$ mass function .................................. 22
1.10 The stacked 21 cm spectrum of 121 $\mathrm{H\,\alpha}$ emitting galaxies .... 30
1.11 A unified model of AGN ................................................ 34
1.12 The cosmic evolution of the star formation rate and the black hole accretion rate ................................................. 37
1.13 Observational measurements of the black hole and galaxy bulge masses 38
1.14 The observed and simulated luminosity function ....................... 40

2.1 A representation of a radio data cube. .................................. 49
2.2 An example of five extracted and aligned $\mathrm{H\,i}$ spectra. ............. 50
2.3 A series of co-added spectra ........................................... 51
2.4 A conceptual representation of the image-plane stacking process. .. 53
2.5 Example spectra overlaid to show the RFI properties and spectral features ................................................................. 55
2.6 Example data from Parkes 2008-2009 observations in the 1335 and 1285 MHz bands. ...................................................... 57
2.7 Example original and flagged data from Parkes 2010 and 2012 observations in the 1285 MHz band. ................................. 59
2.8 The fraction of data masked by AOFlagger due to RFI presence .... 60
2.9 The rms noise level of the co-added spectrum when $N$ galaxies are stacked together, with and without sufficient RFI excision. ..... 62
2.10 An example of a co-added $\mathrm{H\,i}$ spectrum prior to baseline fitting and removal. ......................................................... 63
### LIST OF FIGURES

2.11 Comparison of stacked spectra using spectroscopic versus photometric redshifts. ............................................. 64
2.12 The stack of 1,000 randomly-selected galaxies from the HIPASS source catalogue as various levels of error are introduced to the redshifts and the integrated flux in each case. ................................. 66

3.1 RFI occupancy per channel ........................................... 77
3.2 The heliocentric redshift distributions of the 2dFGRS galaxies .... 79
3.3 The 21 cm flux density spectrum of the only directly detected 2dFGRS source in the SGP field ........................................ 81
3.4 The co-added H I mass spectrum of all 3,277 SGP sources and 15,093 2dFGRS sources covered by HIPASS .............................. 84
3.5 The rms noise level of the co-added spectrum versus the number of individual spectra contributing to the stack ........................ 87
3.6 H I line widths estimated using the Tully-Fisher relation for the sample of SGP and HIPASS galaxies. ................................. 89
3.7 The luminosity distribution of all stacked optical sources in the SGP field and HIPASS field ........................................ 91
3.8 Observational constraints on the cosmic H I mass density $\Omega_{\text{HI}}$ as a function of look-back time. ................................. 95
3.9 The average H I masses and average mass-to-light ratios derived by co-adding all spectra in bins of redshift, $b_J$ luminosity and $(b_J - r_F)$ colour .................................................. 97

4.1 The percentage of data flagged due to RFI presence per spectral channel in the G09 and G15 Parkes data .......................... 111
4.2 The rms per spectral channel for the gridded G09 and G15 Parkes data cubes ......................................................... 112
4.3 The $r$ band luminosity and stellar mass versus redshift for all G09 and G15 sources available for stacking ............................ 116
4.4 Distribution of sources in the G09 and G15 fields in both the low and high redshift $L_r$ complete samples .......................... 117
4.5 The redshift distribution of all G09 and G15 sources suitable for stacking ................................................................. 119
4.6 The co-added spectrum produced by stacking H I mass, H I mass-to-light and H I-to-stellar mass spectra for low and high redshift galaxies ................................................................. 122
4.7 The H I profile widths of individual GAMA galaxies, as predicted by the Tully-Fisher relation ........................................ 126
4.8 The correction factors applied in the confusion correction ........ 128
4.9 The luminosity and stellar mass distributions of all stacked sources before and after correction for confusion ..................... 130
4.10 Average H I mass-to-light and H I-to-stellar mass ratios in luminosity and stellar mass bins ........................................... 133
4.11 The cosmic H I mass density versus redshift ............................ 134

5.1 Examples of a radio galaxies with complex structure .................. 147
5.2 The distribution of all spectroscopic redshifts available, all generated photometric redshifts and the final redshift distribution adopted for this analysis .......................................................... 150
5.3 The $\chi^2$ of the photometric template fit versus the fractional error in the photometric redshift, and the fractional redshift error versus $K_s$-band magnitude ....................................................... 153
5.4 Comparison of GAMA spectroscopic redshifts with those predicted by the $K$-$z$ relation and photometric redshifts generated with LEPHARE . 154
5.5 The photometric and available spectroscopic redshifts versus $K$-band magnitude. .............................................................. 154
5.6 Near-infrared colours of the sample of radio galaxies, as a diagnostic of quasar presence ......................................................... 158
5.7 Upper limits on the possible extent of quasar contamination at 250 $\mu$m versus the measured 250 $\mu$m flux from H-ATLAS. .................. 159
5.8 The co-added 250 $\mu$m SPIRE images of all 250 radio sources, only the 239 sources not in the H-ATLAS catalogue and 239 random positions 163
5.9 The stacked 250 $\mu$m SPIRE images within bins of 1.4 GHz luminosity and redshift. ......................................................... 164
5.10 $L_{250}$ versus 1.4-GHz radio luminosity and redshift with individual points calculated using the isothermal model and temperature fitting to each bin ................................................................................... 165
5.11 1.4 GHz luminosity vs redshift of strong radio sources .............. 166
5.12 Comparison between the stacked $L_{250}$ using various sample cuts in bins of $L_{1.4}$ and $z$ .............................................. 172
5.13 Comparison of the stacked $L_{250}$ in bins of $L_{1.4}$ and $z$ with and without the introduction of random redshift error ..................... 173
5.14 Comparison between the stacked $L_{250}$ in bins of $L_{1.4}$ and $z$ using the source samples of this work and that of V13 ................. 174
Chapter 1

Introduction

1.1 Thesis overview

One of the principle aims of modern cosmology is to achieve a comprehensive understanding of galaxy formation and evolution. Galaxies are intricate systems of gas, dust, stars and dark matter with non-linear and often poorly-understood interactions between these constituents. Many physical properties of the galaxy population have been observed to change over time. For example, the rate of star-formation within galaxies has changed drastically over the past $\sim 10$ billion years. The challenge is to procure a more detailed understanding of the various physical and chemical processes responsible for such evolution.

To probe the evolution of galaxy components and to better understand the processes involved, sensitive measurements are required out to large redshift. Limitations in the current generation of telescopes restricts many observations to the relatively nearby Universe. This thesis aims to extract information at the limits of available and new data sets using the ‘stacking’ analysis technique. Stacking is the process of combining many low signal-to-noise observations of different individual objects in order to retrieve a high-significance statistical detection. This enables studies of the average properties of the distant galaxy population.
The total amount of neutral atomic hydrogen gas ($\text{H}_\text{i}$) in the Universe, the fuel for star formation, has changed over cosmic time. It constituted the dominant baryonic component after the Epoch of Recombination but represents a much smaller fraction of the present Universe. However, it is unclear how exactly the cosmic $\text{H}_\text{i}$ density varies with redshift and how this relates to galaxy evolution. By applying spectral stacking to the 21 cm hyperfine emission line of $\text{H}_\text{i}$, evolution in the gas properties of galaxies can be examined.

Active galactic nuclei (AGN) activity is a powerful phenomenon that is expected to severely disrupt star formation in a galaxy through various feedback and quenching mechanisms, particularly in the radio-loud case. By applying image-plane stacking to deep far-infrared observations, it is possible to study the dust properties, and by proxy the star formation rate, of distant and/or faint host galaxies containing AGN. Thus, the influence of AGN activity on galaxy formation and evolution can be explored.

In this introductory chapter, section 1.2 briefly summarises the current theoretical and observational framework of galaxy evolution. Section 1.3 discusses the role of $\text{H}_\text{i}$ in galaxy evolution processes, describes the different methods used to trace $\text{H}_\text{i}$ content and explains the importance of spectral stacking to push observational boundaries. Section 1.4 considers the role of AGN feedback within the context of galaxy formation and evolution, the influence of radio-loud AGN activity on star formation rates and how stacking of far-infrared imaging can be used to explore this.

Chapter 2 provides a detailed description of the stacking technique applied to both the spectral and image planes. It also examines how poor data quality, namely radio frequency interference (RFI) and redshift error, can impact the results of a stacking analysis and how this can be prevented. Chapter 3 contains the first $\text{H}_\text{i}$ stacking analysis which utilises data collected from the Parkes radio telescope and optical information from the Two-degree-field Galaxy Redshift Survey (2dFGRS). Chapter 4 presents the second stacking analysis of 21 cm data from Parkes, with optical information provided by the Galaxy and Mass Assembly (GAMA) redshift survey. Chapter 5 presents an image plane stacking analysis of far-infrared Herschel
data to examine the star forming properties of strong radio galaxies. Finally, Chapter 6 summarises the conclusions and describes future work.

1.2 Galaxy Evolution

1.2.1 Overview of evolutionary processes

How do galaxies form and evolve? Answering this question presents one of the greatest challenges in modern astronomy and cosmology. Galaxies are complex, multi-component systems which have undergone significant changes since their formation after the Big Bang. This section will provide a brief outline of the current framework of galaxy formation and evolution. It is based on the background cosmology and semi-analytic formalism as described in Cole et al. (2000) and subsequently reviewed in e.g. Baugh (2006), Benson (2010) and Mo et al. (2010). Figure 1.1 summarises the relationships between the key ingredients.

Variations in the matter distribution of the Universe originated from quantum fluctuations in the very early Universe which were seeded and rapidly grown during inflation. Gravitational instability led over-dense regions to collapse, resulting in the formation of dark matter halos. Based on the hierarchical picture of structure formation, these halos merge to form more massive systems (e.g. Springel et al., 2005). Gas falling in to these gravitational potential wells shock-heats to the virial temperature (e.g. Binney, 1977) and then cools, at which point gas flows towards the centre of the potential well and accretes onto the ‘proto-galaxy’ (e.g. White & Frenk, 1991). If residual angular momentum is present, the gas will settle naturally into a disc (e.g. Fall & Efstathiou, 1980).

If cooling is efficient, gas can fragment into many smaller regions of high density (Maller & Bullock, 2004). It is within these molecular clouds that the pressure and
Figure 1.1: A flow chart of key galaxy formation and evolution processes. Reproduced from Figure 7 in Baugh (2006) (adapted from Cole et al., 2000).
temperature can increase enough to allow the formation of stars. Metals are synthesised within stars and can be ejected into the interstellar medium (ISM). The changing metallicity of a galaxy can influence the cooling efficiency of the gas, alter the observed colour of the galaxy and attenuate light via absorption by dust grains (Cole et al., 2000). The rate at which the metallicity of a galaxy changes will depend heavily on the initial mass function (IMF; the mass distribution of newly formed stars) and the star formation efficiency of the galaxy (e.g. Bastian et al., 2010). The most massive stars will die in supernovae explosions which will not only enrich the ISM but can also heat, compress and/or blow gas out of the galaxy in a ‘galactic wind’. These various feedback processes can both inhibit and enhance star formation in different parts of the galaxy and on different time scales (Efstathiou, 2000).

Mergers of multiple galaxies can send stars on disordered trajectories, severely disrupting the morphology and may result in a spheroidal shape or a bulge component (e.g. Bournaud et al., 2005). During gas rich mergers, the gas can be compressed and produce a rapid increase in the star formation rate of the galaxy, referred to as a ‘starburst’ phase (e.g. Mihos & Hernquist, 1996). The merger may also trigger the accretion of matter onto a supermassive black hole (SMBH) that often resides in the centre of galaxies (Di Matteo et al., 2005). Accretion releases ultraviolet (UV) and X-rays and powers jets in what is collectively referred to as active galactic nuclei (AGN) activity. This radiation can produce strong feedback mechanisms which may have a significant effect on the star formation rate of the galaxy. For example, star formation may be quenched by the rapid transfer of gas from the galaxy into the intergalactic medium (IGM) via outflow processes (e.g. Bower et al., 2006; Croton et al., 2006).

Many of the astrophysical processes described above are still poorly understood. While both observational and theoretical studies of galaxy evolution are progressing at a rapid rate, many questions still remain. This thesis will provide further observational constraints to assist our overall knowledge of galaxy evolution.
1.2.2 Observations of galaxy evolution and the role of stacking

Each component of a galaxy (stars, dust, neutral and molecular gas, accreting black holes etc.) plays a distinct role in galaxy evolution processes, as described above. Observations of each baryonic constituent therefore provide a unique perspective on the evolutionary history and contribute important constraints to galaxy formation models. Each component is best traced at different wavelengths due to their different emission properties, as illustrated in Figure 1.2. For example, stars emit predominantly in the optical and near-infrared, molecular emission lines can be detected in the mid-infrared and submillimetre, dust grains are heated by stars and re-radiate in the far-infrared, and neutral hydrogen produces an emission line in the radio (at 21 cm). Compact objects (such as black holes and pulsars) can be studied in the higher energy gamma, X-ray and UV regimes, while synchrotron emission from radio-loud AGN activity can be detected in radio continuum. Therefore, obtaining a unified picture of galaxy evolution requires complementary surveys from the radio to the far-UV and beyond.

It is for precisely this purpose that campaigns such as the Galaxy and Mass Assembly (GAMA) survey are important. GAMA aims to provide a repository of multiwavelength information for thousands of galaxies within three 48°2 equatorial fields (Driver et al., 2009, 2011), shown in red in Figure 1.3. The GAMA redshift survey has so far collected optical spectroscopic redshifts for more than 250,000 galaxies in these regions with the AAOmega spectrograph on the Anglo-Australian Telescope (AAT). Using the Herschel Space Observatory, launched in 2009, the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS) provides the complementary far-infrared data for GAMA (Eales et al., 2010). Initial H-ATLAS observations of a ∼14°2 science demonstration phase (SDP) field, shown in green in Figure 1.3, were conducted in 2010 and have recently been extended to cover all three GAMA fields. Near-infrared imaging data is provided by the Visible and Infrared Survey Telescope for Astronomy (VISTA) Kilo-degree Infrared Galaxy survey (VIKING; Sutherland et al., in prep; Fleuren et al., 2012) which takes advantage of the excellent survey
Figure 1.2: The spectral energy distribution of an example galaxy, showing the components dominating at each wavelength. The instruments contributing data to the GAMA survey in each regime are labelled. Image by Simon Driver for the GAMA team.
capabilities of the 4 m VISTA telescope in Chile.

Other ancillary imaging data in the radio continuum, mid-infrared, optical and UV have been (or will shortly be) provided by surveys with the Giant Metre-wave Radio Telescope (GMRT; Mauch et al., 2013), Wide-field Infrared Survey Explorer (WISE; Wright et al., 2010), VLT Survey Telescope (VST; de Jong et al., 2013) and the Galaxy Evolution Explorer (GALEX; Martin et al., 2005), respectively. Deep radio spectral line information will be provided by the Deep Investigation of Neutral Gas Origins (DINGO; Meyer, 2009) survey with the Australian SKA Pathfinder (ASKAP) which is currently being constructed. By exploiting these cutting edge instruments, the GAMA team have created an unparalleled homogenous, multiwavelength data set which is ideal for observational studies of galaxy evolution. GAMA, H-ATLAS and VIKING are central to the galaxy evolution analyses carried out in this thesis (see Chapters 4 and 5).

In addition to covering many wavelength regimes, extragalactic surveys must also aim to detect a large number of galaxies in a wide range of environments. This ensures that the observed sample is representative of the entire galaxy population and therefore facilitates unbiased statistical analyses of evolutionary properties. This is best achieved by surveying the widest sky area possible. The Two-degree Field (2dF) Galaxy Redshift Survey (2dFGRS) is an excellent example of a wide-field optical spectroscopic survey that has made significant contributions to the field of galaxy evolution. The 2dFGRS was conducted with the 2dF multifibre spectrograph on the Anglo-Australian telescope (Colless et al., 2001). It covers two large strips, one in the northern sky and one in the south, along with a number of random 2 deg² fields. These regions are shown in yellow in Figure 1.3 and altogether cover \( \sim 2,000 \text{ deg}^2 \). By determining accurate spectroscopic redshifts for more than 250,000 galaxies over these large areas, the 2dFGRS revealed the extent of galaxy clustering. Figure 1.4 shows the large-scale filamentary ‘cosmic-web’ structure in the matter distribution of the Universe. This clustering has important implications for cosmological and galaxy formation models (Peacock et al., 2001; Peacock, 2002). Furthermore, this highlights
Figure 1.3: The position of various galaxy survey fields on the sky shown with equatorial coordinates in SIN projection. Yellow indicates regions covered by the 2dFGRS spectroscopic survey. This full region, and the subset near the South Galactic Pole (SGP) shown in blue, are used in the analysis of Chapter 3. The red GAMA fields at 9\textsuperscript{h} and 14.5\textsuperscript{h} (G09 and G15, respectively) are the focus of Chapter 4. The green H-ATLAS SDP field is used in Chapter 5.
the range of possible galaxy environments - from voids to clusters. If a measurement is made over a small sky volume, it may be prejudiced by the particular galaxy distribution within that volume (known as cosmic variance bias). Large-volume surveys will minimise this cosmic variance bias since they will sample galaxies over a wide range of environments. Thus, the wide-field 2dFGRS is an excellent tool for statistical analyses of galaxy evolution and has in fact facilitated the work presented in Chapter 3 of this thesis.

To increase the volume probed by a survey, and therefore to reduce cosmic variance bias, observations must not only be wide-field but also cover a wide redshift range. High redshift observations are vital for galaxy evolution studies since they probe the Universe at an earlier stage and are therefore required to track the evolution of various galaxy properties. For example, multiwavelength observations over the range \(0 < z < 6\) have shown that the star formation rate (SFR) of galaxies has varied significantly over cosmic time (e.g. Madau et al., 1996 and Lilly et al., 1996). Hopkins (2004), Hopkins & Beacom (2006) and, very recently, Behroozi et al. (2013), compiled observational measurements of the SFR density history. This is the rate at which gas is converted into stars per unit volume as a function of redshift. As seen
in Figure 1.5, this revealed that the SFR of the Universe peaked at around $z \sim 2 - 3$ and has decreased by an order of magnitude over the past $\sim 10$ billion years. While this places vital constraints on galaxy evolution models, explaining this trend is not a simple matter since many different factors influence the SFR of a galaxy, as discussed previously.

To obtain a complete picture of galaxy evolution, the overall challenge is to conduct multiwavelength surveys that probe large volumes by covering both wide spatial fields and large redshift ranges. Since a given galaxy will have a fainter apparent magnitude if it is placed at higher redshifts, it is crucial for surveys to be highly sensitive in order to capably detect distant systems. However, due to limitations in telescope capabilities and available observing time, there is often a compromise between sensitivity and sky coverage. For example, the 2dFGRS and GAMA are both magnitude-limited spectroscopic surveys of comparable numbers of galaxies. While the 2dFGRS covers a
1.3. NEUTRAL HYDROGEN GAS  

field $\sim$14 times wider than GAMA, the GAMA survey extends 1.1 magnitudes deeper. Thus the median redshift of GAMA is $\sim$0.21, which is considerably higher than the median redshift of the 2dFGRS, which is $\sim$0.11. Thus, each survey is suitable for probing a different region of parameter space.

Despite recent technological advancements that allow simultaneously more sensitive and wider field surveys than ever before, limitations and trade-offs still exist. This is where the innovative stacking technique, the focus of this thesis, can help. As will be demonstrated below, stacking is a powerful tool for studies of galaxy evolution and has the potential to unveil previously unchartered epochs in the Universe’s history. Rather than requiring direct, individual detections, stacking is the process of making high signal-to-noise (S/N) statistical detections of an ensemble of galaxies. Therefore, this analysis technique can be used to extract information from the high redshift limits of surveys and thereby increases the effective volume. This thesis will demonstrate how stacking can be used to push detection limits of the 21cm line of neutral hydrogen gas, and far-infrared emission from dust, to facilitate studies of important galaxy evolution trends.

1.3 Neutral hydrogen gas

1.3.1 The role of H\textsc{i} in galaxy evolution

Cool gas ($< 10^4$K) is a fundamental component in the life of all galaxies. A large fraction of this gas is in the form of neutral atomic hydrogen (H\textsc{i}). As outlined in section 1.2.1 above, this gas is the primary building block of galaxies as it accretes within galactic halos, condenses into giant molecular (H$_2$) clouds and thereby fuels star formation. Hence, it is the H\textsc{i} which essentially keeps a galaxy ‘alive’ and generating new stars rather than simply evolving quiescently. The presence of H\textsc{i} is
essential to trigger a starburst phase, particularly as a result of a galaxy merger. It is also thought to play an important role in triggering AGN activity, which may also influence star formation (see section 1.4.2 below).

As discussed above, the observed SFR density of the Universe varies significantly over cosmic time and peaks at $z \sim 2–3$ (Hopkins et al., 2006; see shaded region of Figure 1.6). Correspondingly, the stellar mass density of the Universe has built up over cosmic time (e.g. Wilkins et al., 2008; see solid line in Figure 1.6). Since star formation is closely linked to the availability of HI, to explain this observed trend we must also understand how the cosmic mass density of neutral hydrogen gas ($\Omega_{\text{HI}}$) evolves with redshift. HI was the dominant baryonic component in the early Universe but currently comprises less than 1 per cent of the baryon budget (Prochaska & Tumlinson, 2009), with a much larger fraction locked up in stars (as seen in Figure 1.6 at $z \approx 0$). Therefore, the HI density of the Universe, taken here to mean the density within galaxies, must have evolved significantly with cosmic time. Tracing this evolution is now the challenge. Figure 1.6 compares the evolution of the star formation rate density (Hopkins & Beacom, 2006), stellar mass density (Wilkins et al., 2008) and some currently available simulated and observed constraints on $\Omega_{\text{HI}}$, which will be described in the following sections.

Furthermore, the neutral gas component of galaxies varies not only with cosmic epoch, but also with many physical properties such as environment density, galactic morphology and colour. For example, galaxies in clusters are more HI deficient compared to galaxies in the field (e.g. Haynes et al., 1984) and blue, late type galaxies predominantly contain more HI than redder, late type systems (e.g. Spitzak & Schneider, 1998). Moreover, galaxy clusters above $z \approx 0.2$ appear to contain more blue and spiral galaxies than their counterparts in the local Universe, referred to as the Butcher-Oemler effect (Butcher & Oemler, 1978; Couch et al., 1994; Wirth et al., 1994). There is debate over whether this is mostly due to stripping mechanisms, or due to the unrelated evolution of the field galaxies which then accrete onto the cluster.

Tracing the HI component of galaxies will provide useful insights into the internal
Figure 1.6: The evolution of the cosmic H\text{I} mass density (\(\Omega_{\text{H} \text{I}}\)). Observational measurements are shown as points and indicated in the legend. Circular points have been derived using 21 cm emission line observations, squares with damped Lyman \(\alpha\) measurements and diamonds with intensity mapping techniques. The black dashes at \(z = 0.8\) show the uncertainty on the intensity mapping values due to the unknown value of the bias parameter. The Lagos et al. (2011) semi-analytical prediction of \(\Omega_{\text{H} \text{I}}\) evolution is shown as the dashed line. The solid black line also corresponds to values on the left y-axis and shows the evolution of the cosmic stellar mass density, as determined by Wilkins et al. (2008). The grey shading, corresponding to values on the right y-axis, shows the 1 \(\sigma\) confidence region of the evolution of the cosmic star-formation rate density (\(\Omega_{\text{SFR}}\)) established by Hopkins & Beacom (2006).
physical processes at work and the impact on the evolutionary pathway of a galaxy. Thus, it is vital to examine the H\textsc{i} content of galaxies over cosmic time and within a varied sample of galaxies, in order to obtain a complete picture of galaxy evolution. However, as we will see in the following section, simulations and observations of H\textsc{i} both face serious limitations and little is known about this ingredient beyond the local Universe.

1.3.2 Tracing H\textsc{i}

1.3.2.1 Simulations

Simulations can be used to predict the evolution of the cold gas component of galaxies. The two main approaches are to use either semi-analytical models (SAMs) or hydrodynamical simulations, each of which have their strengths and weaknesses. Semi-analytical models (SAMs) use a dark matter halo merger tree framework such as the Millennium Simulation (Springel et al., 2005) and implement a galaxy formation prescription over the top of this (such as that described in section 1.2.1 above). Different semi-analytical ‘recipes’ are used to approximate the various astrophysical processes. These recipes can be varied, allowing searches of a wide parameter space in a computationally efficient manner.

Power et al. (2010) compared the predictions for gas evolution using SAM realisations by Baugh et al. (2005), Bower et al. (2006), De Lucia & Blaizot (2007) and Font et al. (2008). These differ in their treatment of aspects such as AGN feedback and star formation time scales. Despite their differences, these models all reproduce the $z = 0$ H\textsc{i} mass function (HIMF; discussed in section 1.3.3.1 below) relatively closely, down to the resolution threshold. However, they predict quite different evolution of the cold gas density with redshift, as seen in Figure 1.7. This shows that while different galaxy formation models can reproduce $z \approx 0$ observational constraints, they can diverge widely in their evolutionary predictions. Knowing which prescription most
1.3. NEUTRAL HYDROGEN GAS

Figure 1.7: Comparison of the predicted evolution of cold gas density with redshift. The lines show simulated results using galaxy formation models by Bower et al. (2006), De Lucia & Blaizot (2007), Font et al. (2008) and Baugh et al. (2005). Individual symbols show available observational constraints. Reproduced from Figure 2 in Power et al. (2010).

accurately predicts the H\textsc{i} properties of the real Universe is therefore difficult, as is determining the relative contribution of each physical ingredient to the evolution.

Resolution limitations can also contribute ambiguity when interpreting results of simulations. All realisations considered by Power et al. significantly underestimated the HIMF below $M_{\text{cold}} \approx 10^{8.5} \, M_\odot \, h^{-1}$. Since this roughly corresponds to the mass resolution limit in the underlying halo merger tree framework, it was unclear whether the underestimation was an artefact of poor resolution, or whether there was a problem with the physical prescription. Using a higher resolution simulation, Lagos et al. (2011) found that this discrepancy was actually caused by an insufficient decomposition of the cold gas into the neutral and molecular constituents. By invoking a self-consistent method of separating the two components within the simulation itself, Lagos et al. found that low mass galaxies lack molecular gas, preventing star-formation and allowing their atomic content to remain constant. Therefore, Lagos et al. find a larger number of low H\textsc{i} mass galaxies than in Power et al. Using this decompo-
CHAPTER 1. INTRODUCTION

1.3. NEUTRAL HYDROGEN GAS

sition, Lagos et al. examined the evolution of the cosmic atomic and molecular H\textsc{i} gas densities separately. As shown by the dashed line in Figure 1.6 above, they find (relatively) good agreement with $z = 0$ observations of $\rho_{\text{H}1}$ and with higher redshift DLA measurements (discussed in the next section).

Unlike SAMs, cosmological hydrodynamical simulations have the ability to directly model gas dynamics and have thereby shown that accretion and outflow processes play important roles in determining overall H\textsc{i} contents. For example, smoothed particle hydrodynamics (SPH) simulations have shown that the filamentary inflow of gas at temperatures cooler than the virial temperature (i.e. cool mode accretion) contributes a significant fraction of the gas in galaxies, particularly for lower mass systems (e.g. Kereš et al., 2005). Davé et al. (2013) applied various wind models to a hydrodynamical framework to explore the influence of outflows on galaxy H\textsc{i} content. They found that the HIMF is strongly influenced by the outflows, overestimating the low mass end and producing too few high mass galaxies if excluded.

Once again, resolution is a major issue for these simulations. Due to limited processing power, there is often a compromise between particle (or grid) resolution and box size in hydrodynamical simulations. This is analogous to the trade-off between sensitivity and volume in observational surveys. While, in principle, hydrodynamical simulations can numerically track the evolution of all components, in practice this is restricted by the computational resolution and processing power. As a consequence, prescriptions similar to those used for SAMs must be adopted for various aspects including supernovae and AGN feedback, such as that invoked by Davé et al. (2013).

Another challenge faced by hydrodynamical simulations is the accurate estimation of the neutral content of the simulated gas (see Altay et al., 2011; Duffy et al., 2012) since this is not self-consistently calculated. This too can affect the predictions for the evolution of neutral gas in galaxies.

Both SAMs and hydrodynamical simulations are limited by uncertainty in the physics of various processes and by the considerable processing power required to track gas over a wide spatial and temporal range. Therefore, further observational constraints
1.3. NEUTRAL HYDROGEN GAS

on H\textsc{i} physics are crucial for improving simulations and therefore our overall understanding of galaxy formation and evolution.

### 1.3.2.2 Lyman \(\alpha\) observations

At high redshifts, cosmic H\textsc{i} densities can be probed observationally using damped Lyman \(\alpha\) (DLA) systems. DLAs are reservoirs of predominantly neutral gas with H\textsc{i} column densities greater than \(2 \times 10^{20}\) cm\(^{-2}\). They can be identified via the broad Ly\(\alpha\) absorption signature over the top of the Ly\(\alpha\) forest in quasi-stellar object (QSO) spectra. DLAs are the main contributor to \(\Omega_{\text{HI}}\) of all the known absorbing systems (e.g. Ly\(\alpha\) forest lines, Lyman-Limit Systems etc.) As such, the H\textsc{i} mass density can be derived from the DLA distribution function - the number of absorbers per unit H\textsc{i} column density, per unit absorption distance. By examining 738 DLA systems identified in Sloan Digital Sky Survey (SDSS) spectra, Prochaska et al. (2005) and Prochaska & Wolfe (2009) found that the cosmic H\textsc{i} density of DLA systems (\(\Omega_{\text{HI}}^{\text{DLA}}\)) decreases by around 50 per cent over \(2.2 < z < 5.5\). Along with lower redshift results (see below), this seems to suggest that there has been almost no evolution in gas densities of galaxies over the past \(\sim 10\) Gyr but that there was a period of rapid evolution over the \(\sim 2\) Gyr prior to this. However, various systematic errors may be contributing appreciable uncertainties to their measurements. For example, gravitational lensing can magnify background quasars and bias the magnitude-complete sample. Underestimates of the derived gas mass density may also be introduced by dust obscuration.

Although effective at high redshift, tracing H\textsc{i} via the Ly\(\alpha\) damping method becomes difficult at intermediate redshifts. Below \(z \approx 1.65\), the Ly\(\alpha\) line shifts from the optical into the UV and therefore space-based campaigns become the only option. The small number density of DLAs at low redshift mean that these studies require significant amounts of observing time, which is impractical for such instruments. Rao et al. (2006) probed neutral gas densities at \(z < 1.65\) using the Hubble Space Telescope (HST)
by targeting 197 pre-selected systems with strong Mg II and Fe II absorption along sight lines to SDSS-selected QSOs. These metal lines are known indicators of high neutral gas column density, and therefore of possible DLAs. They confirmed that 42 of these absorbers are DLAs using the HST UV spectra. They measured H I column densities for all objects in their sample and found no statistically-significant evolution of $\Omega_{\text{DLA}}^{\text{H I}}$ over $0.5 < z < 5.0$, in apparent contradiction to the results of Prochaska & Wolfe (2009). However, the Rao et al. (2006) measurements are also associated with significant uncertainty since they rely on assumptions of the DLA fraction of Mg II absorption systems and suffer somewhat from low number statistics.

1.3.2.3 21 cm emission line

At low redshifts ($z \approx 0$), H I in galaxies can be detected in emission via the 21 cm line. Within the ground-state neutral hydrogen atom, the quantum spins of the electron and proton can either be parallel or antiparallel. This results in two possible hyperfine states, with the energy of the parallel state $6 \times 10^{-6}$ eV greater than the antiparallel state. A transition of the electron from the higher to the lower energy state, often referred to as a ‘spin flip’, corresponds to the emission of a photon of frequency $1420.4058$ MHz, wavelength $21.105$ cm (in the radio regime). This transition is ‘forbidden’, with a mean lifetime in the parallel state of $1.1 \times 10^7$ yr. However, collisions between atoms (which are much more frequent than the spontaneous emission) will maintain the relative populations of the two energy levels. Therefore, collisions will keep the spin temperature in equilibrium with the kinetic temperature. Using Einstein’s relations for atomic transitions in the applicable Rayleigh-Jeans regime ($h \nu \ll kT$) and assuming an optically-thin medium, the H I mass of a galaxy can be directly derived from the observed 21 cm flux density spectrum:

$$M_{\text{HI}} = \frac{16\pi m_H}{3h\nu_0 A} D_L^2 \int S(\nu) d\nu,$$

where $m_H$ is the mass of the hydrogen atom, $h$ is Planck’s constant, $\nu_0$ is the rest
1.3. NEUTRAL HYDROGEN GAS

CHAPTER 1. INTRODUCTION

The double-horned 21 cm line of the galaxy HIPASSJ0344-44.

Figure 1.8: The double-horned 21 cm line of the galaxy HIPASSJ0344-44.

frequency of the emission line, \( A \) is the Einstein coefficient for spontaneous emission, \( D_L \) is the luminosity distance to the galaxy, \( S(\nu) \) is the observed 21 cm flux density per channel and the integral is over the frequency range of the galaxy’s \( \text{H}_1 \) profile. The spread in the actual frequency of the emitted photons, inversely related to the transition lifetime, is insignificant. However, the width of the observed line will be increased by the velocity dispersion of the gas and the rotational velocity of the galaxy. Due to the rotational profile of disc galaxies, the \( \text{H}_1 \) line often has a ‘double-horned’ shape, as in Figure 1.8.

Since 21 cm detection is the most direct method of tying observables to physical quantities (\( \text{H}_1 \) masses) with minimal assumptions, this method is preferred to the indirect DLA detection method, where possible. A detailed summary of past and present 21 cm surveys will be presented in the next section.
1.3.3 21 cm surveys

1.3.3.1 Surveys of the local Universe

Large numbers of galaxies must be studied to understand the overall distribution of gas in the local Universe and to examine global trends in \( \text{H}_1 \) properties of galaxies. However, prior to the 1990s the small field-of-view of radio telescopes restricted most extragalactic 21 cm studies to pointed observations of known sources. The installation of multibeam receivers on large, single-dish radio telescopes dramatically improved their instantaneous field of view and facilitated blind, all-sky surveys at 21 cm for the first time.

The first such campaign was the \( \text{H}_1 \) Parkes All Sky Survey (HIPASS) conducted using the 13 beam receiver installed in 1996 on the Parkes 64 m telescope (Staveley-Smith et al., 1996). HIPASS mapped the entire sky south of Declination (Dec.) +25° out to \( z = 0.0423 \). Observations at Dec. < +2° were carried out between 1997 and 2000 (Barnes et al., 2001) and the ‘northern’ field at +2° < Dec. < +25° between 2000 and 2002 (Wong et al., 2006). HIPASS detected 5,317 galaxies which are recorded in the HIPASS source catalogue (HICAT) (Meyer et al., 2004; Wong et al., 2006). This was the largest homogenous \( \text{H}_1 \) galaxy sample ever constructed and revolutionised our understanding of the nearby \( \text{H}_1 \) Universe.

Most notably, HICAT allowed the construction of a \( z \approx 0 \) HIMF with far greater accuracy than previously possible (Zwaan et al., 2003, 2005), as shown in Figure 1.9. The HIMF is essentially the radio counterpart to the optical luminosity function. It describes the distribution of baryons between galaxies but probes gas instead of stars. The well-understood completeness of HICAT allowed the accurate calculation of space densities. This, in combination with the large number of HICAT galaxies detected across a wide range of \( \text{H}_1 \) masses, allowed Zwaan et al. (2005) to closely constrain the faint end of the HIMF. This resolved discrepancies in the faint end slope determined in previous studies, mostly caused by low number statistics. As we
have already seen in section 1.3.2.1, accurate $z = 0$ observational constraints on the low mass end of the HIMF are particularly important for determining the success of galaxy evolution simulations. Multiplying the HIMF by the H$\text{I}$ mass (to give the H$\text{I}$ mass density function) and integrating under this, provided Zwaan et al. with a precise measurement of $\Omega_{\text{H}_1}$ in the local Universe. This is indicated by the red point in Figure 1.6 above.

Similarly, a 7 beam receiver called the Arecibo L-band Feed Array (ALFA) was commissioned on the 305 m Arecibo Observatory in Puerto Rico in 2004. The wide field-of-view and excellent frequency resolution provided by ALFA and its spectral line back-end facilitated the recent Arecibo Legacy Fast ALFA (ALFALFA) H$\text{I}$ survey. ALFALFA began in 2005 and is currently ongoing. When complete, it will cover $\sim 7,000$ deg$^2$ of the sky with eight times the sensitivity of HIPASS and four times the angular resolution, thanks to the large dish diameter (Giovanelli et al., 2005). ALFALFA is expected to detect $>30,000$ galaxies out to $z \sim 0.06$. 

Figure 1.9: The $z \approx 0$ H$\text{I}$ mass function derived from the HIPASS source catalogue (closed points) and fit by a Schechter function (solid line) with parameters as indicated. The dotted line is the Schechter function fit to an earlier data set presented in Zwaan et al. (2003). The number of galaxies contributing to each point is shown in the bottom panel. Reproduced from Figure 1 in Zwaan et al. (2005).
Martin et al. (2010) used the ALFALFA survey, at its 40 per cent complete stage, to generate the HIMF and estimate $\Omega_{\text{H}i}$ in the local Universe, independent of the data set used by Zwaan et al. (2005). The particular sample of galaxies used by Martin et al. contained $\sim 10,000$ sources with $6.2 < \log(M_{\text{HI}}/M_\odot) < 11.0$. This allowed them to probe the HIMF down to masses 1 dex below that of Zwaan et al. (2005). Martin et al. also found a value of $M_{\text{HI}}^*$ (the characteristic ‘knee’ of the HIMF) that was 0.1 dex higher than Zwaan et al. (2005) and an $\Omega_{\text{H}i}$ value 16 per cent higher, which disagreed with the Zwaan et al. value by more than $2\sigma$. Martin et al. concluded that this discrepancy was due to under-counting in the HICAT catalogue of high mass galaxies around the range $9.0 < \log(M_{\text{HI}}/M_\odot) < 10.0$, which dominate the $\Omega_{\text{H}i}$ calculation. The ALFALFA source catalogue contains more of the rare, high $\text{H}i$ mass galaxies than HICAT since ALFALFA probes cosmologically larger volumes than HIPASS and at greater sensitivity. This highlights the fact that surveys of large volume are required to obtain a fair census of the distribution of galaxy properties.

1.3.3.2 Deep surveys

As discussed above, $\text{H}i$ surveys with many of the world’s largest radio telescopes have traditionally concentrated on shallower observations covering a wide sky field in preference to deeper integrations. This allows the largest number of galaxy detections per unit observing time and facilitates studies of the global and statistical $\text{H}i$ properties of galaxies in the relatively nearby Universe. However, the next step is to track the cosmic evolution of gaseous properties of galaxies and link this to other observed evolutionary trends such as the variation in the star formation rate and the Butcher-Oemler effect. To do so, we must make significantly deeper observations to higher redshift and higher sensitivity.

As discussed in Section 1.2.2, there is often a trade-off between survey depth and sky coverage. With the capabilities of the current generation of telescopes, extremely long integration times per field are required to achieve sufficient sensitivities for $\text{H}i$
detection much beyond the local Universe. This practically restricts observations to small areas. High redshift 21 cm detection is also restricted by bandwidth limitations of current front-end and correlator systems. Furthermore, radio frequency interference (RFI) in redshifted H\textsc{i} bands is increasingly becoming a problem, as will be shown in Chapter 2. Nonetheless, there have recently been a number of successful studies of the H\textsc{i} properties in a small number of galaxies beyond \( z = 0.1 \) using new receiver and correlator technology.

Zwaan et al. (2001) employed the larger instantaneous bandwidth of the new Westerbork Synthesis Radio Telescope (WSRT) backend to demonstrate that it is possible to detect H\textsc{i} in galaxies at cosmological redshifts. They targeted the cluster Abell 2218 to maximise the number of galaxies per pointing and therefore the chance of detection. Using \( 18 \times 12 \) h of integration in each of two consecutive 10 MHz bands, Zwaan et al. made an 8 \( \sigma \) detection of 21 cm emission from a galaxy on the outskirts of the cluster at \( z = 0.1766 \). This was the highest redshift direct H\textsc{i} detection of the time.

Based on this success, a more extensive survey of galaxy clusters using the WSRT was then launched. The clusters Abell 963 (at \( z = 0.206 \)) and Abell 2192 (at \( z = 0.188 \)) were observed for \( 20 \times 12 \) h and \( 15 \times 12 \) h, respectively. Verheijen et al. (2007) reported the detection of 20 H\textsc{i} sources in A963 and 30 in A2192. They found H\textsc{i} masses over the range \( 5 \times 10^9 - 4 \times 10^{10} \) \( M_\odot \). The excellent spatial and spectral resolution of the WSRT even allowed rough rotation curves to be derived for some of the detected sources. These high redshift observations enabled both Zwaan et al. (2001) and Verheijen et al. (2007) to study the Butcher-Oemler effect and the general evolution of H\textsc{i} in cluster galaxies.

Catinella et al. (2008) studied H\textsc{i} in field galaxies at higher redshifts than previously possible using ALFA, which facilitates observations down to 1120 MHz (\( z < 0.27 \)) with Arecibo. They used a total of \( \sim 280 \) h to conduct pointed observations of H\alpha emitting galaxies identified in the SDSS. With the excellent sensitivity provided by the large collecting area of Arecibo, Catinella et al. (2008) reported the detection
of 10 galaxies over $0.16 < z < 0.27$ with S/N $> 7$. The most distant detection was made at $z = 0.2454$ with 176 minutes of on-source integration. This corresponds to the highest-redshift direct detection of 21 cm emission ever made. To be detected at these high redshifts, these galaxies are of course very massive and gas rich, with the most distant galaxy having $M_{\text{HI}} = 10^{10.9} M_\odot$.

While Zwaan et al. (2001), Verheijen et al. (2007) and Catinella et al. (2008) showed that high redshift detections are possible with pointed observations of clusters or optically-selected galaxies, the next challenge is to achieve high redshift detection of larger numbers of field galaxies with blind surveys to study H I evolution in the galaxy population at large. This has motivated the Arecibo Ultra Deep Survey (AUDS) which is a blind survey of $1.35 \text{deg}^2$ out to $z = 0.16$. AUDS has recently been completed and hopes to detect H I in an unbiased sample of galaxies by using 1,200 integration hours to achieve a sensitivity of 50 $\mu$Jy (Freudling et al., 2011; Hoppmann et al., in prep). Freudling et al. (2011) reported the results of 53 h of AUDS precursor observations targeting an overdensity of SDSS galaxies. They detected 18 H I lines at $0.07 \lesssim z \lesssim 0.15$. From this, Freudling et al. estimated the mean H I mass density at $z = 0.125$ to be $(1.0 \pm 0.3) \rho_0$, where $\rho_0$ is the density at $z = 0$. Evolution has clearly not been detected. Either there is none, or the effect is smaller than the cosmic variance-dominated uncertainties. So while this study demonstrated that H I detection is possible out to $z = 0.16$ with a blind survey, interpretation of the results is difficult due to biases introduced by the necessarily small field size observed.

As backend and receiver technology progresses, the next generation of radio telescopes will be capable of conducting 21 cm surveys that are simultaneously wide field, sensitive and cover a large redshift range. This is the ideal scenario for comprehensive studies of the distant H I Universe. Ultimately, the Square Kilometre Array (SKA; e.g. Schilizzi et al., 2008; Rawlings, 2011) telescope will be the instrument of choice for this work. With the broad frequency range of this facility, it is hoped that direct H I detections in emission out to $z \approx 1$ and beyond will be possible. While the SKA will not be fully operational for more than a decade, precursor instruments are cur-
### Table 1.1: A compilation of HI survey parameters. Values marked with an asterisk are projected estimates. After Hanning.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Instrument</th>
<th>(\Delta z)</th>
<th>Survey area</th>
<th>Target</th>
<th>(z) range</th>
<th>(\Delta z) (HII)</th>
<th># HI</th>
<th>Notes</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zwann et al. (2001)</td>
<td>WSRT</td>
<td>0.1766</td>
<td>A2218</td>
<td>0.4</td>
<td>78.1</td>
<td>31.2</td>
<td>0.0193</td>
<td>VLA</td>
<td></td>
</tr>
<tr>
<td>Verheijen et al. (2007)</td>
<td>WSRT</td>
<td>0.164</td>
<td>A963, A2192</td>
<td>0.81</td>
<td>50</td>
<td>39.0</td>
<td>6.1</td>
<td>ALFAPA</td>
<td></td>
</tr>
<tr>
<td>Catinella et al. (2008)</td>
<td>Arecibo</td>
<td>0.16</td>
<td>SDSS galaxies</td>
<td>10</td>
<td>6.1</td>
<td>31.4</td>
<td>52.1</td>
<td>ALFAPA</td>
<td>0.09</td>
</tr>
<tr>
<td>Meyer et al. (2004)</td>
<td>HIPASS</td>
<td>&lt;0.0423</td>
<td>Blind</td>
<td>17,027</td>
<td>5,317</td>
<td>62.5</td>
<td>31.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wong et al. (2006)</td>
<td>Martin et al. (2010)</td>
<td>ALFAPA</td>
<td>0.05</td>
<td>40</td>
<td>52.1</td>
<td>0.0193</td>
<td>VLA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nath &amp; al. (2005)</td>
<td>Martin et al. (2010)</td>
<td>ALFAPA</td>
<td>&gt;30,000</td>
<td>Blind</td>
<td>7,074</td>
<td>25.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freudling et al. (2011)</td>
<td>AUDS pilot</td>
<td>0.07-0.16</td>
<td>Blind</td>
<td>0.09</td>
<td>1.35</td>
<td>24.4</td>
<td>5.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoppmann et al. (in prep)</td>
<td>AUDS</td>
<td>0.07-0.16</td>
<td>&gt;115</td>
<td>0.16</td>
<td>24.4</td>
<td>5.21</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fernández et al. (2013)</td>
<td>JVLA</td>
<td>0-0.193</td>
<td>Blind</td>
<td>0.32</td>
<td>33</td>
<td>31.2</td>
<td>0.0193</td>
<td>VLA</td>
<td></td>
</tr>
</tbody>
</table>

Parameters for all surveys (2011) after Hanning smoothing. Remaining columns are survey parameters. Values marked with an asterisk are projected estimates. After Hanning.
CHAPTER 1. INTRODUCTION  

1.3. NEUTRAL HYDROGEN GAS

Currently being constructed and include the Australian SKA Pathfinder (ASKAP; e.g. Johnston et al., 2008) in Western Australia, the Karoo Array Telescope (MeerKAT; e.g. Booth & Jonas, 2012) in South Africa and the Aperture Tile in Focus (APER-TIF; Verheijen et al., 2008) in the Netherlands. These should enter commissioning phases within the next year or two.

In fact, the Very Large Array (VLA) in Soccoro, New Mexico has recently been upgraded (now called the Karl G. Jansky VLA; JVLA) to become the first synthesis instrument capable of performing reasonably wide-field, high redshift H\textsc{i} surveys. The improved receivers and new backend of the JVLA provide a larger field of view, improved sensitivity and a wider instantaneous bandwidth (Perley et al., 2011). A deep, blind H\textsc{i} survey over $0 < z < 0.45$ is being planned to capitalise on these new JVLA capabilities, and is referred to as the Cosmological Evolution Survey (COSMOS) H\textsc{i} Large Extragalactic Survey (CHILES). Fernández et al. (2013) have recently conducted a pilot survey for CHILES, covering $0 < z < 0.193$ over a $34 \times 34$ arcmin$^2$ region of the COSMOS field. With only 50 h of observations they made 33 H\textsc{i} detections in a variety of environments - walls, voids, individual and interacting galaxies, over the full redshift range. Fernández et al. predict that with 1000 h of observations of the COSMOS field, CHILES will detect $\sim 300$ galaxies and facilitate a comprehensive study of H\textsc{i} content evolution as a function of redshift and environment.

Table 1.1 compares the various parameters for each shallow and deep H\textsc{i} survey discussed in this and the previous section.

1.3.4 Spectral stacking of the 21 cm line

The statistical data analysis technique of stacking is the focus of this thesis. It can be used to extract information from the limits of currently available data sets and can therefore push the redshift boundary of H\textsc{i} surveys. While other studies have concentrated on direct 21 cm detections, requiring long integrations of small numbers of distant galaxies, the stacking technique does not require individual detections.
1.3. NEUTRAL HYDROGEN GAS

Rather, spectral stacking combines the undetected 21 cm signals of many different sources, identified using external redshift catalogues, such that a strong statistical detection is obtained. Therefore, stacking can efficiently provide information on the average properties of the high redshift galaxy population as a whole. This is particularly useful for measuring evolution in cosmic quantities such as $\Omega_{\text{HI}}$. Chapter 2 will provide a detailed explanation of the stacking process.

The validity of the stacking, or ‘co-adding’, technique was first demonstrated by Zwaan (2000) and Chengalur et al. (2001). Similar to the work of Zwaan et al. (2001) and Verheijen et al. (2007), their aim was to study the Butcher-Oemler effect by detecting $\text{HI}$ in cluster galaxies at high redshifts. Instead of requiring direct detections and therefore only examining a handful of galaxies in the cluster, they used the stacking technique to examine the average properties of all the cluster galaxies with available spectroscopic redshifts. Zwaan (2000) studied Abell 2218 at $z = 0.18$ with the WSRT (see the description of the Zwaan et al., 2001 observations in the previous section) and stacked 45 undetected galaxies in the core of the cluster to achieve a $2.5\sigma$ detection. They measured a very low average $\text{HI}$ mass for these galaxies, $3 \times 10^8 h^{-2} M_\odot$, which is consistent with the $\text{HI}$ deficiency observed in galaxies in the central regions of more nearby clusters. Chengalur et al. (2001) conducted an intentionally shallow survey of Abell 3128 at $z = 0.06$ with the Australia Telescope Compact Array (ATCA). While they directly detected only two galaxies in the cluster, a stacking analysis of the 193 galaxies with available spectroscopic redshifts revealed an average $\text{HI}$ mass of $5 \times 10^8 h^{-2} M_\odot$. This demonstrated that stacking can be used to probe low mass systems, undetected above the noise level in a given set of observations.

Lah et al. (2009) then showed that such stacking analyses of cluster galaxies are possible at much higher redshift. They examined the $\text{HI}$ content of galaxies in the cluster Abell 370 at $z = 0.37$ using $\sim 40$ h of observations with the Giant Metrewave Radio Telescope (GMRT) near Pune, India. They conducted optical imaging of the cluster using the Australian National University 40-inch telescope and obtained optical
spectroscopy and redshifts for 324 cluster galaxies using the AAT. No individual H\textsc{i} detections of these high redshift galaxies were found in the radio data. Therefore, Lah et al. used the stacking technique to examine the H\textsc{i} properties of all available cluster galaxies and of galaxies as a function of various properties such as position within the cluster. Despite the fact that the most significant stacked detection they achieved was 3\(\sigma\), they were still able to conclude that galaxies on the outside of the cluster contain more H\textsc{i} gas than galaxies closer to the cluster centre. They also found that the overall H\textsc{i} density of Abell 370 galaxies is significantly higher than in the nearby Virgo and Coma clusters. Therefore, the use of the stacking technique by Lah et al. suggests that there could have been significant evolution in the H\textsc{i} content of cluster galaxies over the past ~4 billion years and seems to confirm the relationship between galaxy environment and H\textsc{i} content.

Lah et al. (2007) were the first to show that co-adding techniques can be used to examine the H\textsc{i} content of field galaxies (as opposed to solely cluster galaxies) at intermediate redshifts. Lah et al. (2007) used the 2dF and AAOmega spectrographs on the AAT to obtain spectroscopic redshifts for 154 H\textsc{\alpha} emitting galaxies at 0.218 < \(z\) < 0.253, chosen from Subaru Telescope imaging by Fujita et al. (2003). These star forming galaxies were targeted since they should contain significant amounts of neutral gas and therefore maximise the chance of achieving a stacked detection. H\textsc{i} data was collected using ~40\,h (excluding overheads) with the GMRT. No galaxies were individually detected at 21 cm and co-adding all 121 sources produced no obvious detection. However, binning the stacked data to 500\,km\,s\(^{-1}\) revealed a weak statistical detection with a S/N of 2.6 and is reproduced here in Figure 1.10.

Using this stacked detection, Lah et al. (2007) made the highest redshift measurement to date of the cosmic gas density using 21 cm data (excluding measurements with intensity mapping, discussed below). They measured the total cosmic mass density of neutral gas (\(\Omega_{\text{gas}}\)), which includes a 24\,per cent correction for the contribution from neutral helium, to be \((0.63 \pm 0.29) \times 10^{-3}\,h^{-1}\) at \(z = 0.24\). These findings are consistent with DLA measurements, although the small sky field studied by Lah
et al. (2007) (34×27 arcmin$^2$) means there is up to 40 per cent cosmic variance in the result. Furthermore, this measurement of $\Omega_{\text{gas}}$ is technically a lower limit as it only includes galaxies displaying star-formation. Lah et al. (2007) argue that the underestimation will be small since non-star-forming galaxies are assumed to contain insignificant quantities of neutral gas. Therefore, Lah et al. (2007) demonstrated that stacking can be used to trace evolution in cosmic densities at higher redshifts than otherwise possible.

Fabello et al. (2011a) then showed that H I stacking can be performed on large numbers of field galaxies to enable a detailed analysis of various evolutionary trends. Fabello et al. (2011a) examined the 4,726 sources in the ALFALFA 40 per cent data set at $0.025 < z < 0.05$ with stellar masses greater than $10^{10} \, M_\odot$. 77 per cent of these galaxies did not have 21 cm detections in ALFALFA. Co-adding these non-detections in five stellar mass bins produced stacked signals with S/N greater than 6.5. Therefore, Fabello et al. (2011a) were able to demonstrate that strong statistical detections of individually-undetected, low redshift field galaxies are possible when using large
samples. Fabello et al. (2011a) and Fabello et al. (2011b) then used these high significance detections to examine the trends in average H\textsc{i} content with other physical parameters of galaxies such as bulge-presence and AGN activity. They found no evidence that the bulge influences the H\textsc{i} content and also conclude that AGN activity does not strongly influence the evolutionary properties of the host galaxies at low redshifts.

It is also worth noting that the related method of intensity mapping is showing potential for high redshift detection of H\textsc{i}. This uses the cross-correlation between optical and 21\text{cm} data sets to make aggregate detections of H\textsc{i} in the cosmic web, as was shown to be possible by Pen et al. (2009) using very low redshift data. Chang et al. (2010) used intensity mapping of Green Bank Telescope (GBT) data to make the first detection of H\textsc{i} in the large-scale structure at $0.53 < z < 1.12$. Masui et al. (2013) improved upon this work and derived the most precise measurement of $\Omega_{\text{H}\textsc{i}}$ at $z \approx 0.8$. However, this measurement depends upon an H\textsc{i} bias parameter and a galaxy-hydrogen correlation coefficient, which combined can vary between 0.5-2 (black dashes in Figure 1.6). Constraining these values awaits improved redshift space distortion measurements (Chang et al., 2010).

### 1.3.5 This work

The work presented in Chapters 3 and 4 of this thesis uses the spectral stacking technique to study galaxy H\textsc{i} properties at higher redshifts than otherwise possible. Previous studies have already shown that stacking is a powerful tool for probing H\textsc{i} trends in otherwise unattainable galaxy populations. Here, we will apply stacking to larger volumes and larger samples of field galaxies than ever before, to study statistical properties of the entire galaxy population and examine evolutionary trends. The 21\text{cm} data for this work comes from the 64\text{m} Parkes radio telescope, chosen for its rapid survey speed and good sensitivity. The low redshift ($0 < z < 0.04$) data are provided by the all-sky but shallow HIPASS survey. We conduct new Parkes
# Table 1.2: Parameters of various $H_1$ stacking analyses.

<table>
<thead>
<tr>
<th>Source/Criteria</th>
<th>$\Delta \nu$ (kHz)</th>
<th>$\sigma$ (rms)</th>
<th>$N_{\text{galaxies}}$ (mJy)</th>
<th>Stack Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CANA</strong> (&lt; 19.4' field galaxies)</td>
<td>98.93</td>
<td>1.33</td>
<td>6.25</td>
<td>This work (Chapter 4)</td>
</tr>
<tr>
<td><strong>CANA</strong> (&gt; 19.4' field galaxies)</td>
<td>1.124</td>
<td>1.33</td>
<td>6.25</td>
<td>This work (Chapter 3)</td>
</tr>
<tr>
<td><strong>SDSS</strong> (&lt; 160' field galaxies)</td>
<td>0.025 – 0.05</td>
<td>25.1</td>
<td>$\sim$ 2.5</td>
<td>Padella et al. (2011)</td>
</tr>
<tr>
<td><strong>ALFALFA</strong> (&lt; 2000' field galaxies)</td>
<td>1.12</td>
<td>0.018 – 0.25</td>
<td>125</td>
<td>Catin (2007)</td>
</tr>
<tr>
<td><strong>VCT</strong> (&lt; 2000' field galaxies)</td>
<td>1.3</td>
<td>0.06</td>
<td>125</td>
<td>Catin (2009)</td>
</tr>
<tr>
<td><strong>SDSS</strong> (&lt; 160' field galaxies)</td>
<td>1.12</td>
<td>0.018 – 0.25</td>
<td>125</td>
<td>Catin (2009)</td>
</tr>
<tr>
<td><strong>ALFALFA</strong> (&lt; 2000' field galaxies)</td>
<td>1.12</td>
<td>0.018 – 0.25</td>
<td>125</td>
<td>Catin (2009)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Survey</th>
<th>Instrument/Field</th>
<th>$N_{\text{galaxies}}$ (mJy)</th>
<th>$\Delta \nu$ (kHz)</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2dFGRS</td>
<td>Parkes</td>
<td>15,093</td>
<td>62.5</td>
<td>13.3</td>
</tr>
<tr>
<td>GAMA</td>
<td>Parkes</td>
<td>1,152</td>
<td>62.5</td>
<td>13.3</td>
</tr>
<tr>
<td>WISE</td>
<td>Survey</td>
<td>4,873</td>
<td>78.1</td>
<td>4.2</td>
</tr>
<tr>
<td>HIPASS</td>
<td>Instrument/Field</td>
<td>$N_{\text{galaxies}}$ (mJy)</td>
<td>$\Delta \nu$ (kHz)</td>
<td>$z$</td>
</tr>
<tr>
<td>2dFGRS</td>
<td>Parkes</td>
<td>15,093</td>
<td>62.5</td>
<td>13.3</td>
</tr>
<tr>
<td>GAMA</td>
<td>Parkes</td>
<td>1,152</td>
<td>62.5</td>
<td>13.3</td>
</tr>
<tr>
<td>WISE</td>
<td>Survey</td>
<td>4,873</td>
<td>78.1</td>
<td>4.2</td>
</tr>
</tbody>
</table>

*Note: The table provides parameters for various $H_1$ stacking analyses, including source criteria, instrument used, number of galaxies, channel spacing, and redshift.*
observations to obtain high redshift ($0.04 < z < 0.14$) $21$ cm data across chosen sky fields (shown in Figure 1.3). In Chapter 3 we combine the radio data with optical spectroscopic information provided by the wide-field 2dFGRS. For the analysis of Chapter 4, the spectroscopic redshifts are provided by the new, deep GAMA survey.

Table 1.2 shows how the parameters of the stacking analyses presented here compare to other studies from the literature. While our data do not probe redshifts as large as Lah et al. (2007), our sample size is considerably larger and therefore cosmic variance will not bias our results as significantly. Similar to the work of Fabello et al. (2011a), we will demonstrate that strong statistical detections of large numbers of individually-undetected galaxies are possible, however we will push to significantly higher redshifts (and therefore larger volumes) where evolution may be evident. Furthermore, our sample selection simply consists of detection in optical imaging. This means our results will be far less biased to a particular galaxy population by H$\alpha$ emission or stellar mass requirements. Therefore, this thesis will make a significant contribution to our understanding of the evolutionary trends of H$\text{I}$ gas in the Universe and will develop the novel spectral stacking technique for future use with the next generation of radio telescopes.

### 1.4 Active galactic nuclei and radio galaxies

#### 1.4.1 Overview of AGN and radio galaxies

A subset of galaxies in the Universe are observed to be ‘active’. Such galaxies have very luminous compact nuclear regions emitting non-thermal radiation, often across the entire spectrum from X-ray to radio. It is believed that these AGN are powered by the accretion of matter onto a central supermassive black hole (SMBH). AGN have been observed to emit various combinations of X-ray, optical, infrared and radio
Figure 1.11: A unified model of AGN. The different types of AGN observed along each sight line are indicated. The thin inner ring is the accretion disk and the thick outer ring is the dusty torus. The black clumps close to the accretion disk are the broad line clouds and the grey clumps are the narrow line clouds. The radio jet is the collimated emission perpendicular to the accretion disk. Adapted from Figure 1 in Torres (2004).
continuum, broad and narrow optical emission lines, X-ray emission lines and powerful radio jets.

Traditionally, different types of AGN (such as Seyfert 1 and 2 galaxies, quasars, radio galaxies, and blazars) have been classified based on the strength and presence of these various emission signatures. However, a unified model has emerged in which all AGN have common physical components and the different emission signatures (and hence different classes) are explained by the different angles of inclination of the AGN relative to the observer (illustrated in Figure 1.11; see e.g. Antonucci 1993 for a review).

The standard AGN components include: (i) a central SMBH with a mass typically between $10^6$ and a few times $10^9 M_\odot$ (e.g. Peterson et al., 2004), (ii) a disc of material accreting onto the black hole which is heated by viscosity and is responsible for the bright nuclear continuum emission, (iii) a dusty torus which absorbs emission from the accretion disc and re-radiates in the infrared, (iv) gas clouds which produce broad or narrow emission lines based on their proximity to the SMBH, and (v) a hot, X-ray emitting corona. In the case of radio-loud AGN, a collimated jet of outflowing material also exists, emanating from the central nucleus and producing large quantities of non-thermal synchrotron radiation. Under this framework, blazars and quasars are thought to be AGN viewed directly along, or at small angles to, the jet. Whether or not the dusty torus obscures the observer’s view of the accretion disk determines whether broad lines are seen and hence the AGN classification e.g. Seyfert 1 or 2. It should be noted, however, that this unified model does not explain the range of observed AGN luminosities or the absence or presence of radio-loud activity.

Radio-loud AGN (hereafter referred to as radio galaxies) constitute roughly 10-20 per cent of the total AGN population, although this is luminosity-dependent (e.g. Kellermann et al., 1989). As we will see below, these are a particularly interesting sub-population since their radio jets can couple with the gas in the host galaxy, sometimes to dramatic effect. Fanaroff & Riley (1974) defined two classes of radio galaxies (now known as FR I and FR II objects) based on their radio structure. FR
I systems are often brightest in the central core region, generally have continuous two-sided jets and display decreasing surface brightness towards the edges of the jets (i.e. edge-darkened). FR II systems commonly display edge-brightened hotspots, have comparatively fainter cores and often have one-sided jets. Fanaroff & Riley (1974) found that this structure difference was correlated with radio luminosity, with the higher power objects displaying FR II morphology. It is now fairly well understood that this correlation is caused by the different interactions of the transonic (FR I) or supersonic/hypersonic (FR II) jets with their environments (e.g. Bicknell, 1995).

However, a more puzzling distinction between radio galaxies involves the strength of their optical emission lines. High-excitation radio galaxies (HERGs) with strong emission lines have, in general, higher radio luminosities than low-excitation radio galaxies (LERGs) with weak emission lines (e.g. Hine & Longair, 1979; Laing et al., 1994; Jackson & Rawlings, 1997). However, the HERG/LERG distinction does not seem to correlate with the FR I/II separation and therefore cannot be explained in the same manner. Instead, it is possible that different fuelling mechanisms (i.e. different gas origins) could be responsible for the two emission modes. The emerging picture is that LERGs are powered by radiatively-inefficient mechanisms such as spherically-symmetric Bondi accretion or advection-dominated accretion flow of hot gas from the ISM, while HERGs are powered by radiatively-efficient accretion of cold gas supplied through mergers and interactions (Hardcastle et al., 2007; Best & Heckman, 2012). Further multiwavelength observations are required to improve our understanding of these different radio galaxy classes and their importance within the overall framework of galaxy evolution.

1.4.2 The role of AGN in galaxy evolution

Observational evidence suggests that AGN activity can strongly influence the growth and evolution of galaxies. For example, the cosmic evolution of the quasar luminosity density, AGN co-moving number density and black hole accretion rate appear to
Figure 1.12: The cosmic evolution of the star formation rate (red measurements compiled by Hopkins, 2004 and green measurements from Bouwens et al., 2007) and the black hole accretion rate (solid black line with shaded 1σ error region; Aird et al., 2010). The difference between the red and green points is largely due to different assumptions regarding dust presence (Behroozi et al., 2013). Reproduced from Figure 13 in Aird et al. (2010).

track that of the star formation rate density and peak at around $z \approx 2$ (e.g. Boyle & Terlevich 1998, Wall et al. 2005, Silverman et al. 2005, Richards et al. 2006, Croom et al. 2009, Aird et al. 2010; see Figure 1.12). This implies that AGN activity is somehow linked to the star formation rates of galaxies.

Furthermore, there appears to be a tight correlation between the mass of the SMBH ($M_{\text{BH}}$) and the stellar mass of the galaxy bulge ($M_{\text{bulge}}$; e.g. Magorrian et al. 1998; Marconi & Hunt 2003; Häring & Rix 2004). In particular, Häring & Rix (2004) found a relationship of $\log(M_{\text{BH}}) \propto 1.12 \log(M_{\text{bulge}})$ with a scatter of less than 0.3 dex (see Figure 1.13). This suggests that the growth of black holes and their host galaxies are co-dependent.

The overall implication from these observations is that AGN can somehow influence their environments. This is believed to occur through various ‘feedback’ processes which prevent gas from cooling, actively heat gas and/or remove gas entirely from
the host galaxy. If AGN accretion is radiatively efficient, energy can be transferred to the gas in the ISM via Compton scattering and photoionization heating by UV and X-ray photons (e.g. Ciotti & Ostriker, 2001). Furthermore, if the gas is coupled to large amounts of dust in the ISM, radiation pressure can drive the gas outwards. For example, Maiolino et al. (2012) observed dramatic outflows, at a rate of over 3,500 $M_\odot$ yr$^{-1}$, from a $z = 6.4$ quasar. The radiative feedback mode is thought to occur mostly in younger AGN and works primarily on the cold gas (e.g. Fabian, 2012). Since cold gas is an essential ingredient in the formation of new stars, as we have already seen in section 1.3, AGN feedback can therefore inhibit or altogether quench star formation.

In the case of radio-loud AGN, additional mechanical feedback may be provided by the powerful radio jets. It is possible that the jet can both enhance and hinder star formation in the host galaxy. Shocks induced by the jet may compress gas, causing cloud collapse and subsequent star formation (e.g. Fragile et al., 2004). A clear example of this is Minkowski’s object, a starburst triggered by the radio jet from a nearby active galaxy (van Breugel et al., 1985). Indeed, Croft et al. (2006)
detected double H\textsc{i} clouds straddling the jet which were formed by collisions between the jet and the interstellar material. Conversely, jet-induced shock fronts may create X-ray cavities, or `bubbles' of relativistic plasma, which heat the gas and prevent cooling flows and therefore star formation (e.g. Fabian et al. 2006, Forman et al. 2007). Furthermore, mechanical energy from the jet may produce violent outflows of gas, thereby stripping the galaxy of its star forming fuel. This has been seen in 21 cm observations of H\textsc{i} at low redshift \((z \lesssim 0.2\); e.g. Morganti et al. 2005, 2007) and in optical emission lines out to high redshifts \((z \gtrsim 2\); e.g. Nesvadba et al. 2008). Since this gas is also required to feed the AGN itself, the outflows may halt AGN activity thus making the AGN self-regulating. This may help to explain the observed black hole/bulge mass correlation.

As further evidence of the crucial role of AGN activity on galaxy evolution, this feedback is required to reconcile observed and simulated trends in the local Universe. For example, AGN feedback must be invoked in semi-analytical models to halt the formation of the most massive galaxies and reproduce the observed cutoff in the bright end of the luminosity function (e.g. Croton et al., 2006; Bower et al., 2006; see Figure 1.14). AGN-induced gas outflows are also required to accurately reproduce the HIMF, as we have already seen in section 1.3.2.1 (e.g. Davé et al., 2013).

While the relationship between AGN activity and star formation is obviously central to studies of galaxy evolution, the microphysics of the various feedback processes are not well understood and the relationship is clearly complex. Due to the coupling between the jet and the gas of a radio galaxy, we may expect to see strong trends between radio-loud AGN activity and star formation. This may be particularly true for HERGs, which are believed to be radiatively efficient and so can influence both the hot and cold gas. LERGs, on the other hand, are predominantly radiatively inefficient and so are thought to affect only the hot, ionised IGM via the kinetic energy of the jet (e.g. Hardcastle et al., 2009).

The connection between AGN and star formation is yet further complicated by the fact that both phenomena can be triggered by a common mechanism. One likely
scenario for AGN trigger is a gas-rich merger. There is certainly long-standing observational evidence that the host galaxies of powerful radio-loud AGN have disturbed morphologies (Heckman et al., 1986). The merger can create a torque, resulting in a flow of gas towards the central regions of the galaxy where black hole accretion can occur and AGN activity commence (e.g. Granato et al., 2004; Di Matteo et al., 2005). In addition, the larger surface density of gas in the central regions of the merger can promote a significantly higher than normal star formation rate (e.g. Mihos & Hernquist, 1996). However, the time scales of the merger-induced starburst and AGN accretion may be different (e.g. Wild et al., 2010) and so we may not always observe a galaxy in both phases simultaneously (e.g. Dicken et al., 2012).

Therefore, radio-loud AGN activity and star formation are intricately linked. It remains to be seen whether the feedback dominates the relationship or whether both are triggered by the same event and then evolve somewhat independently. To obtain a clearer picture, further observations out to higher redshift are required.
1.4.3 Observations of star formation in radio galaxies

The connection between star formation and radio-loud AGN activity has not been well-studied until recently because of limitations in available data sets. For observations in most wavebands, complications arise in separating out the contribution from the young stars and the emission (thermal or non-thermal) from the AGN. As we will see below, any studies which successfully avoided this contamination were restricted in sample size and/or redshift range.

Star formation can ordinarily be traced in the UV and optical. However, these wavelengths can also be severely contaminated by nuclear or jet emission from the AGN (e.g. Tadhunter et al., 1996; Holt et al., 2007). Nonetheless, there have been a number of successful studies in this regime. Baldi & Capetti (2008) used the high resolution of HST optical and UV imaging to mask AGN emission. They were therefore able to identify 25 non-contaminated, star-forming, radio-loud AGN at low redshift ($z < 0.1$). With this sample, they showed that HERGs are often associated with enhanced star formation and interacting systems, while LERGs are predominantly not. Herbert et al. (2010) conducted optical spectroscopy of 24 powerful radio galaxies at $z \sim 0.5$ using the William Herschel Telescope and Gemini North telescope and used the 4000 Å break strength as a proxy for star formation. Once again, they found that LERGs are usually associated with older stellar populations and HERGs more often with star formation. While both studies were quite successful in mitigating against contamination, they were both restricted to small sample sizes and low redshifts due to practical limitations of the observing instruments (such as available observing time). So while the findings seem to agree with the current understanding of LERG/HERG fuelling, a larger sample must be studied to fairly examine the properties of the overall population.

Similarly, mid-infrared continuum emission is a signature of star formation, but many studies have shown that these wavelengths are also heavily contaminated by thermal emission from dust in the AGN accretion disc and surrounding torus (e.g. Dicken
et al., 2009; Hardcastle et al., 2009). In particular, Haas et al. (2004) showed the existence of a thermal ‘bump’ in AGN spectral energy distributions (SED) at around 60-100 $\mu$m caused by AGN-associated dust. However, using the Spitzer Infrared Spectrograph, Dicken et al. (2012) showed that spectroscopic mid-infrared observations can be used to identify polycyclic aromatic hydrocarbon (PAH) emission lines, and thus trace star formation in radio galaxies. They found evidence for star formation in only $\sim$30 per cent of radio galaxies in their sample. They suggest that this argues against a direct relationship between AGN activity and star formation signatures, or at least that it is rare for AGN activity to occur close to the peak of a merger-induced starburst. However, interpretation of these results within a broader context are once again limited by the moderate sample size (46) and low redshift range ($0.05 < z < 0.7$) probed.

The far-infrared (FIR) is an ideal regime over which to study star formation in AGN. Dust is heated by UV from young stars to temperatures of roughly 15-60 K and re-radiates in the FIR/submillimetre. Therefore, star-formation can be well-traced at these wavelengths (Kennicutt, 1998; Kennicutt et al., 2009). As we have already seen, the thermal emission from dust in the accretion disc and torus peaks in the mid-infrared and should not contribute significantly above $\sim$100 $\mu$m (Haas et al., 2004). Therefore, the only major contaminant will be emission from ‘cold’ dust heated by the older stellar population, rather than newly forming stars. Furthermore, dust and gas is mostly optically-thin at these long wavelengths and so should not obscure the FIR emission.

Ground-based instruments, such as the Submillimetre Common User Bolometer Array (SCUBA), have contributed in this regime. For example, Archibald et al. (2001) and Reuland et al. (2004) used this facility to survey the star formation signatures of 47 and 24 powerful radio galaxies, respectively. They both found a strong redshift dependence on the detection rate of submillimetre emission from these galaxies, but argue that this is likely a general property of massive ellipticals rather than a trend caused by radio-loud activity. However, their sample was restricted to high redshift
(1 < z < 5), where the peak in the rest-frame thermal FIR emission will be shifted into the submillimetre (850 µm). Due to the flux-limited nature of the SCUBA surveys, this meant that only the highest energy AGN were probed.

Fortunately, the Herschel Space Observatory now allows very sensitive, high-resolution FIR surveys, overcoming many of the barriers faced in previous studies. Herschel is a 3.5 m space-based far-infrared/submillimetre telescope launched on 14th May 2009 (Pilbratt et al., 2010). It has two onboard cameras - the Photodetector Array Camera and Spectrometer (PACS) which functions below 210 µm (Poglitsch et al., 2010), and the Spectral and Photometric Imaging Receiver (SPIRE) which can image at 250, 350 and 500 µm (Griffin et al., 2010). The combined sensitivity and wavelength capabilities of Herschel make it the most powerful tool for detecting dust in both nearby and distant galaxies.

The Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS) is one of the key extragalactic surveys conducted with Herschel. The full program uses 600 hours to cover 550 deg² of the sky spread over the three equatorial GAMA fields, a 150 deg² North Galactic Pole field and a 250 deg² South Galactic Pole region (Eales et al., 2010). Using both PACS and SPIRE, it collects imaging data at 100, 160, 250, 350 and 500 µm. The first data to be released in 2010 covered the ~14 deg² science demonstration phase (SDP) field (the green region in Figure 1.3). Data over the full ~161 deg² H-ATLAS ‘Phase 1’ field (the three GAMA regions in red in Figure 1.3) has not yet been made public.

A number of recent studies have employed this sensitive, wide-field H-ATLAS data, together with an image stacking technique, to examine the average star formation properties of larger samples of radio-loud AGN than ever before. The stacking strategy, detailed in Chapter 2, is used to probe an even deeper (fainter and higher redshift) population than can be directly detected by Herschel.

The first of these studies was conducted by Hardcastle et al. (2010). Using the NRAO VLA Sky Survey (NVSS) and Faint Images of the Radio Sky at Twenty-Centimetres
(FIRST) 1.4 GHz radio continuum surveys, they identified the 187 radio galaxies with 1.4 GHz flux densities greater than 2.5 mJy present in the H-ATLAS SDP field. They then identified the host galaxies of these radio sources in near-infrared imaging data provided by the UKIRT Infrared Deep Sky Survey - Large Area Survey (UKIDSS LAS; hereafter LAS; Lawrence et al., 2007), which covers 92 per cent of the SDP field. Hardcastle et al. collected redshifts for all these objects, either from the GAMA spectroscopic redshift catalogue, or from photometric redshifts derived using LAS and SDSS photometry. Using this external information, they performed various stacking analyses of H-ATLAS 250 $\mu$m data at the positions of the radio galaxies. They found an increase in the FIR luminosity (and therefore SFR) of the galaxies with redshift and radio luminosity. However, they found no difference compared to a radio-quiet control population. Therefore, they conclude that any trends they see are a result of evolution of the overall galaxy population, rather than a dependence on the AGN activity. However, their source selection was restricted to galaxies with a $K$-band detection in the LAS and with $r < 22$ mag, thus limiting the analysis to relatively low redshift ($z < 0.85$) and low radio luminosity ($L_{1.4} \lesssim 10^{27}$ WHz$^{-1}$) objects. Thus, it is possible that their sample does not contain many higher luminosity HERGs, which are expected to be hosted by star-forming, merging systems. This could explain the absence of any unusual star formation trends.

Virdee et al. (2013) expanded the work of Hardcastle et al. (2010) over 128.8 deg$^2$ of the H-ATLAS Phase 1 field, again using optical and near-infrared data from SDSS and UKIDSS. Due to the larger sky area probed, their radio galaxy sample was almost an order of magnitude larger than that of H10 (1,599), allowing them to split their sample by 1.4 GHz luminosity, source size, stellar mass and spectral index while still achieving a statistically-significant average FIR detection in each. They found a significant deficit in the 250 $\mu$m luminosity of high stellar mass radio detected galaxies, compared to a $K$ band luminosity and $g - r$ colour matched radio-quiet population. They argue that this is caused by a deficit of dust in the environments hosting the radio-loud galaxies. This is consistent with the popular scenario in which LERGs
are hosted by massive galaxies where the hot, dense halo inhibits star formation and subsequent dust production.

Finally, Hardcastle et al. (2013) conducted a similar stacking analysis over the full H-ATLAS field to explicitly study the difference between the SFRs of HERGs and LERGs. They used spectroscopic observations provided by GAMA, or the GAMA spare fibre program (Ching et al., in prep) to identify radio sources with and without strong emission lines, which they equate to high and low excitation galaxies respectively. They find that high excitation galaxies have four times higher average 250µm luminosity than low excitation galaxies. They attribute this to higher dust temperatures, and therefore increased star formation rates in HERGs. They also see evidence of excess star formation in HERGs compared to a control sample of radio-quiet galaxies. Once again, these findings are consistent with a picture in which HERGs are associated with lower mass, star forming systems and LERGs are predominantly located in the centres of massive elliptical galaxies, groups or clusters where star formation is quiescent or altogether quenched.

1.4.4 This work

In Chapter 5 of this thesis, we will apply an image-plane stacking technique, similar to the work of Hardcastle et al. (2010), Virdee et al. (2013) and Hardcastle et al. (2013), to an early H-ATLAS data release to study the relationship between AGN activity and star formation rates of the host galaxies. Once again, the radio galaxies will be identified in the VLA continuum surveys NVSS and FIRST. However, this work will make use of newly available, deep ancillary data over the SDP field to examine a deeper population of radio galaxies than previously possible. VIKING is a new near-infrared survey which probes deeper than UKIDSS by 1.2 magnitudes (Sutherland et al., in prep; Fleuren et al., 2012). Using this data we can identify the host galaxies of more distant radio sources than was possible with UKIDSS. A new INT survey also provides optical data over the field to a greater depth than the SDSS. Therefore,
photometric redshifts of fainter, more distant systems can be derived. A comparison of the parameter space probed by this and the previous analyses is presented in Table 5.1 in Chapter 5.

Using these new data sets, we will investigate the statistical FIR properties of radio-loud AGN out to $z = 5$. This is much higher than the redshift range of any previous study with H-ATLAS. While some SCUBA observations covered these high redshifts, our sample size is significantly larger and we probe a much wider range of radio luminosities. Therefore, we will examine evolution of the SFRs of radio galaxies over a previously unchartered epoch.
Chapter 2

Stacking techniques and data quality considerations

This chapter overviews the technique of ‘stacking’, the data analysis strategy explored in this thesis. Section 2.1 describes the basic concept of the technique and the specific application to 21 cm spectral observations and the image plane of far-infrared data. Sections 2.2 and 2.3 discuss how the quality of the data can affect the results of a spectral stacking analysis. Respectively, they explore the impact of radio frequency interference (RFI) in the H\textsc{i} data cube and the importance of spectroscopic redshift accuracy.

2.1 Overview of the stacking technique

Stacking is a relatively new data analysis technique which is fast becoming popular. The concept of ‘stacking’ itself is not new to astronomy; multiple exposures (or integrations) of a given field or object can be combined to achieve a better signal-to-noise ratio (S/N) and therefore a stronger detection. However, in the context of this thesis, ‘stacking’ instead refers to the process of combining together observations of
many different objects to achieve a statistical detection. While information on each individual galaxy is lost, stacking allows the average properties of the overall source population to be examined in an efficient manner. The benefit of this strategy is that it does not require individual detections of objects and so it can be used to extract information beyond the sensitivity limits of the data. Chapters 3 and 4 of this thesis apply the stacking concept to the 21 cm emission line of $\text{H}\,\text{i}$. The ‘spectral stacking’ technique used for this work is reviewed in section 2.1.1 below. In Chapter 5, stacking is applied to far-infrared imaging data. While the ‘image plane stacking’ technique is fully described in that chapter, we briefly discuss it here in section 2.1.2, by way of introduction.

2.1.1 Spectral stacking

$\text{H}\,\text{i}$ spectral stacking requires two independent data sets: (i) a spectral line radio cube, and (ii) an external source catalogue containing the positions and redshifts of galaxies in the field. For the work presented in this thesis, all radio data were collected by the Parkes 64 m radio telescope and the external source catalogues are provided by optical spectroscopic redshift surveys conducted on the Anglo-Australian Telescope. Specifically, these are the 2dFGRS (Colless et al. 2001; used in Chapter 3) and GAMA (Driver et al. 2009, 2011; used in Chapter 4). The reduced $\text{H}\,\text{i}$ data cube has axes of right ascension (RA), declination (Dec) and frequency ($\nu$), as illustrated in Figure 2.1. While no, or very few, galaxies have direct detections in the 21 cm data, the optical catalogue contains the positions of all optical sources known to exist in the observed field. Thus, the $\text{H}\,\text{i}$ spectrum can be extracted from the data cube along the pixel corresponding to each optical position.

Owing to the data gridding method we use (detailed in Barnes et al. 2001) and the fact that our sources are all unresolved, the spectral extraction technique which provides the most accurate flux densities involves a convolution with the beam, rather than simply extracting the intensity at the central pixel. Specifically, this involves using
CHAPTER 2. STACKING TECHNIQUES  2.1. OVERVIEW OF STACKING

Figure 2.1: A representation of three spectral channels from a 21 cm data cube. The yellow circle represents the position of an optical galaxy. The radio spectrum at this pixel position is extracted for stacking purposes.

the yaxis = point option within the MBSPect task in MIRIAD. This ensures that all the flux from a given galaxy is optimally incorporated and so maximises the S/N of the resulting stacked spectrum.

The left-most panel of Figure 2.2 shows the extracted H i spectra at the position of five individual optical sources. An RFI spike appears at \( \sim 1316 \) MHz. The spectra are then each shifted to rest frame, based on their optical redshifts:

\[
\nu_{\text{em}} = \nu_{\text{obs}}(1 + z), \tag{2.1}
\]

where \( \nu_{\text{obs}} \) is the observed frequency and \( \nu_{\text{em}} \) is the emitted frequency in the galaxy rest frame. This aligns all the spectra such that the channel(s) containing any H i signal are now at the rest frequency of the H i emission line (1420.406 MHz). The right-most panel of Figure 2.2 shows the same five spectra as in the left panel, but aligned at rest frame. The relative positions of the RFI spike now show that each
2.1. OVERVIEW OF STACKING  

CHAPTER 2. STACKING TECHNIQUES

Figure 2.2: Left: An example of five extracted spectra. Right: The same five spectra, each shifted to rest frame.

The spectrum has been shifted by a different amount.

The final step is to stack, or ‘co-add’ these $n$ aligned spectra by taking the weighted mean:

$$
\langle S \rangle_\nu = \frac{\sum_{i=1}^{n} (w_i S_{\nu,i})}{\sum_{i=1}^{n} w_i},
$$

where $S_{\nu,i}$ is the $i^{th}$ aligned spectrum, $\langle S \rangle_\nu$ is average flux density per channel (i.e. the stacked spectrum), and $w_i$ is the weight applied to each spectrum. In the above case, where we are stacking a flux density spectrum, the S/N of the stacked spectrum is optimised by the weighting scheme $w_i = \sigma_i^{-2}$, where $\sigma_i$ is the root mean squared (rms) noise level of the $i^{th}$ aligned spectrum. However, it should be noted that different weighting schemes may be more appropriate if the flux axis has been converted to an H I mass (or other) axis, as will be seen in Chapters 3 and 4.

Figure 2.3 shows the shape of the resultant stacked spectrum as progressively larger numbers of spectra ($N$) are co-added. For small $N$, the rms noise level is too large to allow a statistical detection. However, as $N$ increases, the rms noise decreases as $1/\sqrt{N}$ until a convincing statistical detection emerges (occurring at approximately
CHAPTER 2. STACKING TECHNIQUES  

2.1. OVERVIEW OF STACKING

Figure 2.3: The result of co-adding $N$ individual H\textsc{i} spectra, where $N$ is indicated at the top of each panel and increases from left to right and top to bottom.
$N = 316$ for the representative data shown in Figure 2.3). Note that the continued $\sqrt{N}$ increase in the S/N relies on the fact that the noise is mostly Gaussian. If this is not the case, such as in the presence of strong RFI, then the stacking technique may become less effective. This is discussed in more detail in section 2.2.2 below. Although individual galaxies may have a double-horned 21 cm profile, the process of averaging over many profiles of different widths results in a single-peaked shape. This stacked detection can now be used to extract information on the average properties of the input sample, as we will see in Chapters 3 and 4.

### 2.1.2 Image plane stacking

The basic principle behind the image plane stacking employed in Chapter 5 is similar to that of the spectral stacking technique described above, in that multiple non-detections can be combined such that a strong, average signal results. In this thesis, we apply this technique to 250 $\mu$m imaging maps from the Herschel SPIRE camera. Again, not all objects of interest were detected in this far-infrared band, however the source positions were identified prior to stacking using various external catalogues. There is no spectral axis associated with such imaging data. Rather, we simply extract a small cutout (e.g. 1 arcmin $\times$ 1 arcmin) from the full map, centred on the source position, as illustrated in the top frame of Figure 2.4. The central pixel of each of these image cutouts are then aligned and the weighted average of the flux of all pixels at each position is taken (middle frame of Figure 2.4), resulting in the stacked image (bottom frame of Figure 2.4).

For the analysis of Chapter 5, we in fact only require the value of the weighted average of the central pixel (blue cross in Figure 2.4). Therefore, we are essentially only performing a one dimensional stack. However, Figure 2.4 illustrates the full two-dimensional process for clarity. This strategy does result in stacked 250 $\mu$m detections, however we caution that the noise in this case, unlike that of the H I data, is highly non-Gaussian. This is due to considerable source confusion caused by the angular
Figure 2.4: A conceptual representation of the image-plane stacking process. The top frame shows the full 250 $\mu$m map of the H-ATLAS science demonstration phase field. The positions of three cutouts are shown (not to scale). The middle frame shows the same cutouts (colour-coded) after they have been aligned to have coinciding central pixels (blue cross). These are then stacked, resulting in the image in the bottom frame.
2.1. OVERVIEW OF STACKING

CHAPTER 2. STACKING TECHNIQUES

resolution of Herschel at 250 µm. Nonetheless, in Chapter 5 we show that image
plane stacking allows the extraction of information from source populations otherwise
difficult to detect.

2.2 RFI in H I data

Since a stacking analysis studies signals well below the detection level, the stacked
result is easily influenced by any signal not generated by the targeted sources. In
the context of H I spectral stacking, the analysis will be particularly sensitive to the
presence of RFI in the data and residual emission from strong continuum sources.
This section will characterise the RFI present in the Parkes observations, describe
the excision and mitigation techniques used and highlight the various ways RFI can
influence the results of a spectral stacking analysis. While the H I data will be de-
scribed in detail in Chapters 3 and 4, here we introduce the basic properties of the
observations to facilitate the RFI discussion.

2.2.1 Characteristics of RFI in Parkes data and mitigation
strategies

H I data for this work were collected using the 64 m single dish Parkes radio telescope
between 2008 and 2012 and are summarised in Table 2.1. The surveyed fields for
each observing run are listed in the table and are shown in Figure 1.3 in Chapter
1. Observations were conducted using two 64 MHz frequency bands. The lower band
was centred on 1285 MHz and the upper band on 1335 MHz. Thus the total frequency
coverage was 1253 - 1367 MHz, corresponding to H I redshifts over the range 0.04 <

z < 0.14.

To illustrate the RFI landscape of this data, Figure 2.5 shows some example spectra
extracted from the radio cubes and overlaid. The different types of RFI experienced
are labelled and will be described in more detail below. Table 2.2 lists the frequency channels worst affected by RFI. The upper band was relatively RFI-free except for strong, narrow-band satellite signals appearing at 1312, 1316, 1340 and 1350 MHz. The top panel of Figure 2.6 shows an example of the data collected in this band.

<table>
<thead>
<tr>
<th>Observation dates</th>
<th>Time (h)</th>
<th>Survey region</th>
<th>Relevant thesis chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008 Nov 20-25</td>
<td>34.25</td>
<td>SGP</td>
<td>3</td>
</tr>
<tr>
<td>2009 Aug 17-20 &amp; Sep 25-29</td>
<td>52.75</td>
<td>SGP</td>
<td>3</td>
</tr>
<tr>
<td>2010 Jan 31 - Feb 15</td>
<td>83.5</td>
<td>G09</td>
<td>4</td>
</tr>
<tr>
<td>2010 Apr 1-15</td>
<td>79.0</td>
<td>G09</td>
<td>4</td>
</tr>
<tr>
<td>2012 Mar 20-31</td>
<td>78.0</td>
<td>G15</td>
<td>4</td>
</tr>
<tr>
<td>2012 Nov 5-8</td>
<td>12.5</td>
<td>G09</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2.1: Details of all Parkes H I observations collected for this thesis.

The lower spectral band is afflicted by more severe RFI presence. Most notably, there is a very strong, broad-band signal (∼10 MHz wide) centred at roughly 1270 MHz. This was present in all beams and both polarisations. The suspected cause of this is the Compass Satellite Navigation system. 100 per cent of the data had to be flagged.
Table 2.2: The frequency and equivalent H\textsubscript{i} redshifts of the strongest RFI seen in Parkes observations at 1353-1267 MHz. \textsuperscript{a} Based on 2008-2009 observations. \textsuperscript{b} Based on 2010 and 2012 observations.

<table>
<thead>
<tr>
<th>Frequency (MHz)</th>
<th>Redshift</th>
</tr>
</thead>
<tbody>
<tr>
<td>1350</td>
<td>0.052</td>
</tr>
<tr>
<td>1340</td>
<td>0.060</td>
</tr>
<tr>
<td>1316</td>
<td>0.079</td>
</tr>
<tr>
<td>1312</td>
<td>0.083</td>
</tr>
<tr>
<td>1280.5-1260\textsuperscript{a}</td>
<td>0.109-0.127</td>
</tr>
<tr>
<td>&lt;1290\textsuperscript{b}</td>
<td>&gt;0.101</td>
</tr>
</tbody>
</table>

wherever this RFI was present.

During the 2008/2009 observations (used in Chapter 3), this signal was only present intermittently. The data files affected by this RFI, such as that shown in the middle panel of Figure 2.6, were identified and all information between 1262.4 - 1276.8 MHz in these files were discarded. Fortunately, 76 per cent of the lower band 2008/2009 observations were free of this RFI and therefore usable. A simple threshold clipping routine was sufficient to mask any remaining RFI in the rest of the data. The mean and rms noise level of this data were calculated within an RFI-free region and any data with an absolute value more than $3\sigma$ from the mean was flagged. The bottom panel of Figure 2.6 shows an example of a flagged data file. This routine successfully masked most RFI, reduced the rms of the final data cube and was sufficient for the analysis presented in Chapter 3.

However, the RFI landscape over the observed frequencies worsened with time. This was particularly true for the broad-band RFI region as the Compass system increased its operations. The observations taken in 2010 and 2012 are severely impacted by RFI, as seen in Figure 2.7. For the 2012 observations, all data below $\sim$1290 MHz was completely unusable. The simple threshold clipping routine was not sufficient to mask all RFI in this heavily-afflicted data. Fortunately, the sophisticated and rigorous RFI-excision software, aoflagger\textsuperscript{1}, has been developed. This was originally created to assist with RFI masking in LOFAR data (Offringa 2012) but has also been adapted

\textsuperscript{1}Available at http://sourceforge.net/p/aoflagger

56
Figure 2.6: Time-frequency plots of Parkes H\textsc{i} data taken in 2008 and 2009 in the 1335 band (top) and 1285 band (middle). The broad-band satellite RFI can be seen at \(\sim 1270\) MHz. The same 1285 data is shown in the bottom panel, with black indicating data flagged by the threshold clipping routine. The colour scale indicates flux density in Jy.
to work with interferometric and single-dish data from other facilities (A.R. Offringa, private communication). The AOFLAGGER pipeline employs a Gaussian high-pass filter, an iterative ‘SumThreshold’ method and a scale-invariant rank operator to identify and excise RFI (Offringa et al. 2010, 2012). This has proven to be one of the fastest and most robust techniques for masking data from a range of radio telescopes. The SumThreshold method iteratively fits to the time-frequency plane and is a new way of thresholding that does not require any manual input yet has a low probability of false RFI detection. The scale-invariant rank operator is then used to find adjacent frequency or time channels that could be impacted by RFI. This uses the principle that RFI behaviour should be scale-invariant. See Offringa et al. (2010) and Offringa et al. (2012) for an extensive description of the AOFLAGGER mechanics. All data from our 2010 and 2012 Parkes observations were passed through AOFLAGGER for use in the analysis of Chapter 4. The right-hand column of Figure 2.7 demonstrates the success of the algorithm in identifying and flagging the RFI.

To highlight the temporal increase in the presence of RFI across the observed frequency range, Figure 2.8 compares the percentage of data flagged in each spectral channel between the 2008/2009, 2010 and 2012 (G15 only) observations. To ensure a fair comparison, the 2008/2009 data used to generate this plot have also been passed through AOFLAGGER. However, only the minimal threshold clipping method was required for the analysis of Chapter 3. It is clear that the RFI landscape at these frequencies worsens significantly over time, particularly below 1290 MHz. Note, also, that the target fields of the 2010 and 2012 observations are equatorial and are therefore most susceptible to RFI from geostationary satellites. By 2012, nearly 100 per cent of the data in these channels are flagged, limiting the analysis of Chapter 4 to $z \lesssim 0.1$.

Along with man-made RFI, strong continuum emission from astronomical sources is also considered a contaminant for the purpose of this work. A standing wave is generated when such a continuum source enters either the primary beam or sidelobes (Barnes et al. 2001) and manifests as a ripple in the spectral data, as seen in Figure 2.5.
Figure 2.7: Example time-frequency plots from 2010 (top) and 2012 (bottom) observations in the lower spectral band (1253 to 1317 MHz). On the left are the raw data files, with the broad-band satellite RFI seen. On the right are the same files after RFI masking by the AOFLAGGER software (Offringa et al. 2010, 2012). Magenta indicates flagged data. The colour scale indicates flux density in Jy.
Additionally, the different response pattern of each frequency band creates a spectral discontinuity at 1310 MHz (the intersection of the two bands) near strong continuum sources. These effects will contribute to a non-Gaussian noise behaviour in the data, which is not ideal for stacking. To mitigate against these disruptive signals, we identified the strong continuum sources in the observed field. We then excluded from further analysis any target galaxies within a beamwidth of the continuum sources. Since continuum emission was not the dominant source of RFI, this simple method was sufficient to prevent significant degradation of the stacked detection.

2.2.2 Impact of RFI on stacking analyses

To highlight the importance of the above RFI mitigation, we now discuss the undesirable repercussions of RFI on a spectral stacking analysis. As we will see, RFI can restrict the accessible spectral channels, limit the achievable S/N, degrade the quality of the stacked detection or even prevent a statistical detection entirely.
CHAPTER 2. STACKING TECHNIQUES  2.2. DATA QUALITY AND RFI

If RFI creates large outlying values, a stacked detection may not be recoverable at all using normal parametric statistics. In such cases, the use of non-parametric (e.g. median) statistics may produce a better outcome. Although median statistics should be approximately equivalent to mean statistics for low S/N and large samples, if direct, high S/N detections are present in the data, median statistics will significantly underestimate the average $H\text{I}$ mass of the sample.

Consequently, if a certain spectral channel is strongly dominated by RFI, all data in that channel must be completely flagged. Any galaxies with $H\text{I}$ redshifts corresponding to the RFI-dominated frequency must then be excluded from the stacking process. If RFI dominates many channels, such as in Figure 2.7, entire redshift ranges may be inaccessible. Since lower frequencies (and hence higher redshifts) are becoming increasingly dominated by satellite RFI, this can limit our ability to probe the evolution of $H\text{I}$ properties.

Furthermore, the flagging of many channels will reduce the number of sources available for stacking and hence limit the achievable S/N. This is particularly problematic when binning the data by different physical properties to examine trends in the $H\text{I}$ content. An insufficient number of galaxies may remain per bin to achieve a statistically-significant detection.

The achievable S/N may also be limited by the presence of a ‘noise floor’ caused by the non-Gaussian properties of the RFI (and residual continuum emission). That is, the rms noise level of the co-added spectrum may cease to decrease with $1/\sqrt{N}$. This is demonstrated in Figure 2.9. The achievable rms noise level of the stacked spectrum is shown when careful RFI mitigation has been employed and when some RFI has been intentionally left unflagged. Insufficient RFI excision not only results in a higher overall rms in the data (i.e. the vertical offset between the two trials) but also causes a deviation from Gaussian behaviour. The expected behaviour of purely Gaussian noise is indicated by the dashed line with a slope of $-0.5$. Once the non-Gaussian RFI components begin to dominate the noise, the open points cease to track this line downwards. The method for generating the error bars for this plot is described in
Figure 2.9: The rms noise level of the co-added \( \text{H} \text{i} \) spectrum when \( N \) galaxies are stacked together. The dashed lines indicate Gaussian behaviour and have slopes of -0.5. The solid points have been generated using careful RFI mitigation, while insufficient RFI mitigation was used to create the open points.

Chapter 3.

Lastly, the presence of RFI in \( \text{H} \text{i} \) data can result in an insufficient baseline removal during the data reduction process. This can generate a complicated baseline in the stacked spectrum. Figure 2.10 shows an example of such a stacked \( \text{H} \text{i} \) spectrum before baseline subtraction. The shape of the baseline can sometimes be difficult to fit and remove, and may make it problematic to define the bounds of the stacked detection profile. This may result in large variations in the integrated flux (or \( \text{H} \text{i} \) mass) recovered. A curious trend found in many of the stacked spectra in Chapters 3 and 4 is an asymmetric stacked profile. This asymmetry manifests itself as a tail towards lower frequencies. While we have not found an explanation for this, inadequate baseline subtraction may be partially to blame.
2.3 Redshift error

Along with RFI, another important data quality consideration for a stacking analysis is the accuracy of the spectroscopic redshifts employed. Here we demonstrate how redshift errors can alter the properties of the stacked detection.

For spectral stacking, correct alignment of the spectra at rest frame will depend upon how closely the optical redshifts match the true redshifts of the gas within the galaxies. If the offset is too large, it is possible that the incorrect alignment will completely exclude a statistical detection. For example, Figure 2.11 compares the results of stacking ∼2,000 galaxies using spectroscopic redshifts from GAMA, and the same galaxies using photometric redshifts derived from SDSS photometry (Driver et al. 2011). The spectroscopic redshifts, which have a 0.0003 (∼90 km s\(^{-1}\)) associated error, produce a strong statistical detection. However, the photometric redshifts, which have an average error of 0.026 (∼7,800 km s\(^{-1}\)), unsurprisingly produce no detection whatsoever. While this is obviously the extreme case, less severe redshift errors can still impact the shape of the stacked spectrum.
2.3. REDSHIFT ERROR

CHAPTER 2. STACKING TECHNIQUES

Figure 2.11: An example of a stacked H i spectrum produced using spectroscopic redshifts (black) and photometric redshifts (red). No stacked detection is is seen in the latter case. The spectroscopic redshifts are from GAMA and the photometric redshifts were derived using SDSS photometry (Driver et al. 2011).

We now use data from the HIPASS source catalogue (HICAT) to examine the influence of spectroscopic redshift errors on the properties of the stacked spectrum. This analysis, using real data, is similar and complementary to that of Maddox et al. (2013) who use simulated H i profiles. The black curve in the top panel of Figure 2.12 shows the co-added H i spectrum of 1,000 galaxies randomly selected from HICAT. Each of these have strong, individual 21 cm detections and the redshifts of these objects have been defined using the H i profiles. Therefore, the HICAT redshifts can be considered to have zero error for the purposes of stacking, since the re-aligned profiles will each be centred exactly at rest frame. The coloured curves in Figure 2.12 show the stacked spectra of the same 1,000 galaxies as progressively larger Gaussian errors are added to the redshifts. With increasing error, the spectrum broadens out and the peak S/N decreases. In each case, the bounds of the stacked profile can be visually identified and the integrated flux within these bounds measured. These fluxes are reported in Table 2.3 and plotted in the lower panel of Figure 2.12. Note that the errors on the integrated fluxes were calculated using a bootstrapping method which will be described in Chapter 3.

Above a $cz$ error of 150 km s$^{-1}$, the stack broadens to an extent where integrated flux
is no longer conserved. Although, the reduction in the total flux is only $\sim 5\%$, even with a $cz$ error of 500 km s$^{-1}$. This difference is mainly due to difficulty in choosing the most appropriate signal bounds since the signal in the outer channels dips below the noise and may be more strongly influenced by the baseline subtraction. This means that flux has been excluded from the average due to misalignment. We note that the median profile width of the individual input spectra also corresponds to 150 km s$^{-1}$.

These findings are in close agreement with the work of Maddox et al. (2013). These authors simulate various H i profiles, with properties based on ALFALFA H i detections, and investigate the impact of redshift error on the stacked profile shape. They too find that the redshift errors begin to dominate the broadening of the stacked profile once these errors are greater than the median profile widths of the individual input H i lines (also $\sim 150$ km s$^{-1}$ for their sample). However, they recover identical integrated stacked fluxes regardless of redshift error. This is likely due to the relative ease of defining clear profile bounds to their stacked signals. Their simulated noise is purely Gaussian and is not affected by the additional baseline and RFI issues that impact our real data.

The spectroscopic redshifts we employ for the stacking analyses of Chapters 3 and 4 have quoted accuracies of 85 and $\sim 100$ km s$^{-1}$, respectively. Since the median profile width of the galaxies we consider is expected to be roughly 150 km s$^{-1}$ (predicted by the Tully-Fisher relation; see Figure 3.6 in Chapter 3), we conclude that spectroscopic redshift errors will not contribute significantly to any broadening of our stacked spectra.

In light of these results, it seems that photometric redshifts will not be sufficient for stacking analyses since they cannot be derived to an accuracy of $< 150$ km s$^{-1}$. This will be an important consideration for large H i surveys with the SKA and pathfinder instruments. The implications are that complementary spectroscopic redshift surveys will be required to facilitate stacking tests with this future H i data. Indeed, many of these H i projects have been designed to overlap with available or planned spectroscopic surveys.
Figure 2.12: Top: The stack of 1,000 randomly-selected galaxies from the HIPASS source catalogue as various levels of Gaussian error are introduced to the redshifts. Bottom: The integrated flux of the corresponding stacks.
Table 2.3: Properties of stacked HICAT spectra as artificial redshift error is introduced.

<table>
<thead>
<tr>
<th>$z$ error ($\times 10^{-3}$)</th>
<th>$cz$ error (km s$^{-1}$)</th>
<th>Profile width (km s$^{-1}$)</th>
<th>Integrated flux ($\times 10^{-2}$ Jy MHz)</th>
<th>Integrated S/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>559</td>
<td>5.56 ± 0.06</td>
<td>93</td>
</tr>
<tr>
<td>0.3</td>
<td>90</td>
<td>796</td>
<td>5.54 ± 0.08</td>
<td>69</td>
</tr>
<tr>
<td>0.5</td>
<td>150</td>
<td>1011</td>
<td>5.50 ± 0.08</td>
<td>66</td>
</tr>
<tr>
<td>0.8</td>
<td>250</td>
<td>1178</td>
<td>5.36 ± 0.09</td>
<td>58</td>
</tr>
<tr>
<td>1.7</td>
<td>500</td>
<td>2408</td>
<td>5.32 ± 0.14</td>
<td>38</td>
</tr>
</tbody>
</table>
Chapter 3

Detection of H I in distant galaxies using spectral stacking

Foreword and context

Stacking allows information to be extracted from the limits of a data set. Chapter 2 of this thesis provided a general introduction to the spectral stacking technique. In this chapter, we develop and describe this technique in detail. We apply the technique to the 21 cm emission line of neutral hydrogen to study the H I properties of galaxies within a significantly larger volume of the Universe than would otherwise be possible. This allows us to determine whether there has been any evolution in the cosmic density of neutral hydrogen gas over the past $\sim$2 billion years, an important constraint on galaxy evolution models. We also show that stacking allows us to examine how the H I content of galaxies changes as a function of various physical properties, even for galaxies below the detection limit of the radio data used.

This chapter is composed of a paper published in the Monthly Notices of the Astronomical Society, Volume 433, Issue 2, p.1398-1410 and is reproduced here exactly as published but reformatted into a thesis layout.
Authors and affiliations

J. Delhaize (1), M. J. Meyer (1,2), L. Staveley-Smith (1,2), B.J. Boyle (2,3)

(1) International Centre for Radio Astronomy Research (ICRAR), M468, University of Western Australia, 35 Stirling Hwy, WA 6009, Australia

(2) ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), Australia

(3) CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1710, Australia

3.1 Abstract

Using the Parkes radio telescope, we study the 21 cm neutral hydrogen (H\textsc{i}) properties of a sample of galaxies with redshifts \(z < 0.13\) extracted from the optical Two-Degree-Field Galaxy Redshift Survey (2dFGRS). Galaxies at \(0.04 < z < 0.13\) are studied using new Parkes observations of a 42 deg\(^2\) field near the South Galactic Pole (SGP). A spectral stacking analysis of the 3,277 2dFGRS objects within this field results in a convincing 12 \(\sigma\) detection. For the low-redshift sample at \(0 < z < 0.04\), we use the 15,093 2dFGRS galaxies observed by the H\textsc{i} Parkes All-Sky Survey (HIPASS) and find a 31 \(\sigma\) stacked detection. We measure average H\textsc{i} masses of \((6.93 \pm 0.17) \times 10^9\) and \((1.48 \pm 0.03) \times 10^9 \, h^{-2} M_\odot\) for the SGP and HIPASS samples, respectively. Accounting for source confusion and sample bias, we find a cosmic H\textsc{i} mass density of \(\Omega_{\text{H} \textsc{i}} = (3.19^{+0.43}_{-0.39}) \times 10^{-4} \, h^{-1}\) for the SGP sample and \((2.82^{+0.30}_{-0.39}) \times 10^{-4} \, h^{-1}\) for the HIPASS sample. This suggests no (12\% per cent) evolution in the cosmic H\textsc{i} density over the last \(~1\, h^{-1}\) Gyr. Due to the very large effective volumes, cosmic variance in our determination of \(\Omega_{\text{H} \textsc{i}}\) is considerably lower than previous estimates. Our stacking analysis reproduces and quantifies the expected trends in the H\textsc{i} mass and mass-to-light ratio of galaxies with redshift, luminosity and colour.
To obtain a complete picture of galaxy evolution, it is crucial to understand how the cold gas content of galaxies varies with cosmic time. Cold gas ($< 10^4 \text{K}$), largely in the form of neutral hydrogen (H\textsc{i}), is supplied through various accretion, merger and feedback processes (eg. Kereš et al. 2005) and is later available to condense into massive molecular clouds and stars. Without this gas, star formation is quenched and evolution continues only passively. To improve our understanding of the evolution in quantities such as star formation rate density (Madau et al. 1996, Hopkins & Beacom 2006), it is therefore fundamental to also improve our understanding of changes in the cosmic gas density.

Models that attempt to predict the variation of cosmic gas density with redshift diverge in their predictions. This is partly due to the wide spatial range required to track the requisite gas physics and partly because of uncertainty in the physics of galaxy formation, in particular the action of feedback from stars and AGN (eg. Somerville et al. 2001; Cen et al. 2003; Nagamine et al. 2005; Power et al. 2010; Lagos et al. 2011). Thus, observational constraints are vital.

A number of observational techniques exist for tracing H\textsc{i} gas within galaxies. At high redshifts, the Gunn-Peterson effect applied to damped Lyman $\alpha$ (DLA) systems can be exploited. Prochaska et al. (2005) and Prochaska & Wolfe (2009) used Sloan Digital Sky Survey (SDSS) spectra to trace systems with H\textsc{i} column densities above $2 \times 10^{20} \text{cm}^{-2}$ at $z > 1.7$. Though they find a statistically significant detection of evolution in the cosmic H\textsc{i} density within $2.2 < z < 3.5$, interpretation of their results is problematic due to systematic effects such as dust obscuration and gravitational lensing. At $z < 1.65$ the Lyman $\alpha$ transition is observed at ultraviolet wavelengths, making it difficult to trace with the limited time available on space-based observatories. Rao et al. (2006) were able to contribute in this regime by optically identifying DLA systems through metal absorption lines in SDSS data. They found no evidence of evolution in the cosmic H\textsc{i} density of DLAs within $0.5 < z < 5$, although their
results are limited by statistical uncertainty.

The preferred method for tracing HI in the local Universe ($z \approx 0$) is via the direct detection of the neutral hydrogen $21\text{cm}$ hyperfine emission line. With the wide instantaneous field of view provided by multibeam receivers on large radio telescopes, it is feasible to conduct blind, all-sky surveys at $21\text{cm}$. This has facilitated the detection of HI in large numbers of galaxies. A prime example of this is the HI Parkes All-Sky Survey (HIPASS), conducted with the 64 m Parkes radio telescope in New South Wales, Australia (Barnes et al., 2001; Staveley-Smith et al., 1996). HIPASS detected HI within 5,317 galaxies at $0 < z < 0.04$ (Meyer et al., 2004; Wong et al., 2006), allowing the construction of the most reliable and complete HI mass function then available, and an estimate of the HI mass density for the local Universe (Zwaan et al., 2005). Similarly, the Arecibo Legacy Fast ALFA (ALFALFA) survey aims to detect $>30,000$ galaxies at $21\text{cm}$ out to $z = 0.06$ over a sky area of $\sim7,000\text{deg}^2$ (Giovanelli et al., 2005). Using the 40 per cent ALFALFA catalogue, Martin et al. (2010) find a cosmic HI mass density 16 per cent higher than that derived by Zwaan et al. (2005). They conclude that this discrepancy is caused by an under-representation in the HIPASS catalogue of the rare, high-mass $[9.0 < \log(M_{\text{HI}}/M_\odot) < 10.0]$ galaxies due to sensitivity limits.

Deeper, blind HI surveys are underway, such as the Arecibo Ultra Deep Survey (AUDS). This survey aims to directly detect significant numbers of field galaxies out to $z = 0.16$ to probe the evolution of cosmic HI properties (Freudling et al., 2011). However, the survey time required to achieve sufficient sensitivity at these higher redshifts restricts the observations to a narrow field ($1/3\text{deg}^2$ for AUDS), thus potentially introducing a strong cosmic variance bias.

To date, the deepest detections of $21\text{cm}$ emission in individual galaxies have been made at $z \approx 0.2$ by Catinella et al. (2008) using pointed observations with the sensitive 305 m Arecibo telescope, and by Zwaan et al. (2001) and Verheijen et al. (2007) using long integrations of cluster galaxies with the Westerbork Synthesis Radio Telescope (WSRT). While future instruments such as the Square Kilometre Array (SKA) and its
pathfinders should have the ability to detect 21 cm emission to much higher redshifts over large sky areas, these will not be available for a number of years.

However, methods other than direct, individual detections are available to push the current redshift limit of H I observations. In particular, stacking is a useful method for probing H I within large samples of distant galaxies. Stacking is the process of co-adding individual non-detections in an attempt to boost the signal-to-noise ratio (S/N) of the data and thereby recover a more significant statistical detection.

Zwaan (2000), Chengalur et al. (2001) and Lah et al. (2009) demonstrated that stacking can be successfully used to examine the environmental influence on the H I properties of cluster galaxies out to $z = 0.37$. Lah et al. (2007) were the first to apply this co-adding method to field galaxies to study the evolution of the cosmic H I density. They examined 154 star-forming galaxies selected in Hα at $z = 0.24$ using 21 cm data from the Giant Metrewave Radio Telescope (GMRT); however, averaged detections were $3\sigma$ at best.

Fabello et al. (2011a) and Fabello et al. (2011b) used similar stacking methods to examine the impact of bulge presence and AGN activity on the H I content of galaxies with stellar masses greater than $10^{10} M_\odot$. They demonstrated that when using relatively low redshift ($0.025 < z < 0.05$) H I data and a large sample of galaxies, strong stacked detections were possible, even for the subsample of galaxies individually undetected at 21 cm.

The related technique of ‘intensity mapping’, which uses cross-correlation of radio and optical redshift data, has been used by Pen et al. (2009), Chang et al. (2010) and Masui et al. (2013) to detect H I in samples at even higher redshift.

The work presented here aims to constrain the cosmic density of H I out to $z = 0.13$ and demonstrate the accuracy of the stacking method over this redshift range. We have employed the fast survey speed of the Parkes telescope to collect 21 cm observations of a very large sample of field galaxies from the Two-Degree-Field Galaxy Redshift Survey (2dFGRS). The sample selection is not biased by environment, star
formation signatures, or any particular physical characteristic other than the optical magnitude limits of the 2dFGRS. By using the wide but low-redshift HIPASS data, in combination with new, spectrally deep Parkes observations of a smaller subsection of the 2dFGRS footprint, we can examine the HI content of galaxies over the full range $0 < z < 0.13$.

Section 3.3 of this paper presents the optical redshift survey and the radio observations used in the analysis. The source selection and stacking procedures are described in Section 3.4. In Section 3.5.1 we present an analysis of noise behaviour in the HI data. Section 3.5.2 discusses the influence of source confusion on our results and illustrates how we correct for this. Section 3.5.3 describes the HI mass density calculation and Section 3.5.4 considers the impact of cosmic variance. Trends in HI properties with redshift, luminosity and colour are measured in Section 3.5.5. Finally, we present our conclusions in Section 3.6. We have assumed a $\Lambda$ cold dark matter cosmology with a reduced Hubble constant $h = H_0/(100\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}) = 1.0$, $\Omega_\Lambda = 0.7$ and $\Omega_M = 0.3$.

### 3.3 The data

#### 3.3.1 The 2dFGRS optical catalogue

The basic principle of the HI stacking analysis employed here is to identify the positions and redshifts of galaxies using an external source catalogue containing optical redshifts, extract the HI information at these coordinates and then co-add the data. For the present analysis, the positions and redshifts are provided by the 2dFGRS. The 2dFGRS is a spectroscopic survey of over 250,000 galaxies conducted with the 2dF multifibre spectrograph on the Anglo-Australian Telescope. This survey covers two large strips, one in the Northern Galactic hemisphere and one in the south, as well as a number of small, random fields. The total sky coverage is $\sim 2,000\,\text{deg}^2$. See Colless et al. (2001) for further details of the surveyed regions.
The input photometric source catalogue for this survey comes from an adapted catalogue of Automated Plate Measuring (APM) machine scans of the Southern Sky Survey taken with the UK Schmidt Telescope. These target sources have extinction-corrected magnitudes brighter than $b_J = 19.45$. All APM $b_J$ magnitudes used in this paper are extinction-corrected, but not $(k+e)$-corrected for bandpass and evolution. These corrections will be incorporated into the corresponding redshifted luminosity density.

As described in Colless et al. (2001), the redshifts were estimated using a combination of automated absorption spectral shape fitting, emission line identification and manual input. Each redshift was assigned a quality factor $Q$ between 1 (no reliable redshift available) and 5 (very reliable redshift). We find that sources with $Q \geq 3$ are sufficient for our stacking analysis. Colless et al. (2001) state that $Q \geq 3$ redshifts have a root-mean-square (rms) uncertainty of $85\,\text{km}\,\text{s}^{-1}$ and can be used with 98.4 per cent reliability.

### 3.3.2 H I observations and data

#### 3.3.2.1 South Galactic Pole observations

We conducted a 21 cm survey with the 64 m Parkes radio telescope over a $42\,\text{deg}^2$ ($7^\circ \times 8^\circ$) sky field centred on right ascension (RA) $00^h42^m00^s$ and declination (Dec.) $-29^\circ00'00''$ (J2000), close to the South Galactic Pole (SGP). The region was chosen as it contains an overdensity of available 2dFGRS spectra (described further in Section 3.4.1). This provides a high number of spectra per square degree available for a stacking analysis. 87 h of telescope time was used for observations of this field. These were conducted on 2008 November 20-25, 2009 August 17-20 and 2009 September 25-29 with an estimated 52 h of on-source integration.

The observations used the multibeam correlator with two adjacent frequency bands, each with 1024 spectral channels over a bandwidth of 64 MHz. The bands were centred
on 1285 and 1335 MHz with a 14 MHz overlap to reduce the impact of the bandpass shape. Observations in the two frequency bands were taken independently. The full frequency coverage of the Parkes data is thus 1253-1367 MHz over two polarizations. We therefore have the potential to detect galaxies with 21 cm redshifts in the range $0.0391 < z < 0.1336$. The resulting channel spacing is 62.5 kHz.

### 3.3.2.2 HIPASS observations

We also present a study using all overlapping data from the 2dFGRS catalogue and HIPASS. HIPASS was also conducted with the Parkes 64 m dish using the 21 cm multibeam receiver (Staveley-Smith et al., 1996) and covers the entire sky south of Dec. $+25^\circ$. The HIPASS velocity range is $-1280 < cz < 12700 \text{ km s}^{-1}$, corresponding to a frequency coverage of 1362.5–1426.5 MHz and redshifts below $z = 0.0423$. Observations of the southern field (below Dec. $+2^\circ$) were completed over the period from 1997 February to 2000 March (Barnes et al., 2001). The northern field (Dec. $+2^\circ$ to $+25^\circ$) was observed from 2000 to 2002 (Wong et al., 2006). The HIPASS and SGP data sets overlap by only 4.5 MHz ($\Delta z = 0.003$) in spectral range, making the two data sets highly complementary.

### 3.3.2.3 Telescope scanning mode

Both data sets were collected using a similar Parkes scanning mode, where the telescope is actively scanned across the required sky area at a rate of $1^\circ \text{ min}^{-1}$. While HIPASS only used scans in declination, the SGP observations used a ‘basket-weave’ pattern, with the entire field also covered by right ascension scans. This reduces the effect of negative sidelobes around bright continuum sources created by the bandpass correction. Using this technique, the SGP field was completely surveyed 15 times in each spectral band within the available time frame.
3.3.2.4 Data reduction and interference mitigation

Bandpass removal and calibration were conducted identically for both data sets using the \textsc{livedata}\textsuperscript{1} reduction package. See Barnes et al. (2001) for a detailed explanation of these processes. The data were smoothed to reduce the effects of ringing created by strong Galactic emission and/or strong radio frequency interference (RFI), such as from satellites. Hanning smoothing was applied to the SGP data. HIPASS utilized Tukey 25 per cent smoothing. However, we have also applied Hanning smoothing to the data for consistency. The spectral resolution after smoothing is 125\,kHz (26.4\,km\,s\textsuperscript{-1} at \( z = 0 \)).

The use of robust statistics in HIPASS data processing (see Section 3.3.2.5 below) was sufficient to mitigate against the majority of RFI. However, the lower frequencies probed during SGP observations suffer from greater RFI. To improve the quality of the SGP data, we applied an automated RFI masking routine to the reduced, smoothed data. The mean and rms noise levels were calculated in an RFI-free region and the resultant 3\( \sigma \) clipping threshold was set at 250.5\,mJy. All pixels with absolute deviations above this threshold were masked. This was not sufficient to completely mitigate against the presence of strong satellite RFI in the frequency range 1262.4–1276.8\,MHz. It was not consistently present and appears in 24 per cent of the observations in the 1285\,MHz band. Therefore, the full range of affected channels was completely masked in all data files containing this RFI. This method was successful in recovering only the good quality data in the broad-band RFI zone. In addition, all data in channels around 1280, 1300, 1312, 1316, and 1350\,MHz were masked as they were also completely contaminated by satellite RFI. The resulting fraction of flagged data per channel is shown in Fig. 3.1.

\textsuperscript{1}\textsc{livedata} is supported by the Australia Telescope National Facility and is available at http://www.atnf.csiro.au/computing/software/livedata/
3.3.2.5 Gridding

The reduced Parkes data were gridded using GRIDZILLA\textsuperscript{2}. Robust median averaging was used to minimize the impact of RFI on the final cubes and the two polarizations were added. Pixel sizes in the gridded data are each $4 \times 4$ arcmin\textsuperscript{2}. For more details, see Barnes et al. (2001).

The SGP field is sufficiently small that it is reasonable to produce a single cube in each frequency band. To account for the band overlap, each cube was cut at 1310 MHz and the two were then concatenated together such that the frequency coverage was continuous. The rms noise per channel in the RFI-free part of the final data cube is $8.5 \text{ mJy beam}^{-1}$. The flux density scale is calibrated against 1934-638 and Hydra A. Calibrations are consistent to within $\sim 2$ per cent.

Since the sky coverage of the HIPASS data is so large, the southern region is split into 388 separate sky fields, each measuring $8 \times 8 \text{ deg}^2$ (Barnes et al., 2001). The northern HIPASS field consists of 102 $8 \times 8 \text{ deg}^2$ cubes and 48 $8 \times 7 \text{ deg}^2$ cubes (Wong et al.,

\textsuperscript{2}GRIDZILLA is supported by the Australia Telescope National Facility and is available at http://www.atnf.csiro.au/computing/software/livedata/
3.4. ANALYSIS

CHAPTER 3. HI IN DISTANT GALAXIES

2006). The rms noise level of the HIPASS data is 13.3 mJy (Barnes et al., 2001), which is 56 per cent higher than the SGP data. The effective spatial resolution of both the gridded SGP and HIPASS data sets is 15.5 arcmin.

3.4 Analysis

3.4.1 SGP source selection

There are 4,240 2dFGRS sources in the SGP field with high-quality redshifts \((Q \geq 3)\) covered by the 21 cm observations. However, not all of these are suitable for use in this stacking analysis. All channels within 2 MHz of the extremes of the spectral range were discarded due to the bandpass shape, and the corresponding sources rejected. Therefore, we have a final redshift coverage of \(0.0405 < z < 0.1319\) for this SGP sample. Additionally, sources were rejected from the sample if their redshifted HI signatures fell at frequencies with 100 per cent RFI occupancy in the data (see Fig. 3.1).

Continuum sources must be avoided as they can induce standing waves in the data (Barnes et al., 2001) and create a discontinuity between the two spectral bands. Using the NRAO VLA Sky Survey (NVSS) source catalogue by Condon et al. (1998), the 17 continuum sources in the SGP field with 1.4 GHz flux densities greater than 200 mJy were identified. The 840 2dFGRS galaxies with angular separations from these continuum sources of less than the spatial resolution of the data were discarded from the sample.

Our final 2dFGRS sample therefore contains 3,277 sources suitable for stacking within the SGP field. The heliocentric redshift distribution of these sources is shown in Fig. 3.2. Many galaxies are distributed around \(z = 0.11\) and the average redshift of the sample is 0.096. This ‘clumpiness’ in redshift space is not attributable to a single galaxy cluster, but seems to be probing genuine cosmic structure. For example, the
Figure 3.2: The heliocentric redshift distributions of the optical samples considered here. The black line shows the 3,277 2dFGRS sources within the observed SGP region and spectral range suitable for stacking. The mean redshift is 0.096. Overlaid in green is the redshift distribution of the full 2dFGRS catalogue across this redshift range, scaled by the ratio of the SGP to total 2dFGRS field areas. The red line shows the distribution of the 15,093 2dFGRS sources with available HIPASS spectra. The average redshift of this sample is 0.029.

2dFGRS Percolation-Inferred Galaxy Group (2PIGG) catalogue by Eke et al. (2004) identifies 17 galaxy groups with 10 or more members at \( z \approx 0.11 \) within the SGP region.

For comparison, the redshift distribution of the full 2dFGRS catalogue (across the relevant \( z \) range) is also shown in Fig. 3.2. It has been scaled down by the ratio of the full 2dFGRS and SGP field areas so the shape of the two distributions can be compared. The full 2dFGRS distribution is far more uniform across redshift and the ratio of the total source counts reveals that the SGP field contains an \( \sim 30 \) per cent overdensity at these redshifts. It is therefore evident that large-scale clustering is present in the chosen SGP field, highlighting the importance of cosmic variance considerations in such an analysis. This is discussed in detail in Section 3.5.4.
3.4.2 HIPASS source selection

The full HIPASS completely covers the 2dFGRS footprint, but is limited in redshift coverage. There are 15,152 2dFGRS sources with reliable redshifts ($Q \geq 3$) within the spectral range of the HIPASS data that are appropriate for stacking. This number excludes any sources with $z < 0.0025$ ($v < 750 \text{ km s}^{-1}$) to minimize contamination from the Galaxy and misclassified stars. We have removed the contribution of any Galactic emission from all spectra by masking frequency channels between 1418.4 and 1422.4 MHz. Our final HIPASS sample consists of 15,093 2dFGRS sources.

The redshift distribution is shown in Fig. 3.2. According to Barnes et al. (2001), the contribution from continuum sources has already been suppressed in the HIPASS data by fitting a weighted average of the spectral shape near these sources and subtracting from all subsequent spectra.

3.4.3 Stacking process

3.4.3.1 Extracting H I spectra

For each 2dFGRS source in our final sample, we extracted the full 21 cm spectrum from the data cube using an optimal weighting according to the beam shape, centred on the pixel corresponding to the RA and Dec. defined in the optical catalogue. This extraction strategy optimizes the S/N of the stacked spectrum but results in an expanded effective beam width of 21.2 arcmin ($\sim 1.6 h^{-1} \text{ Mpc at } z = 0.1$) for the SGP data and 21.9 arcmin ($\sim 0.6 h^{-1} \text{ Mpc at } z = 0.03$) for the HIPASS data.

We find one direct H I detection of a 2dFGRS source in the SGP sample with an S/N of 7.2. The H I spectrum of this source is shown in Fig. 3.3 and the measured parameters are reported in Table 3.1. There is no firm evidence of H I detection in the rest of the sample.

In contrast to this, there are numerous, significant individual detections in the HIPASS spectra owing to the fact that HIPASS surveys the more local Universe. The HIPASS
Figure 3.3: The 21 cm flux density spectrum of the only directly detected 2dFGRS source in the SGP field. The gap in the spectrum is due to RFI flagging. The dashed red lines show the nominal width of the profile and the vertical dotted line indicates rest frequency.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA (2dFGRS)</td>
<td>00:40:04.28</td>
</tr>
<tr>
<td>Dec. (2dFGRS)</td>
<td>-30:35:45.60</td>
</tr>
<tr>
<td>z (2dFGRS)</td>
<td>0.0500</td>
</tr>
<tr>
<td>$\nu_{\text{centre}}$ (MHz)</td>
<td>1352.8</td>
</tr>
<tr>
<td>$\Delta v$ (km s$^{-1}$)</td>
<td>329</td>
</tr>
<tr>
<td>S/N</td>
<td>7.2</td>
</tr>
<tr>
<td>$S_{\text{int}}$ (mJy MHz)</td>
<td>24</td>
</tr>
<tr>
<td>$M_{\text{H}<em>1}$ ($\times 10^9 h^{-2} M</em>{\odot}$)</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 3.1: Measured parameters of the SGP detection in Fig. 3.3.
source catalogue (HICAT) created by Meyer et al. (2004) (southern field) and Wong et al. (2006) (northern field) identifies 5,317 HI sources in total. 226 of these coincide spatially and spectrally with 2dFGRS galaxies in our sample such that they will contribute to the extracted HI spectrum.

The HI mass per unit frequency in the observed frame (hereafter referred to as the mass spectrum $M_{\text{HI},\nu_{\text{obs}}}$) can be calculated from:

$$
\left( \frac{M_{\text{HI},\nu_{\text{obs}}}}{M_{\odot} \text{MHz}^{-1}} \right) = 4.98 \times 10^7 \left( \frac{S_{\nu_{\text{obs}}}}{\text{Jy}} \right) \left( \frac{D_L}{\text{Mpc}} \right)^2 ,
$$

where $S_{\nu_{\text{obs}}}$ is the observed-frame HI flux density and $D_L$ is the luminosity distance.

### 3.4.3.2 Stacking spectra

The first step in the stacking process is to align all extracted spectra at rest frequency (1420.406 MHz). The spectral axis is converted from the observed to the emitted frame via $\nu_{\text{em}} = \nu_{\text{obs}}(1 + z)$. To conserve the total mass (i.e. $\int M_{\text{HI},\nu_{\text{em}}} d\nu_{\text{em}} = \int M_{\text{HI},\nu_{\text{obs}}} d\nu_{\text{obs}}$), the HI mass per unit frequency in the galaxy rest frame is given by

$$
M_{\text{HI},\nu_{\text{em}}} = M_{\text{HI},\nu_{\text{obs}}}(1+z) .
$$

The data around rest frequency now contain any HI emission associated with the galaxies. These $n$ aligned spectra are then ‘stacked’ by finding the weighted mean in each channel:

$$
\langle M_{\text{HI}} \rangle_{\nu_{\text{em}}} = \frac{\sum_{i=1}^{n} (w_i M_{\text{HI},\nu_{\text{em},i}})}{\sum_{i=1}^{n} w_i} ,
$$

82
where $M_{H\text{I},\text{em},i}$ is the value of the $i$th mass spectrum at emitted frequency $\nu_{\text{em}}$. We choose the weighting factor $w_i = \sigma_i^{-2}$ (where $\sigma$ is the rms noise level of each observed-frame flux density spectrum), which improves the S/N of the final co-added spectrum. Using $w_i = (\sigma_i D_i^2)^{-2}$ would further optimize the S/N but would reduce the effective volume and increase cosmic variance in our results.

The stacked spectra produced by co-adding all 3,277 SGP sources and 15,093 HIPASS sources are shown in Fig. 3.4. Although only one galaxy is individually detected in the SGP data, a strong 12$\sigma$ averaged detection is achieved through stacking. As expected, a more significant 31$\sigma$ stacked detection is achieved for the HIPASS sample due to the presence of many strong individual detections and a larger overall sample size.

The stacked SGP profile nominally spans the spectral range indicated by the dashed lines in Fig. 3.4 and is $1418.23 - 1421.97$ MHz ($\Delta v = 790$ km s$^{-1}$). The stacked HIPASS profile spans $1418.83 - 1421.77$ MHz ($\Delta v = 621$ km s$^{-1}$). These are significantly wider than the expected H I line widths of the individual galaxies contributing to the stack (see Section 3.5.2.1).

Possible explanations for the broadened stacked profiles include: (i) systematic differences between optical and radio redshifts, (ii) errors in the optical redshifts (85 km s$^{-1}$) and (iii) source confusion. The principle cause is likely confusion, particularly in the case of the higher redshift SGP data where a larger volume is probed per beam. This will be explored in detail in Section 3.5.2.

Note that the HIPASS spectrum is symmetric as expected, yet the SGP spectrum is $\sim 0.62$ MHz (130 km s$^{-1}$) wider at frequencies below the rest frame. This does not appear to be related to optical redshift accuracy.

The green line in Fig. 3.4 indicates the noise level from averaged mock spectra created by stacking sources from a control catalogue. This catalogue was generated by randomly mismatching 2dFGRS redshifts and the source positions. By using the false redshifts in the extraction process, the channels shifted to 1420.4 MHz should
3.4. ANALYSIS

CHAPTER 3. H I IN DISTANT GALAXIES

Figure 3.4: The co-added H I mass spectrum of all 3,277 SGP sources (left) and 15,093 2dFGRS sources covered by HIPASS (right). The vertical dotted line indicates the rest frequency of the H I line. The dashed red lines represent the nominal width of the stacked signal. The stacked spectrum generated using the control catalogue is shown in green to indicate the noise level in each case. A fourth-order polynomial baseline has been subtracted from all co-added spectra. Data between the dashed red lines were excluded when fitting baselines to the black spectra.

not contain any H I emission. Thus, co-adding these mock spectra will not produce any stacked detection but will closely approximate the noise behaviour. A baseline was removed from all stacked spectra via fourth-order least-squares polynomial fitting over a 14 MHz spectral range, excluding the signal region.

3.4.3.3 Average H I mass

Integrating over the channels containing the stacked detection (ν_{em,1} to ν_{em,2}) gives the average H I mass of the sample:

$$\langle M_{H I} \rangle = \int_{\nu_{em,1}}^{\nu_{em,2}} \langle M_{H I} \rangle_{\nu_{em}} d\nu_{em}. \quad (3.4)$$

We find average H I masses of $\langle M_{H I} \rangle = (6.93 \pm 0.17) \times 10^9 h^{-2} M_\odot$ for the SGP sample and $\langle M_{H I} \rangle = (1.48 \pm 0.03) \times 10^9 h^{-2} M_\odot$ for the HIPASS sample.
The error on the average mass is estimated by producing 10 control catalogues as in Section 3.4.3.2, repeating the stacking process for each, and then calculating the rms deviation in the integrated mass over the signal region.

In both cases, we appear to be sampling galaxies with average masses of less than $M_{\text{HI}}^* = 8.4 \times 10^9 h^{-2} M_\odot$ (Zwaan et al., 2005). The 5σ detection limit at the mean redshift of the SGP sample corresponds to an H I mass of $\sim 1 \times 10^{11} h^{-2} M_\odot$.

Note, however, that the average masses calculated here will be larger than the true value for each sample. This is due to significant source confusion in the data and will be discussed in Section 3.5.2. The significantly smaller mass obtained for the HIPASS sample is due to the 2dFGRS magnitude limit which allows more nearby, low-luminosity galaxies to appear, and due to less confusion with other galaxies in the telescope beam.

3.5 Results

3.5.1 Noise behaviour

The best possible S/N achievable with a stacking analysis will be obtained if the noise behaves in a purely Gaussian manner and decreases with the square root of the number of co-added spectra. To demonstrate the noise behaviour of the Parkes data, we stack $N$ randomly selected spectra taken from the control catalogues described in Sections 3.4.3.2 and 3.4.3.3. Note that in this case we are stacking the flux density spectra, as opposed to the mass spectra. We then find the rms noise per channel calculated over the spectral range indicated by the dashed lines in Fig. 3.4.

Fig. 3.5 shows the noise level of the stacked spectrum as a function of the number of co-added spectra. It is evident that the noise within the signal region of the stacked 21 cm spectra in both data sets displays Gaussian behaviour, with gradients $-0.48 \pm 0.05$ (SGP sample) and $-0.50 \pm 0.03$ (HIPASS sample). This demonstrates
that our data reduction has been successful in mitigating against non-Gaussian noise contributions from eg. RFI, residual continuum emission and insufficient bandpass calibration. If we do not apply these careful corrections to the data, we see evidence of a noise ‘floor,’ i.e. the sensitivity does not continue to decrease with the expected gradient, but flattens out once the non-Gaussian components begin to dominate.

We reach a final sensitivity of 68 $\mu$Jy by stacking all 3,277 SGP spectra and 68 $\mu$Jy by stacking all 15,093 HIPASS spectra. The integration time is 975 s/pixel in the SGP data and 220 s/pixel$^{-1}$ in HIPASS (Barnes et al., 2001). This difference explains the vertical displacement of the SGP and HIPASS noise trends in Fig. 3.5. Despite the greater number of sources contributing to the final HIPASS stack, the longer integration time per pixel for the SGP data results in a similar final sensitivity. Observation of a single source for $\sim 37$ d would be required to attain the same sensitivity as these final stacked spectra.

Note, however, that the mass spectra in Fig. 3.4 incorporate a $D_L^2$ factor (see equation 3.1), which is larger for the higher redshift SGP sample. The rms noise level in the stacked SGP mass spectrum is therefore six times larger than for HIPASS.

### 3.5.2 Confusion

The accuracy of the average quantities we measure via this stacking method will ultimately be limited by confusion in the Parkes data. We are unable to distinguish between multiple objects within the Parkes beam and therefore additional H I flux from neighbouring galaxies can contaminate our measurements.

Based on the simulated results of Obreschkow et al. (2013) and the observed H I line widths of HICAT galaxies, it is reasonable to assume a maximum line width of 600 km s$^{-1}$ for any individual galaxy contributing to our analysis.

As discussed in Section 3.4.3.2 above, the stacked H I signals we achieve are broader than this, particularly for the SGP sample. Therefore, we consider all the flux in
the wings of the stacked H I profile, beyond $\pm 300 \text{ km s}^{-1}$ ($\pm 1.42 \text{ MHz}$) from the rest frame, to exist solely due to confusion with other galaxies in the 2dFGRS catalogue, as well as fainter sources below its magnitude limit.

Restricting the integral in equation (3.4) to this spectral range would partially correct for confusion in the measured $\langle M_{\text{H I}} \rangle$. However, any flux within $\pm 300 \text{ km s}^{-1}$ will still include contributions from confused sources, as well as from the target galaxies themselves. The challenge is therefore to identify which sources are confused over this spectral range, and then to correct for this.

### 3.5.2.1 Line width estimation with the Tully-Fisher relation

We do not directly detect the H I signature of many galaxies in our sample and so cannot determine their H I profile widths. Therefore, we cannot ascertain in advance whether we are integrating over single or multiple H I profiles and hence we cannot easily discern the extent of spectral confusion.
3.5. RESULTS

CHAPTER 3. HI IN DISTANT GALAXIES

One way to overcome this is to estimate the HI line widths of each galaxy in our sample using the Tully-Fisher relation. We choose the $B$-band relation of Meyer et al. (2008) which was derived using optical counterparts to HIPASS galaxies:

$$\log(v_{\text{rot}}) = -(M_B + 5 \log h + 1.4)/8.6,$$

where $v_{\text{rot}}$ is the rotational velocity of the galaxy (in km s$^{-1}$) and $M_B$ is the $B$-band absolute magnitude. $B$ apparent magnitudes are estimated from the SuperCOSMOS $b_J$ and $r_F$ apparent magnitudes available in the 2dFGRS catalogue as follows$^3$:

$$B = b_J - 0.0047 + 0.236(b_J - r_F).$$

SuperCOSMOS magnitudes used in equation (3.6) have not been corrected for extinction by the Galaxy. However, we only require indicative values for $B$ and $v_{\text{rot}}$. The HI line width $W_{\text{HI}}$ (defined at 50 per cent peak flux) can then be estimated using:

$$W_{\text{HI}} = 2v_{\text{rot}} \sin(i),$$

where $i$ is the inclination of the galaxy, calculated from the observed axial ratio as in equation (24) in Meyer et al. (2008). The distribution of the resulting line width estimates is shown in Fig. 3.6. The average width is 150 km s$^{-1}$ for SGP and 100 km s$^{-1}$ for HIPASS. This seems reasonable in comparison to the HI line widths of HICAT galaxies, which have an average of 178 km s$^{-1}$.

3.5.2.2 Identification of confused sources

Using the Tully-Fisher predicted line widths, we can now identify which 2dFGRS galaxies are likely to give rise to confusion over the spectral range of interest. We

\hspace{1cm}^3\text{Derived from magnitude conversions available at www2.aao.gov.au/2dfgr/}
consider a source confused with a particular target galaxy if: (i) it is within one beam width of the target galaxy (±21.2 arcmin for the SGP, ±21.9 arcmin for HIPASS) and (ii) if any part of its predicted H I profile is within ±300 km s\(^{-1}\) of the target galaxy’s recessional velocity.

As expected, we find significantly more confusion in the higher redshift SGP field. On average, any one SGP galaxy is confused with seven others. In the lower redshift HIPASS data, this decreases to three.

3.5.2.3 Luminosity correction

The goal of this paper is to derive an estimate of H I density, based on the mass-to-light ratio of the galaxies. While we cannot fully correct \(\langle M_{\text{H I}} \rangle\) for confusion, we can account for this in the density estimate by artificially ‘confusing’ the optical luminosities. We do this by defining
3.5. RESULTS

CHAPTER 3. H I IN DISTANT GALAXIES

\[ L' = f \sum_{i=1}^{n} L_i b_i + L_0, \]  

(3.8)

where \( L_0 \) is the luminosity of the central galaxy, \( L_i \) are the luminosities of the \( n \) confused galaxies, \( b_i \) is the beam weighting factor defined using a Gaussian with full width at half maximum (FWHM)=21.2 arcmin (SGP) and FWHM=21.9 arcmin (HIPASS), and \( f \) is an extrapolation factor to account for luminosity from galaxies fainter than the 2dFGRS survey limit. The latter assumes the luminosity density function of Norberg et al. (2002), accounting for evolution and \( k \)-corrections.

Fig. 3.7 shows the original \( b_J \) luminosity distribution of each sample, compared to the confusion-adjusted luminosity distribution \((L')\). The average (catalogued) luminosity for the SGP sample is \( 6.62 \times 10^9 \, h^{-2} L_\odot \). Once adjustment for confusion has been carried out, this average increases by a factor of 5 to \( 3.36 \times 10^{10} \, h^{-2} L_\odot \). The average original luminosity of the HIPASS sample is \( 1.93 \times 10^9 \, h^{-2} L_\odot \), which increases by a factor of 2.5 to \( 4.87 \times 10^9 \, h^{-2} L_\odot \) when adjusted. As expected, the correction to the HIPASS luminosities is less significant to that of the SGP, due to less source confusion and a fainter absolute luminosity limit.

3.5.3 H I mass density

The H I density, \( \rho_{H I} \), can be calculated from the luminosity density and the mass-to-light ratio using

\[ \rho_{H I} = \left( \frac{M_{H I}}{L} \right)_{\rho_L} \times \rho_L, \]  

(3.9)

as in eg. Fall & Pei (1993). We use the \( b_J \) luminosity density \( \rho_L = (1.82 \pm 0.17) \times 10^8 \, h \, L_\odot \, Mpc^{-3} \) derived by Norberg et al. (2002) using the full 2dFGRS catalogue,
CHAPTER 3. H I IN DISTANT GALAXIES

3.5. RESULTS

Figure 3.7: The luminosity distribution of all stacked optical sources in the SGP field (left) and HIPASS field (right). The black line represents the original $b_J$ luminosities of the sample, as defined in the 2dFGRS source catalogue. The red line shows the distribution once luminosities have been adjusted to include a beam-weighted sum of the luminosities of all confused galaxies.

with appropriate evolution and $k$-corrections applied (up to 17 per cent at the redshifts of the SGP galaxies).

We calculate $\langle M_{\text{H}1}/L \rangle$ by stacking individual $M_{\text{H}1}/L$ ‘spectra’. We do so via equations (3.3) and (3.4), replacing $M_{\text{H}1,\nu \text{em},i}$ with $M_{\text{H}1,\nu \text{em},i}/L'_i$. As discussed in Section 3.5.2, it is appropriate to restrict the spectral integration range to $\pm 300 \text{km s}^{-1}$ from the rest frame since any contribution outside of this is due only to confused sources. We find $\langle M_{\text{H}1}/L \rangle = 0.43 \pm 0.03 M_\odot/L_\odot$ for the SGP sample, and $0.72 \pm 0.03 M_\odot/L_\odot$ for the HIPASS sample, where the errors are measured as in Section 3.4.3.3.

We require the mass-to-light ratio to be weighted by the luminosity density, rather than the luminosity selection function of the input catalogue (hence the subscript $\rho_L$ in equation 3.9). We therefore apply a weighting correction to the mean mass-to-light ratio as follows:

$$\langle \frac{M_{\text{H}1}}{L'} \rangle_{\rho_L} = \langle \frac{M_{\text{H}1}}{L'} \rangle_s \times \frac{\int N(L) dL}{\int L^\alpha N(L) dL} \times \frac{\int L^\alpha \rho_L(L) dL}{\int \rho_L(L) dL},$$

(3.10)
3.5. RESULTS

CHAPTER 3. H I IN DISTANT GALAXIES

where $\rho_L(L)$ is the luminosity density function and $N(L)$ is the original luminosity distribution of the input redshift sample. The parameter $\alpha$ is the power-law dependence of the average H I mass-to-light ratio of galaxies on luminosity in solar units:

$$\frac{M_{\text{H}I}}{L} = \beta L^\alpha. \quad (3.11)$$

We adopt the $z = 0$ values $\alpha = -0.38$ and $\log \beta = (3.34 - 2\alpha \log h)$ found by Karachentsev et al. (2008) using a subset of HIPASS galaxies with optical counterparts. For the HIPASS sample, the weighting factor applied to the mean mass-to-light ratio is 0.62. For the SGP sample, where galaxies are more luminous and have lower H I mass-to-light ratios, it is 1.35. That is, the mean H I mass-to-light ratio is predicted to be 2.2 times greater for galaxies in the HIPASS sample compared with the SGP sample on the basis of our previous knowledge of the luminosity dependence of $M_{\text{H}I}/L$.

We convert $\rho_{\text{H}I}$ into a fraction of the critical density of the Universe per unit comoving volume, $\Omega_{\text{H}I}$, using

$$\Omega_{\text{H}I} = \frac{8\pi G \rho_{\text{H}I}}{3H_0^2}, \quad (3.12)$$

where $G$ is the gravitational constant and $H_0$ is the Hubble constant. We find $\Omega_{\text{H}I} = (3.19^{+0.43}_{-0.59}) \times 10^{-4} h^{-1}$ for the $0.04 < z < 0.13$ SGP data and $(2.82^{+0.30}_{-0.59}) \times 10^{-4} h^{-1}$ for the $z < 0.04$ HIPASS data. This implies that there has been no $(12 \pm 23$ per cent) evolution in the cosmic H I mass density over the past $1.2 h^{-1} \text{Gyr}$.

The errors in $\Omega_{\text{H}I}$ are a combination of: (i) the $\sim 6$ per cent error in $\langle M_{\text{H}I}/L \rangle$ quoted previously, (ii) a 9 per cent error in $\rho_L$ and (iii) the $1-11$ per cent estimated uncertainty in the confusion correction. The latter was estimated from: (a) variance
in \( \langle M_{\text{H}i}/L \rangle \) in different windows from \( \pm 200 \) to \( \pm 400 \text{ km s}^{-1} \) over the stacked spectrum and (b) variance in \( \langle M_{\text{H}i}/L \rangle \) calculated using different angular resolutions (and corresponding confusion corrections) in the range \( 15.5 - 21.9 \text{ arcmin} \).

Table 3.2 and Fig. 3.8 compare the values calculated here to other observational constraints on \( \Omega_{\text{H}i} \) from the literature. Values at look-back times of less than \( 4 \text{ h}^{-1} \text{ Gyr} \) have been calculated from either direct or stacked 21 cm detections. The Chang et al. (2010) point at \( 4.8 \text{ h}^{-1} \text{ Gyr} \) was derived using 21 cm intensity mapping. All other points estimate the \( \text{H}i \) column density of galaxies from the absorption spectra of background QSOs. The evolutionary trends in Fig. 3.8 are flatter than other examples of this plot in the literature because we have removed the contribution of helium from the values presented in Rao et al. (2006) and Lah et al. (2007) and used the latest DLA measurements of Prochaska & Wolfe (2009) rather than Prochaska et al. (2005).

The \( \Omega_{\text{H}i} \) values we derive here are consistent, within the error margins, with previous estimates over the same redshift range by Zwaan et al. (2005), Martin et al. (2010) and Freudling et al. (2011). It is particularly interesting to note the close agreement between our low-redshift result, derived by stacking HIPASS detections and non-detections, and that of Zwaan et al. (2005), derived from only direct HIPASS detections.

It is possible that we have underestimated our uncertainties in the confusion correction. Perhaps, due to evolution, there are more low-mass galaxies in our beam than we have accounted for. Whether or not this is the case will be apparent in future deep studies with radio telescopes with smaller beam sizes. However, it could also be that the other measurements listed have missed galaxies, perhaps because of the difficulty in defining selection functions or extrapolating the \( \text{H}i \) density function, and are not as accurate as claimed. The use of the stacking method means that the contribution from even low \( \text{H}i \) mass galaxies is included. Furthermore, as will be apparent in the next section, previous surveys may in fact be completely dominated by cosmic variance.
### Table 3.2: Observational constraints on $\Omega_{H_1}$ from the literature. All values have been converted to the same cosmology. The mean redshift and associated look-back time of each sample are quoted. Results presented in this paper are shown in bold. Errors on our values include statistical uncertainty in our stacking analysis, uncertainty in the 2dFGRS luminosity density and systematic errors introduced by correction for confusion.\textsuperscript{a} Original values quoted were $\Omega_{\text{gas}}$ and included a He contribution but have been converted to $\Omega_{H_1}$ here. (See source for the value of the assumed contribution).\textsuperscript{b} Values were calculated using only systems with H I surface densities $> 2 \times 10^{20}$ cm$^{-2}$ (i.e. measures $\Omega_{H_1}^{\text{DLA}}$).
3.5.4 Cosmic variance

When observations are limited to a finite volume of the Universe, a significant contribution to the error in estimates of cosmological values, such as $\Omega_{\text{H}1}$, can come from cosmic variance. That is, variation in galaxy density due to large-scale structure. By employing the stacking technique, we are not restricted to volumes where direct detections are possible. Therefore, the volumes we probe are $3.7 \times 10^5$ and $3.4 \times 10^5 \, h^{-3} \, \text{Mpc}^3$ for the SGP and HIPASS samples, respectively. Within the median 2dFGRS redshift range ($z < 0.11$), the SGP volume reduces to $1.8 \times 10^5 \, h^{-3} \, \text{Mpc}^3$. These are significantly larger than the sampled volumes of Zwaan et al. (2005) using HICAT ($8.2 \times 10^4 \, h^{-3} \, \text{Mpc}^3$, for a median redshift of 0.009), Martin et al. (2010) using the 40 per cent ALFALFA source catalogue ($1.4 \times 10^5 \, h^{-3} \, \text{Mpc}^3$, for a median redshift of 0.027) and Freudling et al. (2011) using the AUDS pilot survey ($7.5 \times 10^2 \, h^{-3} \, \text{Mpc}^3$). Furthermore, the luminosity density $\rho_L$ is calculated by Norberg et al. (2002) from an even higher volume of $7.3 \times 10^6 \, h^{-3} \, \text{Mpc}^3$. Therefore, uncertainty in $\Omega_{\text{H}1}$ due to
3.5. RESULTS

CHAPTER 3. H I IN DISTANT GALAXIES

cosmic variance is dramatically lower in our results than in any previous sample.

As discussed in Section 3.4.1, we have chosen a region which is not representative of the average Universe at $0.04 < z < 0.13$. Rather, the SGP field contains a small overdensity. The selection of an over-dense region gives a higher volume-integrated H I mass. However, to first order, this should not bias our measurement of the average mass-to-light ratio of galaxies, and therefore the H I mass density we calculate will be a fair estimate of the cosmic value. While the local environment is known to influence the H I properties of galaxies (eg. Haynes et al. 1984), little is known about the relationship between the H I mass-to-light ratio and galaxy density over the very large scales and very low overdensities we probe. For comparison, galaxies in the Virgo Cluster, which has a stellar overdensity two orders of magnitude higher than the field (Davies et al., 2012), appear to be four times more H I-deficient compared with background field galaxies (Taylor et al., 2012). Crudely assuming that the interaction rate between galaxies (as for gaseous particles) is proportional to the product of density and velocity dispersion, an overdensity of 30 per cent will only lead to an H I deficiency of 1 per cent or so.

3.5.5 Binning

We now investigate H I trends with redshift, luminosity and colour. We split the parent sample into bins containing roughly equal source counts. We then stack the mass spectra of all sources within each subsample. By splitting the SGP parent sample into three bins and the HIPASS sample into eight bins, we can still achieve a statistically significant detection in each. Tables 3.3 and 3.4 show the properties of each SGP and HIPASS subsample, respectively. The variation of the average H I mass and mass-to-light ratio in each bin is plotted in Fig. 3.9.

To minimize confusion in our results, we have restricted the measurement of both $\langle M_{\text{HI}} \rangle$ and $\langle M_{\text{HI}}/L \rangle$ to the ±300 km s$^{-1}$ spectral range, as in Section 3.5.3. We have adjusted the luminosities to account for confusion (via equation 3.8) before binning.
Figure 3.9: The average H I masses (left) and average mass-to-light ratios (right) derived by co-adding all spectra in bins of (from top to bottom) redshift, $b_J$ luminosity and $(b_J-r_F)$ colour. The values are plotted at the mean value of the binned parameter in each sub-sample. Error bars represent the 1σ spread within the bin. Values derived from SGP data are shown as black points, while the HIPASS values are shown as red diamonds. The red dashed lines in the two lower right panels show the least-squares fit to the HIPASS points. The black dot-dashed lines show the fit to the SGP points, assuming the same gradient as the HIPASS fit. The green dotted line shows the relationship between $M_{\text{H I}}/L$ and luminosity as estimated by Karachentsev et al. (2008).
Table 3.3: Binned parameters for the SGP data. Columns (from left to right) are: the binned property, bin range, mean and rms within the bin range, number of sources, integrated and peak signal-to-noise ratio of the stacked mass spectrum, average mass (units $M_{\odot}$) and average mass-to-light ratio (units $M_{\odot}/L_{\odot}$) calculated from the stacked spectrum.

<table>
<thead>
<tr>
<th>Property</th>
<th>Bin range</th>
<th>Mean $\langle M_{H\text{i}} \rangle$ ($\times 10^9 h^{-2}$)</th>
<th>$\langle M_{H\text{i}}/L_{\odot} \rangle$</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T/\beta$</td>
<td>$z=0.083$</td>
<td>$9.2 \times 10^9 h^{-2}$</td>
<td>$1.43 \times 10^9$</td>
<td>1.271 - 2.089</td>
</tr>
<tr>
<td>$I/\beta$</td>
<td>$z=0.083$</td>
<td>$7.0 \times 10^9 h^{-2}$</td>
<td>$1.01 \times 10^9$</td>
<td>1.091 - 1.271</td>
</tr>
<tr>
<td>$(b-r)$</td>
<td>$z=0.083$</td>
<td>$3.2 \times 10^9 h^{-2}$</td>
<td>$1.27 \times 10^9$</td>
<td>1.271 - 2.089</td>
</tr>
<tr>
<td>$T/\beta$</td>
<td>$z=0.167$</td>
<td>$7.0 \times 10^9 h^{-2}$</td>
<td>$2.6 \times 10^9$</td>
<td>0.081 - 0.167</td>
</tr>
<tr>
<td>$I/\beta$</td>
<td>$z=0.167$</td>
<td>$7.0 \times 10^9 h^{-2}$</td>
<td>$0.36 \times 10^9$</td>
<td>0.081 - 0.167</td>
</tr>
<tr>
<td>$(b-r)$</td>
<td>$z=0.167$</td>
<td>$3.2 \times 10^9 h^{-2}$</td>
<td>$0.63 \times 10^9$</td>
<td>0.081 - 0.167</td>
</tr>
<tr>
<td>Property</td>
<td>Bin range</td>
<td>Mean</td>
<td>rms</td>
<td>N</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>(z)</td>
<td>0.0026 – 0.0171</td>
<td>0.0111</td>
<td>0.0043</td>
<td>1814</td>
</tr>
<tr>
<td></td>
<td>0.0171 – 0.0224</td>
<td>0.0200</td>
<td>0.0014</td>
<td>1981</td>
</tr>
<tr>
<td></td>
<td>0.0224 – 0.0265</td>
<td>0.0245</td>
<td>0.0012</td>
<td>1814</td>
</tr>
<tr>
<td></td>
<td>0.0265 – 0.0298</td>
<td>0.0284</td>
<td>0.0009</td>
<td>1964</td>
</tr>
<tr>
<td></td>
<td>0.0298 – 0.0327</td>
<td>0.0313</td>
<td>0.0008</td>
<td>1799</td>
</tr>
<tr>
<td></td>
<td>0.0327 – 0.0361</td>
<td>0.0345</td>
<td>0.0010</td>
<td>1986</td>
</tr>
<tr>
<td></td>
<td>0.0361 – 0.0388</td>
<td>0.0374</td>
<td>0.0008</td>
<td>1798</td>
</tr>
<tr>
<td></td>
<td>0.0388 – 0.0423</td>
<td>0.0405</td>
<td>0.0010</td>
<td>1937</td>
</tr>
<tr>
<td>(\log (L'/L_\odot h^{-2}))</td>
<td>6.206 – 8.426</td>
<td>7.997</td>
<td>0.413</td>
<td>1812</td>
</tr>
<tr>
<td></td>
<td>8.426 – 8.787</td>
<td>8.623</td>
<td>0.101</td>
<td>1963</td>
</tr>
<tr>
<td></td>
<td>8.787 – 9.073</td>
<td>8.932</td>
<td>0.082</td>
<td>1810</td>
</tr>
<tr>
<td></td>
<td>9.073 – 9.345</td>
<td>9.214</td>
<td>0.079</td>
<td>1962</td>
</tr>
<tr>
<td></td>
<td>9.345 – 9.574</td>
<td>9.459</td>
<td>0.066</td>
<td>1811</td>
</tr>
<tr>
<td></td>
<td>9.574 – 9.809</td>
<td>9.683</td>
<td>0.069</td>
<td>1962</td>
</tr>
<tr>
<td></td>
<td>9.809 – 10.050</td>
<td>9.927</td>
<td>0.070</td>
<td>1811</td>
</tr>
<tr>
<td></td>
<td>10.050 – 10.734</td>
<td>10.248</td>
<td>0.146</td>
<td>1962</td>
</tr>
<tr>
<td>((b_J - r_F))</td>
<td>-0.994 – 0.504</td>
<td>0.309</td>
<td>0.224</td>
<td>1814</td>
</tr>
<tr>
<td></td>
<td>0.504 – 0.629</td>
<td>0.572</td>
<td>0.035</td>
<td>1965</td>
</tr>
<tr>
<td></td>
<td>0.629 – 0.716</td>
<td>0.673</td>
<td>0.025</td>
<td>1816</td>
</tr>
<tr>
<td></td>
<td>0.716 – 0.807</td>
<td>0.762</td>
<td>0.027</td>
<td>1956</td>
</tr>
<tr>
<td></td>
<td>0.807 – 0.908</td>
<td>0.856</td>
<td>0.030</td>
<td>1814</td>
</tr>
<tr>
<td></td>
<td>0.908 – 1.056</td>
<td>0.977</td>
<td>0.043</td>
<td>1955</td>
</tr>
<tr>
<td></td>
<td>1.056 – 1.190</td>
<td>1.125</td>
<td>0.039</td>
<td>1817</td>
</tr>
<tr>
<td></td>
<td>1.190 – 1.958</td>
<td>1.287</td>
<td>0.090</td>
<td>1956</td>
</tr>
</tbody>
</table>

Table 3.4: Binned parameters for the HIPASS data. Columns are as defined in Table 3.3.
3.5. RESULTS

H 1 IN DISTANT GALAXIES

<table>
<thead>
<tr>
<th></th>
<th>HIPASS</th>
<th>SGP (Karachentsev et al., 2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(L′/L⊙k−2)</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>−0.46</td>
<td>4.00</td>
<td>4.35</td>
</tr>
<tr>
<td>(bJ − r_F)</td>
<td>−1.15</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Table 3.5: Parameters for the lines in Fig. 3.9 showing the derived relationships between H 1 mass-to-light ratio and luminosity and colour: log((M_H 1/L′)) = ax + b. The fit to the SGP points has been derived using the HIPASS gradient; hence, only the offset is quoted.

We have also used these corrected luminosities (L′) to determine the mass-to-light ratios. Note, however, that it is not appropriate to apply the M_H 1/L weighting (equation 3.10) to this binning analysis.

The average H 1 mass-to-light ratios we measure are consistent with the results of Doyle et al. (2005) who study HIPASS galaxies with optical counterparts identified in the 6dF Galaxy Survey. They find that the bJ-band mass-to-light ratios of these galaxies are predominantly less than 5 M⊙/L⊙. For comparison, the highest mass-to-light ratio we measure is 2.0 ± 0.2 M⊙/L⊙ and is associated with the lowest HIPASS luminosity bin.

We find that the H 1 mass increases and the mass-to-light ratio decreases with bJ luminosity in both the HIPASS and SGP samples. The former reflects the fact that more luminous galaxies are often larger in size and therefore contain greater total H 1 masses (eg. Toribio et al. 2011). The fact that lower optical luminosity galaxies have higher mass-to-light ratios has been well established through direct H 1 observations (eg. Warren et al. 2006) and we have verified here that such trends can be reproduced with a stacking analysis.

We have performed a least-squares linear fit to the HIPASS data, as shown in Fig. 3.9. Fitting to the SGP data is less robust, owing to the small number of bins. However, we perform the linear fit assuming the same gradient as the HIPASS fit. The dotted line in Fig. 3.9 shows the close match of the mass-to-light versus luminosity derived by Karachentsev et al. (2008), compared with our results (see also Table 3.5). This verifies our adjustment for the luminosity dependence of the mass-to-light ratio. The
only SGP point to substantially deviate is the lowest luminosity bin, which is the main reason for the slightly increased value of $\Omega_{\text{H}I}$ in Table 3.2.

Similar $\text{H}I$ trends are seen with redshift. This does not reveal any real evolution, but arises because we do not have a volume-limited sample and are therefore preferentially detecting the higher mass and luminosity galaxies at higher redshifts. It is precisely this bias that we are attempting to correct via equation (3.10) in our above calculation of the $\text{H}I$ mass density.

We use the SuperCOSMOS $b_J$ and $r_F$ magnitudes, as defined in the 2dFGRS catalogue, to examine trends with colour. For the SGP population, which contains high-mass galaxies in each colour bin, a higher total $\text{H}I$ mass is seen for bluer sources. The HIPASS sample contains a range of masses at all colours, and it is therefore unfair to directly compare the $\text{H}I$ masses in each bin. The large vertical offset between the two samples can again be explained by the preferential inclusion of higher mass sources in the SGP sample due to the magnitude limits of the optical catalogue.

Taking the ratio with the optical luminosity removes the galaxy size dependence, revealing a clear relationship between the $\text{H}I$ mass-to-light ratio and colour. We see that bluer sources have higher mass-to-light ratios. Again, we perform a linear fit to the HIPASS data, and apply this gradient to the SGP fit. Both data sets are remarkably consistent and reflect similar trends found between colour and gas fraction in previous studies using both direct $\text{H}I$ detections (e.g. Spitzak & Schneider 1998) and 21 cm stacking experiments (Fabello et al. 2011a). This supports the notion that galaxies with larger gas supplies are generally associated with increased star formation.

Much of the colour dependence reflects the colour-magnitude relation and is therefore corrected for in our calculation of $\Omega_{\text{H}I}$, which incorporates a luminosity correction to $M_{\text{H}I}/L$. However, it is possible that there remain residual differences between galaxies in the SGP and HIPASS samples. For example, Li et al. (2012) examined the residual dependence of $\text{H}I$-to-stellar mass ratio on a number of parameters, and found that a
simultaneous dependence on stellar mass and colour (as well surface brightness and colour gradient) best explains the properties of luminous SDSS galaxies. However, after application of the luminosity correlation in Table 3.5, we find no residual dependence of $M_{\text{H}i}/L$ on colour to an accuracy of 10 per cent over the mean colour difference between SGP and HIPASS galaxies.

3.6 Conclusions

We have presented an H i spectral stacking analysis of galaxies identified in the 2dFGRS. We have shown that H i stacking can be used to efficiently and accurately probe the statistical properties of high-redshift field galaxies.

Our sample consists of 15,093 galaxies at $0.0025 < z < 0.0423$ and 3,277 galaxies at $0.0405 < z < 0.1319$. 21 cm data for the low-redshift sample were provided by HIPASS. For the high-redshift sample, new 21 cm observations of a 42 deg$^2$ field near the SGP were conducted with the Parkes radio telescope. Many of the low-redshift optical galaxies are found to have strong H i signatures, consistent with the HICAT. Only one galaxy is detected above the 5 $\sigma$ limit in the SGP sample.

We have co-added the H i spectra of all galaxies in our sample, after aligning each at the rest frame. We thus report a strong 31 $\sigma$ average detection of the HIPASS galaxies and a 12 $\sigma$ detection for the SGP galaxies. The sensitivity level we achieve for our stacked spectra is equivalent to observing a single object for $\sim 37$ d. Therefore, we have shown that spectral stacking is an effective method of making statistical detections of large numbers of individually undetected galaxies at high redshift with reasonable integration times.

The rms noise in the stacked signal displays Gaussian behaviour, decreasing with the square root of the number of co-added spectra. We find no apparent intrinsic limitation in the depth to which we can continue stacking experiments with the Parkes telescope. This bodes well for future stacking experiments with deep, large-scale
H\textsc{i} surveys on the next generation of radio telescopes such as the Australian SKA Pathfinder, MeerKAT, APERTIF and ultimately the SKA.

We measure an average H\textsc{i} mass of $\langle M_{\text{H\textsc{i}}} \rangle = (1.48 \pm 0.03) \times 10^9 h^{-2} M_\odot$ from the HIPASS stacked spectrum and $\langle M_{\text{H\textsc{i}}} \rangle = (6.93 \pm 0.17) \times 10^9 h^{-2} M_\odot$ from the SGP stacked spectrum. However, we find that these averages, particularly that of the SGP sample, are over estimated due to source confusion. High-resolution follow-up observations to combat this issue are underway with the Australia Telescope Compact Array and will be presented in a future paper.

For now, we employ the Tully-Fisher relation to estimate source line widths and thus to predict the extent of confusion in our data. We then adjust the optical luminosities to include a beam-weighted sum of all confused galaxies. Finding the average mass-to-light ratio and accounting for sample bias, we derive the cosmic H\textsc{i} mass density ($\Omega_{\text{H\textsc{i}}}$). We find $\Omega_{\text{H\textsc{i}}} = (2.82^{+0.30}_{-0.59}) \times 10^{-4} h^{-1}$ at $0 < z < 0.04$ and $(3.19^{+0.43}_{-0.59}) \times 10^{-4} h^{-1}$ at $0.04 < z < 0.13$, i.e. we find no $(12\pm23$ per cent) evolution over the past $\sim 1 h^{-1}$ Gyr. Within the error margins, these values agree with previous $\Omega_{\text{H\textsc{i}}}$ estimates made via direct detections in the HIPASS, ALFALFA and AUDS surveys. We argue that our results are far more robust to cosmic variance.

Finally, we employ the stacking technique to investigate the variation of H\textsc{i} properties with redshift, luminosity and colour. Trends seen with redshift are obscured by selection effects. We find that lower luminosity galaxies have lower total H\textsc{i} masses but higher mass-to-light ratios. We also see a decrease in the mass-to-light ratio from blue to red galaxies. These results agree with direct observations.

**Acknowledgements**

We are grateful to staff at the Parkes radio telescope for technical support, particularly Ettore Carretti and Stacy Mader. Thanks to Morag Scrimgeour and Chris Harris for observing assistance. The Parkes radio telescope is part of the Australia
Telescope National Facility which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. We acknowledge the efforts of the 2dFGRS and HIPASS collaborations. JD wishes to thank the International Centre for Radio Astronomy Research for funding support, Laura Hoppmann and Stefan Westerlund for coding assistance and Tobias Westmeier, Aaron Robotham, Chris Power, Martin Zwaan, Marc Verheijen and Jason Prochaska for useful discussions. Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. We acknowledge the use of the TOPCAT software in our data analysis (http://www.starlink.ac.uk/topcat/).
Chapter 4

GAMA: Studying H I properties of galaxies to \( z = 0.1 \) with spectral stacking

Foreword and context

This chapter again explores the uses of spectral stacking for probing the H I properties of galaxies at higher redshifts than normally possible. This work uses a similar analysis strategy to Chapter 3, using new radio observations collected with the Parkes radio telescope and also the HI Parkes All Sky Survey. However, it makes use of spectroscopic redshifts provided by the new Galaxy and Mass Assembly (GAMA) survey. The benefit of using GAMA redshifts instead of 2dFGRS redshifts will be discussed. The sky fields studied in this chapter are shown in red in Figure 1.3 in Chapter 1.

A major goal of this chapter is to develop the stacking technique further. Employing the availability of multiwavelength data and stellar masses for all GAMA sources, we explore different strategies for deriving \( \Omega_{\text{H}1} \) from a stacking analysis. We focus particularly on the impact of completeness and sample bias on the final result. The
4.1 Abstract

We present a study of the neutral atomic hydrogen (\(\text{H} \text{I}\)) properties of galaxies in the Galaxy and Mass Assembly (GAMA) survey at \(z < 0.11\) using spectral stacking techniques. New observations of the 21 cm emission of H I at \(0.04 < z < 0.1\) in two \(48\,\text{deg}^2\) equatorial GAMA fields at right ascension \(9^h\) and \(14^h30^m\) were conducted using the Parkes 64 m radio telescope. Low redshift \((z < 0.04)\) 21 cm data from the H I Parkes All Sky Survey (HIPASS) is also available over these fields. Within the \(r\)-band luminosity completeness limits of GAMA, 990 GAMA galaxies at \(z < 0.04\)
and 4,898 galaxies at $0.04 < z < 0.11$ are available for stacking. We also define subsets of these samples which are complete in stellar mass and minimise colour selection bias. These stellar mass selected catalogues contain 654 and 3,361 in the lower and higher redshift ranges, respectively. By co-adding the 21 cm spectra of these galaxies and correcting for beam confusion, we derive average $^{1}\text{H}i$ mass-to-light ratios of $\langle M_{\text{H}i}/L_r \rangle = 0.77 \pm 0.21 \, M_\odot / L_\odot$ and $0.16 \pm 0.02 \, M_\odot / L_\odot$ for the lower and higher redshift samples, respectively. We find average $^{1}\text{H}i$-to-stellar mass ratios of $\langle M_{\text{H}i}/M \rangle = 2.19 \pm 0.19$ at $z < 0.04$ and $0.15 \pm 0.01$ at $z > 0.04$. The cosmic $^{1}\text{H}i$ mass density ($\Omega_{\text{H}i}$) is calculated using both the $r$-band luminosity and stellar mass information and using two separate weighting schemes to account for sample incompleteness. We find that the results of the two weighting schemes are consistent but that a systematic difference between luminosity and stellar mass-derived $\Omega_{\text{H}i}$ estimates exists. At $0.04 < z < 0.11$ we find $\Omega_{\text{H}i} = (3.22 \pm 0.53) \times 10^{-4} \, h^{-1}$ using stellar masses and $(0.92 \pm 0.09) \times 10^{-4} \, h^{-1}$ using luminosities. This inconsistency cannot be attributed to cosmic variance but may be partially explained by evolution in the optical luminosities.

4.2 Introduction

To achieve a comprehensive understanding of galaxy formation and evolution, it is necessary to study the properties of the galaxy population as a whole. This requires wide-field surveys which detect large numbers of galaxies in a variety of environments. Surveys must also be highly sensitive in order to probe galaxies at cosmologically large distances and at different evolutionary stages. Such large surveys have traditionally been conducted in the optical (e.g. the Sloan Digital Sky Survey; Abazajian et al., 2009, and the Two Degree Field Galaxy Redshift Survey; Colless et al., 2001) and more recently in the near-infrared (e.g. VISTA Kilo-Degree Infrared Galaxy survey; Fleuren et al., 2012) and far-infrared (e.g. the Herschel Astrophysical Terahertz Large Area Survey; Eales et al., 2010). Over the next decade, it will become possible to
probe large space volumes in the radio regime using the next generation of radio telescopes such as the Square Kilometre Array (SKA) and the pathfinder instruments (e.g. ASKAP, MeerKAT and APERTIF). This will allow detailed studies of cold ($< 10^4$ K) gas - a vital galactic component and the primary reservoir for future star formation.

Tracing this gas, whose most significant component is usually neutral atomic hydrogen (H\textsc{i}), over $0 < z < 2$ should help to explain the rapid decline in the cosmic star formation rate density over this range (Hopkins & Beacom, 2006; Behroozi et al., 2013). However, H\textsc{i} is notoriously difficult to trace at intermediate redshifts. At high redshift, H\textsc{i} can be probed through damped Lyman $\alpha$ absorption along sight lines to quasi-stellar objects (e.g. Prochaska & Wolfe, 2009). This becomes difficult below $z \approx 1.65$, where the Lyman $\alpha$ line shifts into the ultraviolet and can only be probed using space-based instruments (e.g. Rao et al., 2006).

In the local Universe, H\textsc{i} can be directly detected via its 21 cm emission line. Below $z \approx 0.06$, the H\textsc{i} Universe has been well mapped by all-sky or wide-field surveys such as the H\textsc{i} Parkes All Sky Survey (HIPASS; Barnes et al., 2001) and the Arecibo Legacy Fast ALFA (ALFALFA; Giovanelli et al., 2005) and the local H\textsc{i} mass function has been well constrained (Zwaan et al., 2005; Martin et al., 2010). However, the current generation of telescopes are limited in sensitivity, bandwidth and survey speed. This has restricted the most distant direct detections of 21 cm emission to $z \approx 0.2$ and has only been achievable using pointed observations of small samples of optically-selected galaxies (Catinella et al., 2008) or clusters (Zwaan et al., 2001; Verheijen et al., 2007).

Wide-field, blind 21 cm surveys at higher redshift are now becoming possible with the installation of more advanced correlators and backends. For example, observations for the Arecibo Ultra-Deep Survey have recently been completed and are expected to reveal $>110$ H\textsc{i} detections out to $z = 0.16$ (Freudling et al., 2011; Hoppmann et al., in prep). In addition, the newly upgraded Karl G. Jansky Very Large Array (JVLA; Perley et al., 2011) has recently detected H\textsc{i} in 33 galaxies in a blind survey out to $z = 0.193$ (Fernández et al., 2013). These observations are now being expanded into a
larger survey which is expected to detect over 300 galaxies out to \( z = 0.4 \) (Fernández et al., 2013). The SKA and its pathfinders will have the potential to directly detect \( 21\text{cm} \) emission out to \( z \approx 1 \). However, they are still in planning or construction phases.

In the meantime, ‘stacking’ of individually-undetected \( \text{H}\text{I} \) spectra is becoming a popular technique for achieving statistical detections and determining average properties of faint and/or distant systems (e.g. Chengalur et al., 2001; Lah et al., 2007; Fabello et al., 2011a; Delhaize et al., 2013; Rhee et al., 2013; Geréb et al., 2013). Stacking requires an external spectroscopic redshift catalogue to identify the positions and redshifts of sources below the noise level of the \( \text{H}\text{I} \) data. In Delhaize et al. (2013; hereafter referred to as D13), we stacked \( \text{H}\text{I} \) spectra, collected with the Parkes 64 m radio telescope, of 15,093 galaxies in the Two-Degree Field Galaxy Redshift Survey (2dFGRS) at \( z < 0.04 \) and 3,277 galaxies at \( 0.04 < z < 0.13 \). We achieved strong statistical detections in each case and demonstrated that \( \text{H}\text{I} \) spectral stacking can be used to examine the cosmic \( \text{H}\text{I} \) mass density \( (\Omega_{\text{H}\text{I}}) \) over the largest volume to date. This minimised the impact of cosmic variance in our results, and allowed a precise measurement out to \( z = 0.13 \).

We now conduct a similar stacking experiment, again with \( \text{H}\text{I} \) data from Parkes. This time, we exploit spectroscopic redshifts from the Galaxy and Mass Assembly (GAMA; Driver et al., 2011) redshift survey. While both the 2dFGRS and GAMA were conducted with the Anglo-Australian Telescope, GAMA probes galaxies \( \sim 1.1 \) magnitudes deeper than the 2dFGRS with seven times the redshift density. Therefore, GAMA is sensitive to fainter galaxies at a given redshift and provides more spectra per square degree available for stacking. Additionally, the completeness of the GAMA survey is well-characterised and sample biases are well-understood. These are both important considerations for a stacking analysis. We also take advantage of the multiwavelength data available over the GAMA fields to investigate various methods of calculating \( \Omega_{\text{H}\text{I}} \).

Section 4.3 of this paper describes the optical and multiwavelength information provided
4.3. **The data**

4.3.1 **GAMA**

The optical source positions and redshifts required for this stacking analysis are provided by GAMA - a large, multiwavelength galaxy evolution survey (Driver et al., 2009). Central to this campaign is the spectroscopic redshift survey (GAMAz) of over 220,000 galaxies (Driver et al., 2011). GAMAz is operating with the AAOmega spectrograph and 2dF instrument on the Anglo-Australian Telescope. Observations began in 2008 and are ongoing. The first phase of GAMAz (referred to as GAMA-I), targeted three equatorial $12 \times 4 \text{deg}^2$ fields centred at right ascension (RA) $9^h$ (G09), $12^h$ (G12) and $14^h30^m$ (G15). The second phase (GAMA-II) commenced in 2010 and expands each equatorial region by $12 \times 1 \text{deg}^2$, as well as targeting two additional southern fields. Target selection is described in Baldry et al. (2010) and consists of galaxies from the Sloan Digital Sky Survey with $r_{\text{petro}} < 19.4$ (GAMA-I) and $r_{\text{petro}} < 19.8$ (GAMA-II). There is a high level of completeness ($\sim 97$ per cent) to $r_{\text{petro}} < 19.4$ in all GAMA-I fields (Robotham et al., 2010; Driver et al., 2011). The pipeline for spectroscopic reduction and analysis is presented in Hopkins et al. (2013). Determination of redshifts with the automated runz code and via manual identification is described in Driver et al. (2011).

The work presented here uses data within the G09 and G15 fields. G12 is not suitable for such studies as it is close to the bright radio continuum source 3C 273 which...
would adversely affect the H\textsc{i} observations. Only GAMA-I data is used for this work, although available GAMA-II redshifts provide a quality check. We consider only galaxies within the main catalogue (survey\_class $\geq 6$), within the completeness limit ($r_{\text{petro}} < 19.4$) and with ‘normalised’ redshift quality factors $nQ \geq 3$, which have a probability of being correct of $> 90$ per cent (Driver et al., 2011).

The internal GAMA catalogues used for this work are: TilingCatv16 (GAMA-I source catalogue and spectroscopic redshifts; Baldry et al., 2010), TilingCatv31 (GAMA-II source catalogue including automatically-generated redshifts; Baldry et al., in prep), StellarMassesv08 (stellar mass catalogue; Taylor et al., 2011) and ApMatchedPhotomv02 (photometric catalogue; Driver et al., in prep).

### 4.3.2 21 cm observations

Low redshift ($z < 0.0423$) 21 cm data over the GAMA fields is provided by HIPASS. This all-sky (Dec. $< +25^\circ$) survey was conducted with the multibeam receiver on the
4.3. THE DATA

CHAPTER 4. H I STACKING WITH GAMA

Figure 4.2: The rms per spectral channel for the gridded G09 (black) and G15 (red) Parkes data cubes.

64 m Parkes radio telescope from 1997 to 2002. See Barnes et al. (2001) and Wong et al. (2006) for further details of the observations, data reduction and final data products. The HIPASS data cubes have an rms noise level of 13.3 mJy at Dec. < +2° (Barnes et al., 2001) and ~14 mJy at Dec. > +2° (Wong et al., 2006). For the purpose of this work, we mask all HIPASS data between 1418.4 and 1422.4 MHz to avoid contamination by Galactic H I.

Higher redshift 21 cm observations of G09 were conducted with the Parkes telescope over 83.5 h on 2010 January 31 - February 15, 79 h on 2010 April 1 - 15 and 12.5 h on 2012 November 5 - 8 giving a total of 175 h. Due to overheads, we estimate that 104 h of on-source integration was obtained. G15 was observed for 78 h over 2012 March 20 - 31 with an estimated 61 h of on-source integration.

An identical observing strategy to that of D13 was used. Two 64 MHz bands were centred on 1285 MHz and 1335 MHz, providing a 14 MHz band overlap to reduce the impact of the bandpass subtraction. This resulted in a bandwidth coverage of 1253-1367 MHz and a 62.5 kHz frequency channel spacing. A basket-weave scanning
pattern was used. The data were reduced with LIVEDATA and Hanning smoothing was applied, resulting in a frequency resolution of 125 kHz (26.4 km/s at $z = 0$).

Radio frequency interference (RFI) in the data was identified and flagged using the default strategy of the RFI-excision software AOFLAGGER (Offringa et al., 2010, 2012). AOFLAGGER employs a Gaussian high-pass filter, iterative ‘SumThreshold’ method and scale-invariant rank operator and is one of the most efficient and robust RFI masking pipelines available (Offringa et al., 2010, 2012).

The RFI occupancy in the data per spectral channel is shown in Figure 4.1. In the 1335 MHz band, only two major narrow-band RFI signals are present at 1312 MHz and 1316 MHz. The 1285 MHz band is more heavily afflicted. There is a wide band of severe RFI centred on ~1275 MHz. It is caused by a navigation satellite and worsened over time as the satellite system advanced. It is present in all beams and both polarizations. Much of the data in these channels is flagged in the G09 observations and 100 per cent of the G15 data (which were collected later) below ~1285 MHz is unusable.

A data cube with pixel widths $4 \times 4$ arcmin$^2$ was constructed for each of the two fields using GRIDZILLA$^1$. Flux densities were calibrated against 1934-638 and Hydra A. To reduce the impact of the bad bandpass shape, only data at $\leq 1310$ MHz in the lower spectral band and at $> 1310$ MHz in the upper spectral band were used. Complete coverage over the full 114 MHz bandwidth was still achieved due to the overlap of the band placement. The spatial resolution of both the gridded HIPASS data and the new, higher redshift gridded Parkes data is 15.5 arcmin.

Figure 4.2 shows the rms per spectral channel in the final, higher redshift gridded cubes. The rms of the G15 data is higher at all frequencies due to the shorter integration time. Due to RFI flagging in the lower spectral band, the rms of the remaining data is significantly higher than in the relatively RFI-free upper band.

---

$^1$LIVEDATA and GRIDZILLA are supported by the Australia Telescope National Facility and are available at www.atnf.csiro.au/computing/software/livedata/
The average rms in channels above 1310 MHz is 4 mJy for the G09 data and 5.6 mJy for the G15 data.

4.4 Analysis

4.4.1 Source selection

4.4.1.1 Selection based on redshift and position

We identify all G09 and G15 galaxies with available magnitudes, stellar masses and photometric information and with GAMA-I redshifts within the volume defined by the H I data cube. We also require that the GAMA-II redshift is in this range. This eliminates the few sources (~0.2 per cent) with extremely incorrect redshifts. Their redshifts may have been misidentified during manual GAMA-I redshifting processes (Driver et al., 2011), but are likely correct in the GAMA-II catalogue, which uses an automated cross-correlation method (Baldry et al., 2014).

Using this criteria, we create our ‘low redshift’ sample by identifying all such GAMA sources with \( z < 0.04 \) that will be present within the HIPASS data. We also apply a \( z > 0.003 \) cut to exclude misclassified stars and sources contaminated by emission from the Galaxy. There are 1,403 GAMA-I galaxies (over both G9 and G15) available for stacking in this range.

To create a ‘higher redshift’ sample, we identify GAMA sources available within the new Parkes data. The rms of the G15 data below 1310 MHz is insufficient for a stacking analysis due to RFI (see Figure 4.2). Therefore, we exclude this data from all further work and only select G15 galaxies at \( z < 0.084 \). Similarly, G09 data below 1280.5 MHz is poor quality and we therefore only select G09 galaxies at \( z < 0.109 \). Excluding a small number of channels at the high frequency edge of the spectral range due to bad bandpass shape, our low redshift criterion for both G09 and G15
is $z > 0.0396$. Additionally, any galaxies with redshifts corresponding to the RFI-affected channels at 1312 and 1316 MHz are rejected from the sample. A total of 6,769 GAMA galaxies remain in this higher redshift sample.

To reduce the impact of residual continuum emission in our data, we excluded all sources within a beam width of the 25 strong ($S_{1.4 \text{GHz}} > 200 \text{ mJy}$) continuum sources in the G09 field and the 33 in G15. These continuum sources were identified in the NRAO VLA Sky Survey (NVSS; Condon et al., 1998). We also excluded any galaxies within half a beamwidth of the field edges, to reduce edge effects that would be introduced by the confusion correction described in section 4.4.3 below. Our final ‘higher redshift’ ($z \gtrsim 0.04$) sample of combined G09 and G15 galaxies consists of 4,917 objects, and our final ‘low redshift’ sample ($z \lesssim 0.04$) contains 995 objects.

### 4.4.1.2 Completeness and sample biases

Our first goal is to examine the evolution of the cosmic H i mass density ($\Omega_{\text{H}i}$) with redshift. Selection biases present in the sample as a function of redshift may inhibit our ability to identify evolution in the overall galaxy population. The availability of the multiwavelength, well-selected GAMA sample over the observed fields allows us to identify and minimise such sample biases. In section 4.4.4 below, we will use both an H i mass-to-light ratio and an H i-to-stellar mass ratio to calculate the cosmic H i mass density. It is therefore particularly important to examine the completeness and biases in the $r$ band luminosities ($L_r$) and stellar masses ($M$) of our sample, as a function of redshift.

We use stellar masses and $r$ band absolute magnitudes provided in the GAMA-I stellar masses catalogue (Taylor et al., 2011). The absolute magnitudes have been $K$-corrected and extinction corrected. We convert these to luminosities using a solar magnitude of $M_{\odot,r} = 4.71$, as in Hill et al. (2010). Figure 4.3a shows $L_r$ versus redshift for the G09 and G15 objects defined in section 4.4.1.1 above. The vertical dotted lines show the overlap region between the low redshift sample (where H i data
4.4. ANALYSIS

CHAPTER 4. H I STacking WITH GAMA

Figure 4.3: The $r$ band luminosity (a) and stellar mass (b) versus redshift for all G09 and G15 sources available for stacking. The colour scale represents $g - i$ galaxy colour. The black dashed lines indicate the completeness limit in each case. In (a), this is defined by the GAMA-I survey magnitude limit ($r < 19.4$). In (b), the stellar mass completeness function has been derived by Robotham et al., (in prep) and minimises residual colour bias. Red dashes show the range over which our data extend beyond that studied by Robotham et al. We have conservatively extrapolated to lower redshift. The $L_r$ or $\mathcal{M}$ selected samples used for the rest of the analysis consist of only the objects above the indicated limits. Vertical dotted lines show the redshift range over which both the low and high redshift Parkes observations overlap.
is provided by HIPASS) and the high redshift sample (with H I data provided by the new Parkes observations). The corresponding $r_{\text{petro}} < 19.4$ completeness limit of GAMA-I is indicated by the dashed black line. Some points lie below this line since the apparent magnitudes derived by Taylor et al. (2011) using SED template fitting differ slightly to the SDSS magnitudes used to select the input sample for the GAMA survey. We define a $L_r$ complete sample which excludes the 24 sources lying below this line. By defining such a $L_r$-selected sample, we are able to use the luminosity function to make appropriate corrections later in the analysis (e.g. the confusion correction presented in section 4.4.3). The final $L_r$-selected catalogue contains 990 galaxies in the $z < 0.04$ sample and 4,898 galaxies in the $z > 0.04$ sample.

It is well established in the literature that H I mass content is closely related to galaxy colour. For example, D13 showed that a stacking analysis reproduces the known trend that, on average, bluer galaxies (with $b_J - r_F \approx 0.3$) have $\sim 12$ times larger H I gas fractions than redder galaxies (with $b_J - r_F \approx 1.4$). To be confident that any observed gas trends are caused by genuine evolution in the galaxy population, we must first
ensure that there is no colour bias with redshift in the GAMA source selection.

Figure 4.3b shows the variation of $M$ with redshift. The colour scale indicates the $g - i$ colour of each galaxy as defined in the stellar mass catalogue presented in Taylor et al. (2011). Robotham et al., (in prep) define a stellar mass selection function which produces a sample with minimal colour bias. This was determined by finding the 95$^{th}$ percentile of the colour distribution per stellar mass bin, as a function of redshift. This is indicated by the black dashed line in Figure 4.3b. Robotham et al., (in prep) only define this function down to $z = 0.014$. We conservatively extrapolate horizontally to define the limit at all redshifts below this.

Down to this limit, the sample does not display any significant colour trend, i.e. for a given stellar mass, the mix of galaxy colours stays roughly constant with redshift. Below this limit, however, redder galaxies of a given stellar mass have been excluded from selection at higher redshifts. We note that this artificial selection bias is different to the evolutionary trend towards bluer galaxies at higher redshifts. We therefore define a $M$-selected sample consisting of only the galaxies above this limit. Our final $M$-selected sample consists of the 654 galaxies with $z < 0.04$ and 3,361 with $z > 0.04$.

Down to the completeness limit, no significant colour bias exists for a given luminosity as a function of redshift (as can be seen in Figure 4.3a). Therefore, no further selection criteria are required for the $L_r$-selected sample. In addition to colour, we note that H$\I$ content may also vary with other observable properties such as surface brightness. However, we find no obvious surface brightness bias in our sample.

### 4.4.1.3 Summary of source selection

To summarise, we have constructed four samples. We have a low redshift ($0.003 < z < 0.042$) sample where HIPASS data is available. New Parkes observations provide H$\I$ data for a higher redshift sample ($z > 0.039$). RFI issues restrict the upper redshift limit of the data to $z = 0.109$ in G09 and $z = 0.084$ in G15. Within each
Table 4.1: Summary of source counts in each subsample used for this analysis.

<table>
<thead>
<tr>
<th></th>
<th>Low redshift</th>
<th>High redshift</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$(0.003 &lt; z &lt; 0.042)$</td>
<td>$(0.039 &lt; z &lt; 0.109)$</td>
</tr>
<tr>
<td>$L_r$-selected</td>
<td>990</td>
<td>4,898</td>
</tr>
<tr>
<td>$M$-selected</td>
<td>654</td>
<td>3,361</td>
</tr>
</tbody>
</table>

Figure 4.5: The redshift distribution of all G09 and G15 sources suitable for stacking in the low redshift and high redshift populations. The $L_r$-selected samples are displayed with circular hatching, while the $M$-selected samples are displayed with diagonal hatching, as indicated in the legend.

of the high and low redshift catalogues, a $L_r$-selected and a $M$-selected sample are defined such that the sample is complete and there is no colour bias with redshift. Table 4.1 summarises the source counts in each catalogue.

Note that the $M$-selected samples are subsets of the $L_r$-selected samples in each case. The positions of the full $L_r$-selected sample are shown in Figure 4.4 and the redshift distributions are shown in Figure 4.5. The average redshifts of the low and high samples are 0.029 and 0.072, respectively. The spatial and spectral clustering are real features and highlight the importance of cosmic variance considerations on the results. This clustering is a well known feature within the GAMA fields - see, for example, Figure 5 in Baldry et al. (2012).
4.4. ANALYSIS

4.4.2 Stacking H I spectra

For each source in our sample, we extract the H I spectrum from the Parkes data at the central pixel corresponding to the GAMA source position. As in D13, this extraction involves a beam weighting process that results in an effective beamwidth of 21.2 arcmin for the new Parkes data and 21.9 arcmin for the HIPASS data (the difference is due to the slightly different beam parameters used in the gridding procedures for each). There are 24 known H I sources at \( z < 0.04 \) across G09 and G15, and their positions are defined in the HIPASS source catalogue (Meyer et al., 2004; Wong et al., 2006). No direct 21 cm detections of any of the GAMA galaxies appear in the higher redshift Parkes data at the > 5 \( \sigma \) level.

We now wish to calculate physical quantities for our samples such as the average H I mass \( \langle M_{\text{H} \text{I}} \rangle \), the average H I mass-to-light ratio \( \langle M_{\text{H} \text{I}}/L_r \rangle \) and the average H I-to-stellar mass ratio \( \langle M_{\text{H} \text{I}}/M \rangle \). Since the majority of sources are undetected at 21 cm, we must invoke a stacking analysis to do so. We follow the same methodology as in D13. Please refer to this paper for a more detailed description of the stacking process. To summarise, we first convert each flux density spectrum \( (S_{\nu,\text{obs}}) \) to an H I mass spectrum \( (M_{\text{H} \text{I},\nu,\text{obs}}) \) via:

\[
\left( \frac{M_{\text{H} \text{I},\nu,\text{obs}}}{M_\odot \text{ MHz}^{-1}} \right) = 4.98 \times 10^7 \left( \frac{S_{\nu,\text{obs}}}{\text{Jy}} \right) \left( \frac{D_L}{\text{Mpc}} \right)^2,
\]

where \( D_L \) is the luminosity distance of the galaxy. These mass spectra can be converted into gas mass fraction ‘spectra’ by dividing each channel of a mass spectrum by the corresponding \( L_r \) or \( M \) of that source.

We then align each spectrum at rest frame (1420.406 MHz; \( M_{\text{H} \text{I},\nu,\text{em}} \)) and ensure that total mass is conserved (see D13 equation 2). The final stacking process involves taking the weighted average of the \( n \) aligned spectra:
\[
\langle M_{\text{H}1} \rangle_{\nu_{\text{em}}} = \frac{\sum_{i=1}^{n}(w_i M_{\text{H}1,i,\nu_{\text{em},i}})}{\sum_{i=1}^{n}w_i}.
\]

Again, \(M_{\text{H}1}\) can be replaced by a gas mass fraction. In D13, we chose to use the weighting factor \(w_i = \sigma_i^{-2}\), where \(\sigma_i\) is the rms noise level of the \(i^{th}\) flux spectrum. We find that this weighting scheme is insufficient for this work since the GAMA sample contains many more low stellar mass systems with low \(\text{H}1\) gas content. Here we choose to use \(w_i = \sigma_i^{-2}D_{L,i}^{-1}\) which provides a good signal-to-noise ratio (S/N). A comparison of the resulting spectra using these two weighting schemes is shown in Figure 4.6. Particularly at \(z < 0.04\), the S/N is insufficient in some cases. While \(w_i = \sigma_i^{-2}D_{L}^{-4}\) would maximise the S/N, this biases the sample too heavily towards low redshift where cosmic variance can become an issue, particularly for the \(z < 0.04\) sample.

We achieve clear stacked detections in all cases using \(M_{\text{H}1}\), \(M_{\text{H}1}/L_r\) and \(M_{\text{H}1}/M\). These are shown in Figure 4.6 for both the low and high redshift samples. The \(L_r\)-selected sample has been used to construct the average mass and average mass-to-light spectra. The \(M\)-selected sample has been used to construct the average \(\text{H}1\)-to-stellar mass spectra. We note that the gas fraction spectra have been corrected for beam confusion prior to stacking (see section 4.4.3). A fourth-order polynomial baseline has been subtracted in each case. The significance of each detection is indicated on the plots.

The typical widths of the stacked spectra are around 300-400 km s\(^{-1}\). In many of the co-added profiles, an additional extension to frequencies below rest frame (on the order of \(\sim 100\) km s\(^{-1}\)) is seen. We do not believe this is caused by spectroscopic redshift inaccuracy. Driver et al. (2011) quotes an accuracy of 65 km s\(^{-1}\), however a comparison to SDSS redshifts suggests that for galaxies with \(nQ \geq 3\), the average redshift error may be closer to 100 km s\(^{-1}\). This is still less than the expected median \(\text{H}1\) profile width of the individual galaxies (see section 4.4.3 below for Tully-Fisher predictions of these widths). Maddox et al. (2013) show that redshift errors do not dominate the width of the stacked spectrum if these errors are less than the median
4.4. ANALYSIS  

CHAPTER 4. H I STACKING WITH GAMA

Figure 4.6: The co-added spectrum produced by stacking H I mass (top), H I mass to r band luminosity (middle) and H I-to-stellar mass (bottom) spectra for $z < 0.04$ (left) and $z > 0.04$ (right) redshift galaxies. The black line uses our chosen weighting scheme of $w_i = \sigma_i^{-2} D_{L,i}^{-1}$. The magenta line uses the weighting scheme of D13 - $w_i = \sigma_i^{-2}$. The vertical dotted line indicates rest frame and the red dashed lines are plotted at 300 km s$^{-1}$ either side of this. Individual $M_{HI}/L_r$ and $M_{HI}/M$ spectra have been corrected for confusion (via the method described in section 4.4.3) before stacking. The significance of each detection (for the black line only) is stated for each figure. The green line in each case shows an example of a mock stacked spectrum created using the control catalogue. This indicates the noise level.
width of the individual H I profiles. Furthermore, Maddox et al. (2013) compare the redshifts derived from H I data and from various optical lines and find no offsets larger than $\sim 10 \text{ km s}^{-1}$.

Rather, the asymmetry is likely due to a poor baseline subtraction. This may be partially due to residual effects of continuum emission in our data. Although we have been careful to exclude spectra severely affected by such emission, there remain many continuum sources with $S_{1.4} < 200 \text{ mJy}$ in the field, which may cause ripples in the spectra and discontinuities between the bands (Barnes et al., 2001). As we will see below, there is a high level of confusion in the H I data, the extent of which increases with redshift. There is also a considerable amount of clustering of galaxies in the field, as seen in Figures 4.4 and 4.5. Therefore, a large number of spectra may be influenced by a single continuum source. This signal will not be reduced by the shifting process during stacking, but rather will be amplified by the $D_L^2$ factor in the conversion between 21 cm flux and H I mass.

Although, this does not explain why the asymmetry (which was also seen prominently at $z > 0.04$ in D13) is always towards lower frequencies whenever it is present. While this could hint at some sort of offset in the Parkes H I data, Maddox et al. (2013) compared the H I redshifts of HIPASS and ALFALFA sources and found no offset. The exception was for objects that were blended in the beam and spectroscopically confused, in which case offsets of up to $60 \text{ km s}^{-1}$ were possible.

To find the average quantities $\langle M_{\text{H}I} \rangle$, $\langle M_{\text{H}I}/L_r \rangle$ and $\langle M_{\text{H}I}/M \rangle$, we must integrate the stacked spectra shown in Figure 4.6 over the profile bounds. However, as discussed in detail in D13, the maximum expected linewidth for a single galaxy is $\sim 600 \text{ km s}^{-1}$. Therefore, all the signal beyond $\pm 300 \text{ km s}^{-1}$ of rest frame is likely due to confused sources, i.e. the H I signals of nearby galaxies are ‘contaminating’ the spectrum of the target galaxy. Integrating over these ‘wings’ in the stacked spectrum would result in an overestimated average quantity. We therefore restrict the integral to within $\pm 300 \text{ km s}^{-1}$ of rest frame. These bounds are indicated in Figure 4.6. Limiting the integration bounds will also reduce the extent to which the continuum ripple can
influence the results. We note, however, that the measured averages are still expected to be overestimates due to confusion even within the $\pm 300\,\text{km}\,\text{s}^{-1}$ range. While it is difficult to account for this in the $\langle M_{\text{H}_1} \rangle$ measurement, section 4.4.3 will show how the gas mass fractions can be corrected.

Table 4.2 summarises the average H$\,\text{I}$ masses derived for the $L_r$ and $M$-selected samples at the lower and higher redshifts. At low redshift, the average masses are higher for the $M$-selected samples. This is unsurprising since the $M$ completeness cut preferentially selects the higher mass objects from the $L_r$-selected sample. At $0.04 < z < 0.1$, the average H$\,\text{I}$ masses are consistent within the errors between the $L_r, M$-selected samples. Over this same range, the average H$\,\text{I}$ masses measured for the GAMA samples are more than 2 times lower than those measured at $0.04 < z < 0.13$ in the 2dFGRS sample of D13 ($6.93 \pm 0.17 \times 10^9\,\text{M}_\odot$). This is caused by the fact that GAMA samples fainter, and therefore lower-mass, galaxies at a given redshift, due to the better sensitivity. The average H$\,\text{I}$ masses measured at $z < 0.04$ for the GAMA field are almost double that measured for the 2dFGRS sample of D13 over the same redshift range ($1.48 \pm 0.03 \times 10^9\,\text{M}_\odot$). This is likely due to cosmic variance - the low redshift GAMA sample covers a much smaller volume than the low redshift but wide-field 2dFGRS sample ($3.9 \times 10^4\,\text{Mpc}^3$ versus $3.4 \times 10^5\,\text{Mpc}^3$).

As in D13, the errors on the measured average quantities were calculated by stacking control spectra. Ten control catalogues were created by randomly mismatching the GAMA positions and redshifts in the original catalogue. Mock stacked spectra were then generated using these control catalogues. Of course, this produces no statistical detection since incorrect redshifts were used for each source. Such control co-added spectra are overlaid in Figure 4.6 to give an idea of the noise level. The error on the average quantity of interest is found by integrating each mock spectrum over the $\pm 300\,\text{km}\,\text{s}^{-1}$ range and then determining the rms deviation on the result.
### 4.4.3 Correction for confusion

As in D13, a significant challenge is to accurately correct for source confusion in the radio data. Many individual galaxies are located within the large Parkes beam. If these galaxies are also at similar redshifts, we may unknowingly be integrating over multiple \( \text{H} \text{I} \) profiles per spectrum. Therefore, we will be overestimating the measured average \( \text{H} \text{I} \) mass and therefore also the average \( \text{H} \text{I} \) mass-to-light and \( \text{H} \text{I} \)-to-stellar mass ratios. As will be seen in section 4.4.4, these are important quantities for the calculation of \( \Omega_{\text{H} \text{I}} \).

We follow the method of D13 to correct for this confusion. We apply the Tully-Fisher relation to the \( r \) band magnitudes to predict the \( \text{H} \text{I} \) profile widths expected for each individual galaxy. This allows us to estimate which galaxies will be confused over the spectral range of interest.

Toribio et al. (2011) examined ALFALFA galaxies and found a relationship between \( W_{50, \sin i} \), the full-width at half-maximum of the \( \text{H} \text{I} \) profile, and absolute \( r \)-band magnitude. Converting to luminosity, \( L_r \), and our cosmology, this relationship is:

\[
\log W_{50, \sin i} = (\log L_r - 2.11)/3.26
\]  

(4.3)

Note that \( W_{50, \sin i} \) is the intrinsic linewidth of the galaxy and must be multiplied by \( \sin i \) to achieve the observed line width, where \( i \) is the inclination of each galaxy. We calculate \( i \) using the same method as in D13 and Meyer et al. (2008):

<table>
<thead>
<tr>
<th>Subsample</th>
<th>( z &lt; 0.04 )</th>
<th>( z &gt; 0.04 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_r )-selected</td>
<td>( 1.91 \pm 0.08 )</td>
<td>( 2.56 \pm 0.09 )</td>
</tr>
<tr>
<td>( M )-selected</td>
<td>( 2.36 \pm 0.06 )</td>
<td>( 2.48 \pm 0.09 )</td>
</tr>
</tbody>
</table>

Table 4.2: Average \( \text{H} \text{I} \) masses \((\times 10^9 \text{M}_\odot)\) calculated by stacking galaxies in each subsample.
4.4. ANALYSIS

CHAPTER 4. HI STACKING WITH GAMA

Figure 4.7: The HI profile widths of individual GAMA galaxies in the $L_r$-selected samples, as predicted by the Tully-Fisher relation.

$$\cos^2(i) = \frac{(b/a)^2 - \epsilon^2}{1 - \epsilon^2}, \quad (4.4)$$

where $a$ and $b$ are the semimajor and semiminor axes of the galaxy and $\epsilon$ is the aspect ratio of an edge-on spiral, taken to be 0.12 (Masters et al., 2003). We have used the values of $a$ and $b$ defined in the photometric GAMA-I catalogue (Driver et al., in prep).

The distribution of estimated linewidths of the individual HI profiles of each galaxy in the full $L_r$-selected sample are shown in Figure 4.7. The average predicted widths are $103 \text{ km s}^{-1}$ at $z < 0.04$ and $143 \text{ km s}^{-1}$ at $z > 0.04$ (where preferentially larger galaxies with higher rotation velocities will be sampled). These seem reasonable in comparison to the distribution of line widths of detected HIPASS galaxies, and the average Tully-Fisher predicted widths in D13 ($100$ and $150 \text{ km s}^{-1}$).
Using the known spatial positions and predicted line widths, we estimate that, on average, any one galaxy in the higher redshift sample is confused with nine other GAMA galaxies. This means that an average of nine other galaxies are within the effective Parkes beamwidth and will have H i profiles within ±300 km s$^{-1}$ of the galaxy in question. For the lower redshift sample, the average number of confused galaxies is four. This is more confusion than was identified in the 2dFGRS analysis of D13 since GAMA samples fainter galaxies and therefore contains more galaxies per square degree.

Once confused sources are identified, we can account for the overestimate in the average gas fractions by artificially confusing the $L_r$ and $\mathcal{M}$ to match the confusion in the H i masses.

As in D13, the confusion adjustment is carried out using:

$$X' = f_X \sum_{i=1}^{n} X_i b_i + X_0,$$

where $X_i$ is either the $L_r$ or $\mathcal{M}$ of the $i^{th}$ galaxy confused in the H i data, $X_0$ is the corresponding parameter of the central galaxy in question, $X'$ is the corresponding confusion-adjusted parameter, $b_i$ is a beam weighting factor and $n$ is the total number of sources confused with the galaxy in question. The extrapolation factor, $f_X$, accounts for galaxies that are not in the GAMA source catalogue because they are beyond the $L_r$ or $\mathcal{M}$ completeness limits of the survey, but will still contribute to confusion in the H i data. In D13 we assumed one fixed value of $f_{L_r}$ and applied it to all sources within a given redshift interval. In this work, the completeness function is well-characterised and so we can calculate it for each individual galaxy as follows.

At the redshift of each central galaxy, we know the minimum luminosity and stellar mass to which the GAMA catalogue is complete ($L_{lim}$ and $\mathcal{M}_{lim}$; the dashed lines in Figure 4.3). We can calculate the fraction of the total integrated luminosity or stellar mass density function that is probed down to this limit and thus the fraction.
of all galaxies we expect to miss due to these survey limits. The inverse of this is the completeness extrapolation factor, $f_{L_r}$ or $f_M$, required to account for these sources missed in the confusion adjustment. A detailed description of the $f_X$ calculation and the density functions used for this are presented in Appendix B. We use the $L_r$ density function derived by Driver et al. (2012) and the $M$ density function derived by Baldry et al. (2012).

The distribution of the resulting $f_X$ scale factors required are shown in Figure 4.8. This shows that the contribution of galaxies below the GAMA magnitude limit to the overall $L_r$ and $M$ confusion correction is relatively small. The correction factor is less than 16 per cent in all cases and is always less than 4 per cent for the low redshift sample, where the sample is complete to lower luminosities and stellar masses. The original $L_r$ and $M$ distributions are shown in Figure 4.9 along with the corresponding artificially-confused distribution (i.e. $L'_r$ and $M'$ - the output of equation 4.5). The average original and corrected $L_r$ and $M$ are quoted in Tables 4.3 and 4.4 across both


**Table 4.3:** Average properties derived for the lower and higher redshift $L_r$-selected GAMA samples. Asterisks indicate values that will be overestimated due to confusion. Dashes indicate quantities that have been corrected for confusion via the process described in section 4.4.3.

<table>
<thead>
<tr>
<th></th>
<th>$z &lt; 0.04$</th>
<th>$z &gt; 0.04$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle L_r \rangle \left( \times 10^9 \ h^{-2} \ L_\odot \right)$</td>
<td>2.526</td>
<td>6.08</td>
</tr>
<tr>
<td>$\langle L'<em>r \rangle \left( \times 10^9 \ h^{-2} \ L</em>\odot \right)$</td>
<td>9.21</td>
<td>38.58</td>
</tr>
<tr>
<td>$\langle M_{HI} \rangle \left( M_\odot \ L_\odot^{-1} \right)$</td>
<td>3.21 ± 0.33</td>
<td>1.09 ± 0.04</td>
</tr>
<tr>
<td>$\langle M'<em>{HI} \rangle \left( M</em>\odot \ L_\odot^{-1} \right)$</td>
<td>0.77 ± 0.21</td>
<td>0.16 ± 0.02</td>
</tr>
<tr>
<td>$\langle M_{HI} \rangle \left( M_\odot \ L_\odot^{-1} \right)$</td>
<td>0.96 ± 0.04</td>
<td>0.45 ± 0.02</td>
</tr>
</tbody>
</table>

**Table 4.4:** Average properties derived for the lower and higher redshift $M$-selected GAMA samples. Asterisks indicate values that will be overestimated due to confusion. Dashes indicate quantities that have been corrected for confusion via the process described in section 4.4.3.

<table>
<thead>
<tr>
<th></th>
<th>$z &lt; 0.04$</th>
<th>$z &gt; 0.04$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle M \rangle \left( \times 10^9 \ h^{-2} \ M_\odot \right)$</td>
<td>4.86</td>
<td>12.07</td>
</tr>
<tr>
<td>$\langle M' \rangle \left( \times 10^9 \ h^{-2} \ M_\odot \right)$</td>
<td>11.04</td>
<td>37.86</td>
</tr>
<tr>
<td>$\langle M_{HI} \rangle \left( M_\odot \ M_\odot^{-1} \right)$</td>
<td>4.15 ± 0.27</td>
<td>0.72 ± 0.03</td>
</tr>
<tr>
<td>$\langle M'<em>{HI} \rangle \left( M</em>\odot \ M_\odot^{-1} \right)$</td>
<td>2.19 ± 0.19</td>
<td>0.15 ± 0.01</td>
</tr>
<tr>
<td>$\langle M_{HI} \rangle \left( M_\odot \ M_\odot^{-1} \right)$</td>
<td>0.61 ± 0.02</td>
<td>0.22 ± 0.01</td>
</tr>
<tr>
<td>$\langle M'<em>{HI} \rangle \left( M</em>\odot \ M_\odot^{-1} \right)$</td>
<td>0.28 ± 0.01</td>
<td>0.071 ± 0.003</td>
</tr>
</tbody>
</table>

129
redshift ranges. The confusion-corrected values ($L'_r$ and $M'$) can now be used when creating the gas fraction ‘spectra’ (i.e. the middle and lower panels of Figure 4.6) and when calculating the subsequent average gas fractions. These gas fractions will be closer to the true, unconfused values, compared to the fractions calculated using the original luminosities and stellar masses. Both the original and confusion-adjusted gas mass ratios are also quoted in Tables 4.3 and 4.4. As expected, both the $\langle M_{\text{HI}}/L'_r \rangle$ and $\langle M_{\text{HI}}/M' \rangle$ are smaller for the higher redshift sample which will preferentially include more luminous, higher stellar mass objects than the low redshift sample. The $\langle M_{\text{HI}}/M' \rangle$ found for the lower redshift sample is slightly larger than that found using direct detections in ALFALFA (1.5; Huang et al., 2012) over a similar range. This is a result of the high $M$ limit imposed on the $M$-selected sample.

### 4.4.4 Cosmic H I mass density

To calculate the cosmic H I mass density ($\Omega_{\text{HI},1}$), we use two methods:

$$\Omega_{\text{HI},1} = \langle \frac{M_{\text{HI}}}{X'} \rangle \times \mathcal{C}_{1,X} \times \frac{\rho_X}{\rho_c},$$

(4.6)
and

$$
\Omega_{H_1,2} = \frac{\langle M_{H_1} \rangle}{\langle X' \rangle} \times c_{2,X} \times \frac{\rho_X}{\rho_c},
$$

(4.7)

where \( \rho_c = \frac{3H_0^2}{8\pi G} \) is the critical density of the Universe, and \( X \) and \( X' \) are the original and confusion-adjusted \( L_r \) or \( M \) values. The luminosity density, \( \rho_{L_r} \), is \((2.29 \pm 0.06) \times 10^8 \, L_\odot \, h \, \text{Mpc}^{-3}\) (Driver et al., 2012) and the stellar mass density, \( \rho_M \), is \((3.2 \pm 0.5) \times 10^8 h \, M_\odot \, \text{Mpc}^{-3}\) (Baldry et al., 2012).

The fraction \( \langle M_{H_1} \rangle / \langle X' \rangle \), required for equation 4.6, is found by stacking \( H_1 \) mass spectra, then dividing by the weighted mean of the confusion-adjusted \( L_r \) or \( M \) of the whole sample. These weighted means are calculated in a similar manner to equation 4.2 using \( w_i = \sigma_i^{-2} D_{L,i}^{-1} \). The fraction \( \langle M_{H_1} / X' \rangle \), required for equation 4.7 is found by stacking \( M_{H_1} / X' \) 'spectra' (shown in Figure 4.6). All average values are quoted in Tables 4.3 and 4.4.

The correction factors, \( c_{1,X} \) and \( c_{2,X} \), are applied to each of these values to account for the fact that the weighted luminosity (and stellar mass) distributions of the galaxy samples being stacked may differ from the true, volume-limited, galaxy population. This is an issue since \( H_1 \) mass is a function of both luminosity and stellar mass, and so any selection effect dependent on these quantities will lead to a bias away from the true mass-to-light values required by equations 4.6 and 4.7. The required correction factors are given by:

$$
c_{1,X} = \frac{\int W(X)dX}{\int X^\alpha W(X)dX} \times \frac{\int X^\alpha \rho(X)dX}{\int \rho(X)dX},
$$

(4.8)

and
$C_{2,X} = \frac{\int XW(X)dX}{\int X^{\alpha+1}W(X)dX} \times \frac{\int X^{\alpha}\rho(X)dX}{\int \rho(X)dX},$  \hspace{1cm} (4.9)

where $\rho(L_r)$ is the luminosity density function, $\rho(M)$ is the stellar mass density function and $W(X)$ is the weight distribution of the sample. The latter is essentially the sum of the weights ($w_i$ used in equation 4.2) for all sources in a particular $X$ bin. These are characterised by Schechter and double-Schechter functions, respectively, and are presented in Appendix A. The $\Omega_{H_1}$ calculation of D13 made use of equation 4.6. A recent stacking analysis by Rhee et al. (2013) using Westerbork Synthesis Radio Telescope data across a similar redshift range used equation 4.7. However, these studies used $N(X)$ in place of $W(X)$, where $N(X)$ is the number distribution of the input sample. This is sufficient when weighting factors, $w_i$, are based only on rest-frame quantities such as the rms in the flux spectrum (as was the case in D13 and Rhee et al., 2013). The weight distribution $W(X)$ is required with distant-dependent $w_i$ factors, such as used here (see section 4.4.2).

In equations 4.8 and 4.9, the factor $\alpha$ relates $M_{H_1}$ and $X$:

$$\frac{M_{H_1}}{X} \propto X^\alpha \hspace{1cm} (4.10)$$

We derive these factors from the low redshift ($z < 0.04$) GAMA data. Figure 4.10 shows the $\langle M_{H_1}/L_r' \rangle$ and $\langle M_{H_1}/M' \rangle$ derived by stacking the low redshift $L_r$-selected sample within bins of $L_r$, and the $M$-selected sample within bins of $M$, respectively. We perform a least-squares fit to the data in each case. Due to low S/N in the co-added $M_{H_1}/L_r$ spectra, the data can only be split into four $L_r$ bins resulting in larger errors on the fitted parameters than in the $M$ case. For $X = L_r$, we find $\alpha = -0.4 \pm 0.1$. This is consistent with the value of $-0.40$ found by Toribio et al. (2011) using an
Figure 4.10: Left: The average $M_{\text{HI}}/L_r$ in four $L_r$ bins. The fit to the points is indicated by the red dashed line and is $\log(M_{\text{HI}}/L_r) = -0.4 \log L_r + 3.2$. Right: The average $M_{\text{HI}}/M$ in eight $M$ bins. The fit to the data is $\log(M_{\text{HI}}/M) = -0.7 \log M + 6.4$.

<table>
<thead>
<tr>
<th>$z &lt; 0.04$</th>
<th>$z &gt; 0.04$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X = L_r$</td>
<td>$X = M$</td>
</tr>
<tr>
<td>$C_1$</td>
<td>$0.38$</td>
</tr>
<tr>
<td>$\Omega_{\text{H}_1}(\times10^{-4})$</td>
<td>$2.42 \pm 0.65$</td>
</tr>
<tr>
<td>$\bar{C}_2$</td>
<td>$0.96$</td>
</tr>
<tr>
<td>$\Omega_{\text{H}_1}(\times10^{-4})$</td>
<td>$2.21 \pm 0.11$</td>
</tr>
</tbody>
</table>

Table 4.5: The cosmic $\text{H} \, \text{I}$ mass densities and sample correction factors calculated for each subsample of GAMA galaxies.

$\text{H} \, \text{I}$ selected ALFALFA galaxy sample and luminosities from the SDSS. For $X = M$, we find $\alpha = -0.70 \pm 0.06$. This is consistent with Huang et al. (2012) who measure $\alpha = -0.724$ for ALFALFA galaxies with $\log(M) > 9$. We assume that these gas fractions also hold at $0.04 < z < 0.11$.

Table 4.5 summarises the correction factors and the derived values of $\Omega_{\text{H}_1,1}$ and $\Omega_{\text{H}_1,2}$ for both the low and high redshift samples. The $\Omega_{\text{H}_1}$ estimates are also compared to other values in the literature in Figure 4.11. The error bars on our measurements only show statistical error.
4.4. ANALYSIS  

CHAPTER 4. H I STACKING WITH GAMA

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zwaan et al. (2005)</td>
<td>$\langle M_{\text{HI}}/L \rangle$</td>
</tr>
<tr>
<td>Martin et al. (2010)</td>
<td>$\langle M_{\text{HI}} \rangle / \langle L \rangle$</td>
</tr>
<tr>
<td>Delhaize et al. (2013)</td>
<td>$\langle M_{\text{HI}}/M \rangle$</td>
</tr>
<tr>
<td>Freudling et al. (2011)</td>
<td>$\langle M_{\text{HI}} \rangle / \langle M \rangle$</td>
</tr>
<tr>
<td>Rhee et al. (2013)</td>
<td>$L_{\text{og}}$ et al. (2011)</td>
</tr>
</tbody>
</table>

Figure 4.11: The cosmic H I mass density versus redshift. Observational (points) and simulated (dashed line) measurements from the literature are as indicated in the legend. Open points shown in the legend are measurements from this work using GAMA samples. Open diamonds show $\Omega_{\text{HI}}$, measurements calculated using $\langle M_{\text{HI}}/X \rangle$ (equation 4.6) and open squares show values calculated using $\langle M_{\text{HI}} \rangle / \langle X \rangle$ (equation 4.7). GAMA values in grey use $X = M$ and in black use $X = L_r$. 

134
4.5 Discussion of results

The \( \Omega_{\text{HI}} \) values derived through this stacking analysis span a range of values. Importantly, we find consistency between the two different methodologies used in the derivation of \( \Omega_{\text{HI}} \): equation 4.6 used by D13 and equation 4.7 used by Rhee et al. (2013). These differ in the form of the average mass-to-light ratio used and consequently the appropriate correction factor. Although these two methods provided consistent results between D13 and Rhee et al. (2013), different sample selection had been used in each case. Here we have confirmed that applying the two separate methods to a common sample also gives consistent results.

At low redshift \( (z < 0.04) \), the use of \( M_{\text{HI}}/L_r \) fractions to derive \( \Omega_{\text{HI}} \) gives results which are consistent with direct detection techniques at \( z \approx 0 \) by Zwaan et al. (2005) and Martin et al. (2010). However, the use of \( M_{\text{HI}}/M \) fractions results in \( \Omega_{\text{HI}} \) values much higher than expected. While the \( L_r \)-selected sample contains more low mass systems than the \( M \)-selected sample, the correction factors in equations 4.8 and 4.9 should account for the selection bias. Rather, the discrepancy may largely be explained by cosmic variance; the \( z < 0.04 \ M \)-selected sample has a smaller number of galaxies and effective volume \( (1.09 \times 10^4 \text{ Mpc}^{-3}; \text{after weighting}) \) compared to the other samples used in this analysis. According to Driver & Robotham (2010), a cosmic variance of 46 per cent is predicted over a field this size. Indeed, we have confirmed that average \( M_{\text{HI}}/M \) measurements are particularly sensitive to the choice of weighting \( (w_i) \) used in equation 4.2. For example, using \( w_i = \sigma_i^{-2}D_L^{-4} \), which optimises the stacked S/N, results in a co-added spectrum which is effectively dominated by very few \( (\sim 7) \) very low redshift, high mass galaxies.

At \( 0.04 < z < 0.11 \), the \( \Omega_{\text{HI}} \) values derived from \( M_{\text{HI}}/M \) fractions are consistent with other measurements at similar redshifts. These include the stacking results of D13 and Rhee et al. (2013) and the direct detection results of Freudling et al. (2011) using the Arecibo Ultra Deep Survey. The measurements are also consistent with the \( z \approx 0 \) results of Zwaan et al. (2005) and Martin et al. (2010), implying no evolution...
in $\Omega_{H_1}$, over the past $\sim 0.9 \, h^{-1} \, \text{Gyr}$. Since the star formation rate is evolving rapidly over this epoch, this implies that the H$\textsc{i}$ gas, which is consumed in star formation processes, must be replenished by various accretion mechanisms. This can include the infalling of recycled gas expelled from the galaxy by supernovae driven winds, the cooling of ionised gas in the walls of supernovae-induced supershells and the accretion of pristine gas onto the halos of galaxies (e.g. Hopkins et al., 2008; Lagos et al., 2011; Davé et al., 2013). This is consistent with the $\Lambda$CDM picture of galaxy formation and evolution in which the total neutral gas content of galaxies is regulated by the balance between accretion and outflow of gas in galaxies (e.g. Lagos et al., 2011; Obreschkow & Rawlings, 2009). These theoretical models find that the slow evolution of the neutral gas content of galaxies, in addition to the decreasing surface density of gas as time progresses (which drive a decreasing H$_2$/H$\textsc{i}$ fraction) lead to little evolution of the global density of H$\textsc{i}$. The prediction of Lagos et al. (2011) is shown in Figure 4.11.

However, over the same redshift range we find that $\Omega_{H_1}$ values derived from the $L_r$-selected sample are 3 times lower than for the $M$-selected sample; a difference which is significantly larger than their combined errors. This systematic difference cannot be explained by cosmic variance since the effective volume of these higher redshift samples are large ($1.58 \times 10^5 \, \text{Mpc}^{-3}$) and the stacked averages are found to be robust to the choice of different weighting schemes. With a brighter magnitude cut ($r < 18.6$), we find that only 1 per cent of the systematic offset can be explained by the difference in the luminosity range spanned by the sample. The remainder of the difference is currently unexplained. It may, for example, also reflect the $k + z$ correction factors in GAMA, which Driver et al. (2012) assumed to be small over the redshift range measured here. These corrections affect luminosity much more than stellar mass.
4.6 Conclusions

We present an $\text{H} \text{I}$ stacking analysis of galaxies in the GAMA survey. We focus on the two 48 deg$^2$ GAMA fields at RA 09$^h$ and 14$^h$30$^m$. Low redshift ($z < 0.042$) $\text{H} \text{I}$ data over these regions is provided by HIPASS. We collected 175 h of higher redshift ($z > 0.039$) 21 cm observations over G09 and 78 h over G15 using the Parkes 64 m radio telescope. Severe RFI at low frequencies restricts the range of the useful data to $z < 0.109$ in G09 and $z < 0.084$ in G15.

We co-add the $\text{H} \text{I}$ spectra of 990 GAMA galaxies at $z < 0.04$ and 4,898 galaxies at $0.04 < z < 0.11$. We measure average $\text{H} \text{I}$ mass to $r$-band luminosity fractions of $\langle M_{\text{H} \text{I}}/L_r \rangle = 0.77 \pm 0.21 M_\odot/L_\odot$ and $0.16 \pm 0.02 M_\odot/L_\odot$ in the lower and higher redshift samples, respectively. Stellar masses are also available for all GAMA sources. Multiwavelength data available over the GAMA fields allows the construction of a stellar mass ($M$) selected sample for which colour bias with redshift is minimised. This ensures that any observed evolutionary trends in the $\text{H} \text{I}$ content with redshift are not simply a result of sample selection bias. We measure average $\text{H} \text{I}$-to-stellar mass fractions of $\langle M_{\text{H} \text{I}}/M \rangle = 2.19 \pm 0.19$ and $0.15 \pm 0.01$ over the lower and higher redshift ranges, respectively. When calculating the above gas mass fractions, the $L_r$ and $M$ have first been adjusted to match the significant extent of beam confusion experienced in the $\text{H} \text{I}$ data. We find that stacking very closely reproduces known trends in gas mass fractions with $L_r$ and $M$ established via direct detection methods.

Various strategies for calculating the cosmic $\text{H} \text{I}$ mass density $\Omega_{\text{H} \text{I}}$ from a stacking analysis are explored. The GAMA sample selection is well-characterised and allows us to carefully account for the incompleteness of the optical data when measuring $\Omega_{\text{H} \text{I}}$. We find that the two different weighting schemes used by Delhaize et al. (2013) and Rhee et al. (2013) to account for incompleteness give equivalent results. We find a systematic difference between the $\Omega_{\text{H} \text{I}}$ values derived using average $M_{\text{H} \text{I}}/L_r$ fractions and average $M_{\text{H} \text{I}}/M$ fractions. At low redshift, this may be due to cosmic variance but is currently unexplained in the higher redshift sample. However, it is unlikely
that these results reveal any real evolution in $\Omega_{\text{HI}}$ over $0 < z < 0.11$, consistent with the results of other observations and simulations.

Acknowledgments

We thank Bi Qing For, Alan Duffy, Danny Price and Scott Meyer for observing assistance. We are very grateful to staff at the Parkes radio telescope for assistance during the observing process, especially Ettore Carretti. The Parkes radio telescope is part of the Australia Telescope National Facility which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. We acknowledge the efforts of the HIPASS team.

JD thanks Andre Offringa for assistance with aoflagger, Claudia Lagos for useful discussions and the International Centre for Radio Astronomy Research and the University of Western Australia for financial support.

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is http://www.gama-survey.org/.

We acknowledge use of the TOPCAT software in our data analysis (www.starlink.ac.uk/topcat/).
Chapter 5

Herschel-ATLAS/GAMA: Infrared properties of bright radio galaxies in the science demonstration phase field

Foreword and context

This chapter examines the star formation rates of the host galaxies of radio-loud AGN. In Chapters 3 and 4, the stacking technique was applied to the spectral axis. In this chapter, stacking is conducted in the image plane, via the procedure described in Chapter 2, section 2.1.2. Applying this technique to newly available, sensitive far-infrared, near-infrared and optical data, allows a deep investigation of the influence of AGN activity on star formation and galaxy evolution.

This chapter is presented in the form of a paper in preparation for publication. A catalogue of selected parameters for all radio sources used in this paper can be found in Appendix C at the end of this thesis. Section 5.3.2 of this paper mentions VIKING image cutouts with overlaid radio contours. The full set of cutouts, with overlaid radio contours, are presented in Appendix D. This shows our decisions as to the structure and position of each radio galaxy in the sample.
5.1. ABSTRACT

CHAPTER 5. RADIO GALAXIES IN H-ATLAS

Authors and affiliations

J. Delhaize (1), M. J. Hardcastle (2), M.J. Jarvis (3,4), J.S. Virdee (3), T. Mauch (3,4), D. Bonfield (2), S. Rawlings (3), H-ATLAS team (various affiliations), GAMA team (various affiliations).

(1) International Centre for Radio Astronomy Research, University of Western Australia, 7 Fairway, Crawley 6009, Western Australia

(2) School of Physics, Astronomy and Mathematics, University of Hertfordshire, College Lane, Hatfield AL10 9AB

(3) Astrophysics, Department of Physics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH

(4) Physics Department, University of the Western Cape, Cape Town, 7535, South Africa

5.1 Abstract

We investigate the far-infrared and star-forming properties of radio galaxies in the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS) science demonstration phase field out to $z = 5$. We compile a complete sample of 250 radio galaxies with 1.4 GHz flux densities greater than 10 mJy using the NRAO VLA Sky Survey (NVSS) and Faint Images of the Radio Sky at Twenty Centimetres (FIRST) catalogues. We identify near-infrared and optical counterparts to the radio sources using the most sensitive available data in the field. This is provided by the VISTA Kilo-Degree Infrared Galaxy (VIKING) near-infrared imaging survey and new optical imaging data from the Isaac Newton Telescope (INT). We generate photometric redshifts for all sources and collect any spectroscopic redshifts available from the Galaxy and Mass Assembly (GAMA) redshift survey. We perform a stacking analysis of Herschel data...
to probe the monochromatic 250 $\mu$m infrared luminosity, and by proxy the star formation rate, of a deeper population of radio galaxies than previously possible. We demonstrate that statistical detections of far-infrared emission from radio-loud AGN are possible out to $z = 5$. We find an increase in the star formation rate with redshift and radio luminosity, although we are currently unable to identify any specific influence from the radio-loud AGN activity.

5.2 Introduction

Active galactic nuclei (AGN) are thought to play a vital role in galaxy evolution processes. Feedback from AGN is required to reconcile observed and simulated properties of the overall galaxy population, such as the cutoff at the bright end of the galaxy luminosity function (e.g. Croton et al., 2006; Bower et al., 2006). Observational evidence also suggests that AGN can strongly influence their environments. For example, the mass of the central supermassive black hole and the mass of the host galaxy bulge are tightly correlated (e.g. Magorrian et al., 1998; Marconi & Hunt, 2003; Häring & Rix, 2004), suggesting a co-dependence in their growth. Furthermore, the cosmic evolution of the luminosity density of quasars seems to track that of the star formation rate, with both peaking around $z \approx 2$ (e.g. Boyle & Terlevich, 1998; Wall et al., 2005; Hopkins & Beacom, 2006; Aird et al., 2010; Behroozi et al., 2013). Therefore, AGN activity must somehow be linked to the build-up of stellar mass in their host galaxies.

This may be particularly true in the case of radio-loud AGN, where the powerful radio jets can couple to the gas, causing dramatic feedback effects. These include the production of X-ray cavities caused by work done on the hot phase (e.g. Fabian et al., 2006; Forman et al., 2007) and appreciable outflows of both the hot and cool phases (e.g. Morganti et al., 2007; Nesvadba et al., 2008; Hardcastle et al., 2012). The heating and stripping of cool gas removes the fuel required for ongoing star formation in the galaxy host. Conversely, the radio jet may compress gas clouds and
therefore be associated with increased star formation in some circumstances (e.g. van Breugel et al., 1985; Fragile et al., 2004; Croft et al., 2006; Kalfountzou et al., 2012; Silk, 2013). Additionally, both star formation and AGN activity may be triggered by merger events (e.g. Granato et al., 2004; Di Matteo et al., 2005; Mihos & Hernquist, 1996), although the peak of each phenomenon may occur on different time scales (e.g. Wild et al., 2010).

Although the complex relationship between AGN activity and star formation is clearly fundamental to our understanding of galaxy evolution, it has not been well studied until recently. A major complication has been contamination by the AGN itself at the ultraviolet (UV), optical and mid-infrared wavelengths used to trace the star formation (e.g. Tadhunter et al., 1996; Haas et al., 2004; Holt et al., 2007; Dicken et al., 2009; Hardcastle et al., 2009). A number of studies have successfully avoided or minimised contamination, such as Baldi & Capetti (2008) using Hubble Space Telescope optical and UV imaging, Herbert et al. (2010) using optical spectroscopy, and Dicken et al. (2012) using Spitzer spectroscopy of polycyclic aromatic hydrocarbon lines. Baldi & Capetti (2008) and Herbert et al. (2010) both found evidence for enhanced star formation in high excitation radio galaxies (HERGs) but not in low excitation radio galaxies (LERGs). This is consistent with the emerging picture whereby HERGs (distinguished by their strong optical emission lines) are fed by cool gas made available via mergers or galaxy interactions, while LERGs (displaying weak optical lines) experience radiatively-efficient accretion of hot gas from the intergalactic medium surrounding massive galaxies, groups and clusters (e.g. Hardcastle et al., 2007; Best & Heckman, 2012). The small fraction of radio galaxies displaying star formation in the work of Dicken et al. (2012) seems to argue against the existence of a direct relationship between star formation and radio-loud AGN activity. However, these studies were restricted to small samples and/or small redshift ranges and so interpretation of their results within the context of the overall radio galaxy population is difficult.

Fortunately, many of these barriers can now be overcome using far-infrared (FIR)/submillimetre data from the sensitive Herschel Space Observatory (Pilbratt et al.,
2010) launched in 2009. Dust in galaxies is heated by young stars and re-radiates at wavelengths detectable by *Herschel*. These long wavelengths are ideal for tracing star formation since they are unattenuated by dust and gas and should not experience much contamination by thermal emission from the dusty torus associated with the AGN (which is expected to peak in the mid-infrared; Haas et al., 2004). Ground-based instruments such as the Submillimetre Common User Bolometer Array have been used to detect redshifted FIR emission from distant galaxies, but these observations were restricted to small samples of very powerful AGN (e.g. Archibald et al., 2001; Reuland et al., 2004).

The *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS) (Eales et al., 2010) provides wide-field, sensitive data at 100, 160, 250, 350 and 500 \(\mu\)m. This can be used to study dust emissions from large numbers of galaxies, ultimately over 550 deg\(^2\). There have been a number of successful studies of star formation in radio galaxies using these data. The first of these was conducted by Hardcastle et al. (2010; hereafter H10). They used the first H-ATLAS data release, over an equatorial \(\sim 14.4\) deg\(^2\) science demonstration phase (SDP) field, to study the far-infrared properties of 187 radio-loud AGN. Using a stacking analysis, H10 found a continued rise in the average FIR luminosity and, by proxy, the star formation rate with both redshift and radio power. Although, the trends were no different to those observed in a radio-quiet comparison population. However, the particular optical and near-infrared sample selection of this study restricted the analysis to fairly low-redshift (\(z \lesssim 0.8\)) and low-luminosity (and predominantly low-excitation) radio galaxies.

This work was then extended to a much larger sample by Virdee et al. (2013; hereafter V13). These authors identified 1,599 radio galaxies within \(\sim 128.8\) deg\(^2\) of the H-ATLAS Phase 1 field, thus improving on the previous sample size by almost an order of magnitude. This allowed them to examine the star formation trends of the radio galaxy hosts within subsets of stellar mass (and other various properties) while maintaining statistically-significant stacked detections. Their primary result was that high stellar mass radio galaxies appear to contain a deficit of dust with respect to a
comparative radio-quiet population. This agrees with the current view that LERGs are mostly hosted by massive elliptical galaxies with little dust content. However, their use of LAS and SDSS data again restricted their study to $z \lesssim 0.8$ objects.

In a companion paper, Hardcastle et al. (2013; hereafter H13) explicitly classified 1,342 radio galaxies in the Phase 1 field as low or high excitation using their emission line strengths in Galaxy and Mass Assembly (GAMA) optical spectra. By stacking 250 $\mu$m H-ATLAS data for these two populations separately, they found a factor of 3-4 excess in the star formation rate of HERGs with respect to the LERGs. Once again, this agrees with the picture in which the environments of HERGs are gas-rich merging or interacting systems which are likely to have more star formation than the quiescent environments of LERGs.

In this work, we conduct a similar far-infrared stacking analysis to H10, V13 and H13; however, more sensitive near-infrared and optical data is now available in the field and so we are able to probe the star forming properties of a significantly higher redshift population of radio-galaxies, i.e. our sample extends to optically fainter objects. We use the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING; Edge et al., 2013), which is 1.2 magnitudes more sensitive than the United Kingdom Infra-Red Telescope Infrared Deep Sky Survey - Large Area Survey (UKIDSS LAS; hereafter LAS; Lawrence et al., 2007) used in the previous studies. We also use new Isaac Newton Telescope optical imaging which extends $\sim 1$ magnitude deeper than the Sloan Digital Sky Survey (SDSS; Abazajian et al., 2009). Table 5.1 compares the basic parameters of this study with those of H10, V13 and H13.

We describe our data and define our samples for subsequent study in section 5.3 of this paper. We examine trends in the far-infrared luminosity and star-formation rates of our sample in section 5.4. We present our discussion and conclusion in sections 5.5 and 5.6, respectively. We have assumed a $\Lambda$ cold dark matter cosmology with $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_\Lambda = 0.7$ and $\Omega_M = 0.3$. All magnitudes are quoted using the Vega system, unless otherwise stated.
### Table 5.1: Comparison of sample parameters from the recent analyses of star-formation rates in radio galaxies with H-ATLAS data: Hardcastle et al. (2010) (H10), Virdee et al. (2013) (V13), Hardcastle et al. (2013) (H13) and this work.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>H10</th>
<th>V13</th>
<th>H13</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field size (deg$^2$)</td>
<td>14.4</td>
<td>128.8</td>
<td>161</td>
<td>14.4</td>
</tr>
<tr>
<td>No. radio galaxies</td>
<td>187</td>
<td>1,599</td>
<td>1,342</td>
<td>250</td>
</tr>
<tr>
<td>$z$ range</td>
<td>$&lt; 0.85$</td>
<td>$&lt; 0.8$</td>
<td>$&lt; 1.0$</td>
<td>$&lt; 5.0$</td>
</tr>
<tr>
<td>1.4 GHz flux range (mJy)</td>
<td>$&gt; 2.5$</td>
<td>$&gt; 2.25$</td>
<td>$&gt; 0.5$</td>
<td>$&gt; 10$</td>
</tr>
<tr>
<td>1.4 GHz luminosity (W Hz$^{-1}$)</td>
<td>$10^{21}$-$10^{27}$</td>
<td>$10^{22}$-$10^{27}$</td>
<td>$10^{22}$-$10^{27}$</td>
<td>$10^{23}$-$10^{28}$</td>
</tr>
<tr>
<td>NIR data</td>
<td>LAS</td>
<td>LAS</td>
<td>-</td>
<td>VIKING</td>
</tr>
<tr>
<td>NIR data limit (Vega)</td>
<td>$K &lt; 18$</td>
<td>$K &lt; 18$</td>
<td>-</td>
<td>$K &lt; 19.2$</td>
</tr>
<tr>
<td>Optical photometry</td>
<td>SDSS</td>
<td>SDSS</td>
<td>SDSS</td>
<td>INT</td>
</tr>
<tr>
<td>Optical mag limit (AB)</td>
<td>$r' &lt; 22$</td>
<td>$r' &lt; 22$</td>
<td>$i' &lt; 20.5$</td>
<td>$r &lt; 23$</td>
</tr>
<tr>
<td>Comparison populations</td>
<td>radio-quiet galaxies</td>
<td>radio-quiet galaxies</td>
<td>radio-quiet galaxies</td>
<td>V13</td>
</tr>
</tbody>
</table>
5.3 The data

5.3.1 1.4 GHz radio data

Our study focuses on the 14.38 deg$^2$ of the H-ATLAS science demonstration phase (SDP) field. This field is centred on right ascension (RA) 09$^h$05$^m$30$^s$, declination (Dec) 00$^\circ$30$'$00$''$. We created a catalogue of radio-loud galaxies by selecting all sources within this field detected at 1.4 GHz by the NRAO VLA Sky Survey (NVSS). NVSS surveyed the entire sky north of Dec -40$^\circ$ to a 50 per cent completeness limit of $\sim$2.5 mJy (Condon et al., 1998). As we are only interested in radio-loud AGN we can afford to adopt a conservative flux density cut to the sample. We therefore select only sources with $S_{1.4} > 10$ mJy.

We then found counterparts for these sources in the higher-resolution Faint Images of the Radio Sky at Twenty-centimetres (FIRST) survey (Becker et al., 1995) which provides better positional information. FIRST is also a 1.4 GHz survey conducted with the VLA but using a more extended configuration. Almost all our NVSS sources are detected in FIRST, except one diffuse object which is resolved out.

The process of source matching between NVSS and FIRST involved making some decision as to the morphology of the radio sources. For example, some NVSS sources were associated with two or more FIRST counterparts. Examples of this are shown in Figure 5.1. Based on the morphology of these multiple counterparts, we made a decision as to whether these were independent sources, simply unresolved by NVSS, or whether they were different components of the same source e.g. lobes or hotspots associated with the same object. In the former case, we scaled the NVSS flux density by the ratio of the individual to summed FIRST flux densities. We adopted these scaled flux densities for all further analysis. Our final catalogue contains 257 individual radio sources.
Figure 5.1: Examples of radio galaxy structure with NVSS and FIRST 1.4 GHz emission shown as black and magenta contours, respectively. NVSS contours are shown at 1, 3, 5, 7, 10 and 20 mJy. FIRST contours are shown at 0.3, 0.5, 0.7, 1, 2, 3, 5 and 7 mJy. The background grayscale image is the VIKING K-band map. VIKING catalogue sources are indicated by green circles. The central positions of all objects in the FIRST and NVSS source catalogues are indicated by blue crosses and the red pluses, respectively. In (a) the chosen VIKING counterpart to the radio source is indicated by the yellow diamond. In (b) no NIR counterpart is identified and the catalogued NVSS position is considered the best source position. In (c) the source position has been manually defined as the weighted mean between the two lobes, as shown by the orange square. The radio source towards the bottom of this image is considered to be an entirely separate object.
5.3.2 Near-infrared counterparts

To identify any near-infrared (NIR) counterparts to our radio sources, we examined VIKING images. The VIKING survey covers the full H-ATLAS SDP field in the $Z$, $Y$, $J$, $H$ and $K_s$ bands using the VISTA 4m telescope. The $5\sigma$ magnitude limit is $K_s \approx 19.2$ in the Vega system (Fleuren et al., 2012). It extends 1.2 magnitudes fainter than the UKIDSS LAS used in the H10 and V13 analyses. VIKING images and source catalogues used in this paper were accessed from the 2011 April 14 internal data release on the VISTA Science Archive (similar to the public VIKING DR2). The VISTA Data Flow System pipeline processing and science archive are described in Irwin et al. (2004) and Hambly et al. (2008).

Emission in the NIR is mostly dominated by the old stellar component (provided the object is not a quasar). Therefore, a VIKING counterpart to a radio source is very likely to correspond to the host galaxy. Due to the relatively low resolution of the radio data and the often complex radio emission structure, we identified optical counterparts manually. We extracted $3 \times 3$ arcmin$^2$ cutouts of VIKING images in all bands centred on the radio source positions and overlaid FIRST and NVSS contours, as in Figure 5.1. We visually inspected each cutout to ensure that the correct VIKING source was chosen as the most likely host galaxy candidate. We note that in some cases the choice of the best matching VIKING source was highly subjective, as in Figure 5.1a.

We find unique VIKING counterparts to 215 radio sources (84 per cent of our sample). 212 of these appear in the VIKING source catalogue and we use these catalogued positions in the further analysis. We manually recorded the positions of the three additional sources which appeared to have faint (low signal-to-noise) VIKING counterparts but were not in the source catalogue.

For the remainder of objects, either no detection was evident in the VIKING data or the radio emission was too complicated to determine where the host galaxy should lie (four cases). For such sources, the most accurate available position was provided by
the FIRST catalogue in 28 cases and by NVSS in 8 cases. The latter was appropriate when, for example, two or more FIRST sources were deemed to originate from one physical object and no VIKING counterpart was available, as in Figure 5.1b. A further six positions had to be defined manually where, for example, two jets or lobes were apparent at 1.4 GHz but not the core (as in Figure 5.1c). In such instances, the central position was chosen to be the flux-weighted distance between the lobes.

5.3.3 Optical counterparts

In order to derive photometric redshifts (section 5.3.4), we collected additional photometry in the optical $g$, $r$ and $i$ bands provided by the 2.5 m Isaac Newton Telescope (INT). These observations were taken as part of a survey which aims to provide deep, optical data over the full H-ATLAS Phase 1 fields using the Wide Field Camera (WFC). Observations were conducted on 2010 March 10-19 and 2011 March 1-9. Integration times were 300 s in $g$ and $r$ band and 600 s in $i$ band. The magnitude limit of the data is approximately $r_{AB} < 23$ for a 7 arcsec aperture. The data reduction processes are identical to those described by González-Solares et al. (2011) for WFC observations of the ELAIS, First Look Survey and Lockman Hole fields. The optical data do not cover the full field, however we find that 216 sources in our sample (84 per cent) are covered in all three bands, and a further 4 in at least one band. We found 117 sources above the $5\sigma$ level in at least one INT band. For those sources without VIKING counterparts, we verified that no counterpart could be identified in INT bands either.

5.3.4 Obtaining redshift information

5.3.4.1 Spectroscopic redshifts

To calculate luminosities, we require redshift information for every source in our sample. We firstly collected any available spectroscopic redshifts by conducting a
Figure 5.2: The distribution of all spectroscopic redshifts available (magenta), all generated photometric redshifts (cyan) and the final redshift distribution adopted for this analysis (black). For the latter, photometric redshifts were replaced by spectroscopic redshifts where available.

1 arcsec position cross-match between the VIKING position and the Galaxy and Mass Assembly (GAMA) redshift catalogue using TOPCAT. The GAMA redshift survey is on-going and utilises the AAOmega spectrograph on the Anglo-Australian Telescope, Australia (Driver et al., 2009, 2011). GAMA data used in this analysis cover the full H-ATLAS SDP field and are complete down to magnitude $r < 19.4$. 59 objects in our sample were associated with high quality GAMA redshifts. Of these objects, one has a very low spectroscopic redshift of $\sim 5 \times 10^{-5}$ and we suspect it is a star. We therefore exclude it from further analysis. The spectroscopic redshift distribution of the remaining objects is shown in Figure 5.2.

5.3.4.2 Photometric redshifts

We derived photometric redshifts for all sources in our sample. The photometry at the most accurate available source position (see section 5.3.2) was extracted from all VIKING and INT bands using the PHOT task in the IRAF image analysis package. For VIKING bands, we used an aperture diameter of 5.7 arcsec to match that of the
VIKING source catalogue for comparison. We verified that our measured magnitudes agreed closely with those in the VIKING catalogue. However, we use our photometrically measured photometry for all cases in the further analysis. We used a 7 arcsec aperture for measuring INT photometry since these data have poorer seeing. The slightly larger aperture ensures that we measure the total flux of the radio galaxy host.

We identified 37 blended sources in the VIKING images and 18 in the INT images. For these, we extracted the photometry using a smaller, 2 arcsec aperture and scaled up by an appropriate correction factor. We took this factor to be the average ratio of all measured fluxes using the larger aperture to that found using the 2 arcsec aperture. For most bands, this correction was close to 1 magnitude.

We used LEPHARE (Arnouts et al., 1999; Ilbert et al., 2006) to determine the best fit photometric redshift for all sources. LEPHARE generates photometric redshifts by attempting to fit the photometry of each object with input spectral models. For this analysis, we used a set of 30 different galaxy SED templates from Salvato et al. (2009). This suite of models contains various starburst, hybrid AGN/star-forming and quasar templates. Initially a standard elliptical model was also included, but was found not to be the best fit to any source and so was subsequently removed. We restricted luminosities to $-28 < M_K < -20$ and redshifts to $z < 5$, choosing these values to exclude as few correct fits as possible while eliminating obvious bad fits.

One source was undetected in all optical and NIR bands, and therefore it was not possible to obtain a photometric redshift. Since no spectroscopic redshift was available either, we had no choice but to discard it from the rest of the analysis. Five sources have photometric redshifts of exactly 5.0, and do not have spectroscopic redshifts. Since these are clearly hitting the priors of the photometric redshift code, we do not believe them and discard them from the rest of the analysis.

Our final sample contains 250 objects. The distribution of photometric redshifts for all 250 sources is shown in Figure 5.2. Spectroscopic redshifts were used in preference to photometric redshifts for the rest of the analysis for the 59 sources where they
were available. The final redshift distribution adopted for this analysis is also shown in Figure 5.2.

Using these, we converted the 1.4GHz NVSS flux density \( S_{1.4} \) (units W m\(^{-2}\) Hz\(^{-1}\)) of each source into a luminosity \( L_{1.4} \) (units W Hz\(^{-1}\)), adopting a spectral index of \( \alpha = 0.8 \). H13 confirmed this spectral index to be reasonable by cross-matching their sample with 325MHz data from the GMRT (Mauch et al., 2013).

### 5.3.4.3 Photometric redshift quality

Since we must rely on photometric redshifts for the majority of our sample, we would like to test the accuracy of these redshifts. It is not clear what the best diagnostic of photometric redshift quality is. LEPHARE can only examine a finite set of input spectral templates and therefore the redshift errors and \( \chi^2 \) values it outputs may not necessarily indicate whether or not the true spectrum has been tested. Nonetheless, Figure 5.3 compares the \( \chi^2 \) of the fitted template and the fractional redshift errors \( \left( z_{\text{error}}/z \right) \) to provide some characterisation of redshift quality. We flag the 19 photometric redshifts with \( \left( z_{\text{error}}/z \right) > 1 \) or \( \chi^2 > 200 \) as unreliable. We do not flag objects for which more accurate spectroscopic redshifts are available. High \( \chi^2 \) values are likely caused by underestimated photometric noise (e.g. where a source falls on a particularly bad part of the image) or because the source itself is variable and measurements are not simultaneous (e.g for quasars).

Figure 5.4 compares the spectroscopic redshifts of the 59 sources detected in GAMA with the photometric redshifts generated with LEPHARE. There is an encouragingly good correlation at all redshifts \( (r = 0.79) \), even for most known quasars (see section 5.3.6 below). This is likely due to the inclusion of QSO and QSO hybrid templates during LEPHARE fitting. We note that some photometric redshifts are large overestimates, even below \( z = 1 \). In the more extreme cases, this is due to the QSO template having the lowest \( \chi^2 \) even though the object is not a quasar. Fortunately, we can use the spectroscopic redshifts for these objects.
Figure 5.3: The $\chi^2$ of the photometric template fit versus the fractional error in the photometric redshift. Only sources bounded by the two dotted lines are considered to have reliable photometric redshifts. Red points indicate sources with available spectroscopic redshifts. Objects with unreliable photometric redshifts and no available spectroscopic redshifts are marked with green circles.

As a further reality check of our photometric redshifts, we can examine how well these reproduce the $K$-$z$ relation observed by Willott et al. (2003) for powerful radio galaxies, where:

$$K = 17.37 + 4.53 \log z - 0.31 (\log z)^2. \quad (5.1)$$

Figure 5.5 shows the photometric redshift versus the $K$-band apparent magnitudes for the 237 sources in our sample where these magnitudes are available. Most photometric redshifts lie close to the $K$-$z$ relation, although the scatter is $\sim 1.2$ mag as opposed to the $\sim 0.6$ mag quoted by Willott et al. (2003). We address the impact of this on our results in section 5.4.4. We note that much of the downwards scatter is caused by quasars in the sample which are not expected to follow equation 5.1 due to contamination by non-thermal emission.
5.3. THE DATA

CHAPTER 5. RADIO GALAXIES IN H-ATLAS

Figure 5.4: Comparison between GAMA spectroscopic redshifts and photometric redshifts generated with LEPHARE. Black diamonds indicate spectroscopically confirmed quasars. Perfect correlation is indicated by the dashed line.

Figure 5.5: The photometric (red) and available spectroscopic (black) redshifts versus $K$-band magnitude. The dashed line shows the $K$-$z$ relation of Willott et al. (2003). Diamonds indicate spectroscopically confirmed quasars and squares show quasars identified via their colours. Green circles show objects with unreliable photometric redshifts.
5.3.5 Far-infrared counterparts and photometry

Finally, we obtained H-ATLAS photometry for all objects in our catalogue. The SPIRE instrument provides 250, 350 and 500 $\mu$m data across the SDP field (Griffin et al., 2010). Maps in each of these bands have been constructed and are described in Pascale et al. (2011). Data at 100 and 160 $\mu$m are provided by the PACS instrument (Poglitsch et al., 2010) with maps constructed and described in Ibar et al. (2010). Rigby et al. (2011) created a catalogue of sources in the SDP detected with $> 5\sigma$ significance in background-subtracted point spread function (PSF) convolved SPIRE maps. The most sensitive data is available at 250 $\mu$m, where $5\sigma = 33.5$ mJy beam$^{-1}$ (Rigby et al., 2011).

Fleuren et al. (2012) used a maximum likelihood approach to identify VIKING counterparts to objects in the H-ATLAS catalogue. 14 objects in our catalogue are associated such a VIKING counterpart and therefore have direct detections in H-ATLAS. We also conducted a 5 arcsec position cross-match between the H-ATLAS source catalogue and our sample using TOPCAT (Taylor, 2005). In addition to those identified by Fleuren et al. (2012), we found one further match with an object faintly detected in VIKING bands but not formally in the catalogue. For these 15 objects, we use all SPIRE and PACS fluxes as recorded by Rigby et al. (2011) since this catalogue properly deals with extended sources. For objects not in the Rigby et al. (2011) catalogue (i.e. not detected above the $5\sigma$ limit), we extract H-ATLAS photometry following the method detailed by H10, V13 and H13 and briefly summarised here. We assume that the sources are unresolved in SPIRE; the 250 $\mu$m data has an 18 arcsec beam and our sources are expected to be smaller than the point-spread function (PSF). We estimate SPIRE fluxes by taking the value of the pixel at the source position in the background-subtracted PSF-convolved H-ATLAS Phase-1 maps, which is the maximum-likelihood estimator of the true flux. We note that the large PSF of this map may mean that we are measuring additional contributions from the nearby environments of each galaxy. This contribution might be significant if the galaxy is part of a merging system. Conversely, we may be slightly under-estimating the total flux.
for any sources larger than the PSF; however it is unlikely that many physically large
but undetected sources exist. We extracted photometry from PACS bands using a
more traditional aperture integration technique since sources are more likely to be
resolved in the higher-resolution PACS maps.

We determine whether each source is individually detected with greater than $2 \sigma$
significance. The noise is not strictly Gaussian due to a high level of confusion in the
SPIRE maps and so the $2 \sigma$ level cannot be calculated by simply doubling the value of
the corresponding pixel in the noise map (Pascale et al., 2011). Rather, we measured
the flux in a large number of apertures randomly placed down on the SPIRE maps
and took the $2 \sigma$ level to be the value below which 97.7 per cent of randomly measured
fluxes sit. At 250 $\mu$m, $2 \sigma = 24.6$ mJy. In addition to those in the Rigby et al. (2011)
catalogue, 55 of our sources are directly detected at $> 2 \sigma$ in the 250 $\mu$m map.

5.3.6 Quasar identification

The FIR fluxes we measure for all sources could of course include emission from AGN
activity as well as the emission from the host galaxy itself. For non-quasars this is
not much of a problem as there will be little or no contribution from the dusty torus
which peaks in the mid-IR (Hardcastle et al., 2009; Dicken et al., 2009).

However, the jet of a radio-loud quasar makes a small angle to our line-of-sight and
thus it is more likely that non-thermal emission could be contaminating our meas-
urements. Therefore, we wish to identify any quasars in our sample and carefully
consider the impact of non-thermal contamination on our final results.

Visually inspecting the 59 available GAMA spectra, we identified 11 quasars via their
broad optical emission lines and strong blue continuum. These confirmed quasars
are marked with diamonds in Figures 5.4 and 5.5. All sources with GAMA redshifts
above $z = 0.8$ are quasars. This is unsurprising since objects at high redshift must
be extremely luminous to be seen above the GAMA magnitude limits.
To identify quasars where no GAMA spectrum is available, we use the near-infrared colour-colour diagram in Figure 5.6. Maddox et al. (2012) showed that quasars, as well as having a $K$-band excess with respect to stars, also have a $J$-band excess with respect to galaxies (although dust reddening should move quasars downwards and to the right). We note that this diagram only shows the 166 sources with fluxes in $g$, $J$ and $K$ bands (66 per cent of the full sample). Most spectroscopically identified quasars have $(g - J) < 2$, in agreement with Maddox et al. (2012). We therefore flag all objects in this region as suspected quasars. The 10 additional objects in this population without available GAMA spectra are indicated by the black squares in Figure 5.5.

Quasars usually appear as point sources (i.e. star-like). To check the validity of our quasar selection criterion, we use the pstar (probability that a source is a star) information in the VIKING source catalogue to identify point-like sources. We note, however, that a point-like source may not necessarily be a quasar but simply a galaxy at the limits of the data resolution. The 17 VIKING sources appearing in Figure 5.6 with pstar>0.5 are marked with black diamonds. Orange squares show the 133 VIKING-detected sources with pstar<0.5. From this we can see that the majority of point-like sources lie at $(g - J) < 2$, which partially justifies our cut. We have gained no further information on the 16 sources that are not detected in VIKING. Many objects in the top-right part of Figure 5.6 fall in to this category and so we have no indication of whether or not they are quasars.

It is worth noting that of the 21 objects we have identified as quasars, LEPHARE found that a QSO template was the best fit to the photometry of 17 of them.

We estimated the potential quasar ‘contamination’ at 250 $\mu$m for all objects in our sample by assuming an extreme spectral index of $\alpha = 0.5$ and extrapolating to 250 $\mu$m based on the known 1.4 GHz flux. This upper limit on the extent of contamination is shown in Figure 5.7 in comparison to the measured H-ATLAS 250 $\mu$m flux (for sources with $> 2\sigma$ H-ATLAS detections) or the $2\sigma$ upper limit on the 250 $\mu$m flux (for H-ATLAS non-detections). We find that there is no correlation between possible
Figure 5.6: Near-infrared colours of the sample of radio galaxies, as a diagnostic of quasar presence. Quasar and non-quasar classifications identified through GAMA spectral features are indicated by cyan squares and magenta dots, respectively. Objects without GAMA spectra are represented by grey dots. VIKING point-like and non-point-like classifications are indicated by the black diamonds and orange squares, respectively. Error bars are omitted for clarity. The median $J - K$ errors on grey, magenta and cyan points are 0.18, 0.01 and 0.04 magnitudes, respectively. The corresponding median $g - J$ errors are 1.48, 0.10 and 0.02 magnitudes, respectively. All objects above the dashed line are considered to be quasars.
quasar contamination and the actual measured H-ATLAS 250 µm flux. Only one object with a strong 250 µm detection could be suffering appreciable contamination, but we have already identified this object as a quasar and have dealt with it cautiously in the analysis.

5.4 Analysis

5.4.1 FIR luminosity and $K$-correction

Of the H-ATLAS maps available, the 250 µm band is the most sensitive and has the smallest beam size of the SPIRE bands, thus minimising source confusion. Therefore, we choose to focus on the 250 µm properties of our sample in the following analysis. To convert the 250 µm flux density ($S_{250}$) into luminosity ($L_{250}$), we require both the redshift and the $K$-correction ($K$) (e.g. Dunne et al. 2011):
\[ L_{250} = \frac{4\pi D_L^2 S_{250} K}{1 + z}, \quad (5.2) \]

where \( D_L \) is the luminosity distance. Since the large majority of our sample are undetected in one or more SPIRE or PACS bands, determination of the appropriate \( K \)-correction represents the greatest difficulty in a stacking analysis of this nature.

Following Dye et al. (2010), we assume a modified blackbody SED:

\[ S_\nu = \frac{N_\nu \nu^{\beta+3}}{e^{h \nu/kT} - 1}, \quad (5.3) \]

where \( \nu \) is frequency and \( \beta \) is the spectral index of the power-law emissivity of dust.

Hardcastle et al. (2013) found an optimal value of \( \beta = 1.8 \) for their sources (see also Smith et al., 2013). We will assume the same value as it should also be valid for our sources, at least at low redshift. The two remaining free parameters are the temperature (\( T \)) and the normalisation (\( N \)). The most straightforward approach is to assume an isothermal model for all sources in our sample, as in H13 and V13. H13 determined a best-fitting value of \( T = 20 \text{K} \) for their sample, based on sources with significant detections in more than one Herschel band. We adopt this value for all objects in our sample and fit the SED allowing the normalisation to vary. However, we note that it is not immediately obvious whether this value of \( T \) holds for our sample, which contains sources with higher radio luminosity and higher redshifts than in H13. Indeed, in section 5.4.3 we will investigate how \( T \) varies between the bins.

For the rest of the analysis, we will concentrate on the monochromatic infrared luminosity at \( 250 \mu\text{m} \) (\( L_{250} \)) rather than the bolometric infrared luminosity (\( L_{\text{IR}} \)). The advantage of this is that only the potentially incorrect \( K \)-correction can influence our results. If we were to use \( L_{\text{IR}} \), the choice of an incorrect SED template could also skew our results significantly.

H13 determined a direct relationship between \( L_{250} \) and SFR by comparing their calculated luminosities to star-formation rates derived via the method of Brinchmann
et al. (2004) using SDSS information:

\[
\log(L_{250}/\text{W Hz}^{-1}) = 23.64 + 0.96 \log(\text{SFR}/\text{M}_\odot \text{ yr}^{-1}).
\] (5.4)

We adopt the same relationship to calculate the SFR of each source in our sample. However, we note that the predicted SFR will be an upper limit since this relation attributes all 250 $\mu$m emission to hot dust and assumes there is no significant contribution from the cold dust component. See Smith et al. (2012) for a discussion of the separation of these two components for H-ATLAS sources.

### 5.4.2 Far-infrared stacking

Many radio galaxies in our sample are undetected at 250 $\mu$m. However, we can still recover some information on the infrared properties of this population via stacking. This involves finding the weighted average 250 $\mu$m flux ($S_{250}$) of the $n$ sources of interest:

\[
\langle S_{250} \rangle = \frac{\sum_{i=1}^{n} (w_i S_{250,i})}{\sum_{i=1}^{n} w_i},
\] (5.5)

We use a weighting factor, $w_i = 1/\sigma_{250,i}^2$, where $\sigma_{250,i}$ is the value of the 250 $\mu$m noise map at the $i^{th}$ source position. Stacking the 250 $\mu$m flux of all sources in the catalogue and repeating this process for all pixels within $\sim$1.2 arcmin of the central source position results in Figure 5.8 (left panel) and allows us to visually verify that we achieve a stacked detection. Since a number of sources are individually detected at 250 $\mu$m, it is unsurprising that this gives a strong stacked detection with $\langle S_{250} \rangle = 19.13 \pm 0.41$ mJy. Stacking only the 239 objects in our sample which are not formally detected in the H-ATLAS source catalogue, i.e. with $S_{250} < 5 \sigma$, we still find a very
significant detection with $\langle S_{250} \rangle = 6.68 \pm 0.42$ mJy. The corresponding stacked image is shown in the middle panel of Figure 5.8.

Since there is a high level of confusion in the 250 $\mu$m map, the noise is not Gaussian. Therefore, to determine whether the stacked signal is significantly detected above the noise, we use a Kolmogorov-Smirnov (K-S) test. As described in detail by H10, we create a ‘background’ population by measuring the flux densities at 100,000 random positions in the field. The stacking tests shown in the left and middle panels are both distinguished from the background population with $>99.99$ per cent confidence with a K-S test. For comparison, the right panel of Figure 5.8 shows an example of stacking at 239 random positions within the SDP field. Of course, this is indistinguishable from the background with a K-S test.

We now apply this stacking method to examine how the average $L_{250}$ of the population varies with $L_{1.4}$ and redshift. We split the sample into four $L_{1.4}$ or $z$ bins such that each subset contains roughly equal numbers of sources. The stacked 250 $\mu$m flux map in each bin is shown in Figure 5.9. Convincing detections are evident in all cases. We then find the stacked $L_{250}$ for all sources in each bin via equation 5.5, replacing $S_{250}$ by $L_{250}$. We again used a K-S test to verify that we achieve a statistically significant detection in each bin. As in H13 and V13, errors on the stacked luminosities are calculated using bootstrapping with replacement in each bin. The stacked results are shown in Table 5.2 and marked in the top panel of Figure 5.10. Individual points are also shown for illustrative purposes, colour coded by their detection level at 250 $\mu$m. Quasars are marked on this plot but have been excluded from the stacked calculations.

Some of the radio sources in our catalogue may be (non-active) starbursts with radio fluxes high enough to enter our sample. However, we are only interested in studying galaxies displaying AGN activity. Therefore, we must identify and exclude such starburst galaxies from the analysis. Figure 5.11 shows the $L_{1.4}$ versus the redshift of each individual object in the sample. The lower horizontal line shows where the starburst luminosity function cuts off sharply at $5 \times 10^{23}$ W Hz$^{-1}$, as determined by Sadler et al. (2002) and Mauch & Sadler (2007). The dashed line indicates the 10 mJy flux cut
Figure 5.8: The stacked 250 µm SPIRE images of all 250 radio sources (left), only the 239 sources not in the H-ATLAS catalogue (centre) and 239 random positions (right). Note that in stacking these images, each pixel of a given image was weighted by the value of the noise map at the position of the central pixel. The small green cross marks the relative position from the central pixel. The colour scale shows the flux in Jy.
Figure 5.8: The stacked 250 µm SPIRE images within bins of 1.4 GHz luminosity (top) and redshift (bottom). Labels and units are as in Figure 5.8.
Figure 5.10: $L_{250}$ versus 1.4-GHz radio luminosity (left) and redshift (right). Individual points are calculated using the isothermal model (top row), and temperature fitting within the bins (bottom row). Red crosses show sources detected above the 5σ level at 250µm, orange squares at 2-5σ and black points at < 2σ. The black points are plotted at the upper 2σ limit, as indicated by the arrows. The dashed lines represent the bounds of the chosen bin ranges. The bars show the weighted average within the bin calculated using the isothermal model (magenta) and temperature fitting within the bins (cyan). They are plotted at the average $L_{1.4}$ or redshift within the bin. Vertical error bars were calculated with a bootstrapping method and horizontal bars show the bin range. The dashed line shows the FIRC cut applied. Objects above this line and which are suspected to be starburst galaxies are marked with blue squares. Quasars are marked with green diamonds (spectroscopically-determined) and circles (colour-determined).
Figure 5.11: 1.4 GHz luminosity vs redshift. The legend indicates the detection level in the 250 µm H-ATLAS SPIRE map. Known and suspected quasars are marked by green diamonds and circles, respectively. The imposed 10 mJy cut is indicated by the dashed line. The lower horizontal line shows the luminosity above which starburst galaxies are expected to make a minimal contribution and the upper line indicates where HERGs will begin to dominate the AGN population.
imposed during source selection. It is clear that this cut precludes many starbursts from entering our sample. Nonetheless, to identify any starbursts remaining in the sample, we employ the FIR-radio correlation (FIRC). We adopt the notation of Jarvis et al. (2010) and define the FIRC by the quantity:

\[ q_{250} = \log\left(\frac{L_{250}}{L_{1.4}}\right) \]  

(5.6)

Jarvis et al. (2010) find that objects obeying the FIRC within one standard deviation sit at \(1.4 < q_{250} < 2.1\). As can be seen in Figure 5.10, three sources at low \(L_{1.4}\) seem distinctly separated from the rest of the population and are likely to be starbursts. We use a \(q_{250} > 1.1\) cut, indicated by the dashed line, to ensure that all three of these suspected starbursts are excluded from the stacking calculation.

The individual star-formation rates and \(L_{250}\) values shown in Figure 5.10 seem unrealistic, extending above the FIRC cut in the highest \(L_{1.4}\) bin (which are very unlikely to be starburst galaxies). This suggests either that there is strong AGN contamination or that the fixed-temperature method for \(K\)-correction breaks down in the higher bins.

5.4.3 Temperature fitting within bins

Since there is no physical reason to assume that a fixed temperature is suitable for all our sources, we tried a second method for calculating the \(K\)-correction by allowing the temperature to vary between the bins, again following the method of H13 and V13. For each \(T\) in the range 5–55 K, we fit the template in equation 5.3 to all sources in a particular bin and calculated the \(\chi^2\) of each fit. The temperature resulting in the minimum total \(\chi^2\) was deemed to be the best fitting single temperature that characterises sources in that bin. The \(L_{250}\) for each source was then re-calculated using this temperature for the \(K\)-correction. These individual values are shown in the bottom row of Figure 5.10 and the stacked measurements are indicated by the
Table 5.2: The mean 250 µm luminosities and star-formation rates derived using a fixed temperature within bins of 1.4 GHz luminosity and redshift.

<table>
<thead>
<tr>
<th>Range (168)</th>
<th>Log(SFR/10^{30} M_{\odot} yr^{-1})</th>
<th>Log(L_{250}/10^{25.7} {\rm W Hz^{-1}})</th>
<th>K-S probability (%)</th>
<th>No. stacked Mean S_{250}^{\odot} mJy</th>
<th>( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0 - 7.0</td>
<td>0.0 ± 0.2</td>
<td>0.8 ± 0.2</td>
<td>91</td>
<td>0.1 - 0.6</td>
<td>0.6 - 0.6</td>
</tr>
<tr>
<td>7.0 - 8.0</td>
<td>0.9 ± 0.2</td>
<td>0.7 ± 0.2</td>
<td>97</td>
<td>1.0 - 2.0</td>
<td>1.0 - 2.0</td>
</tr>
<tr>
<td>8.0 - 9.0</td>
<td>0.6 ± 0.2</td>
<td>0.8 ± 0.2</td>
<td>94</td>
<td>0.6 - 1.0</td>
<td>0.6 - 1.0</td>
</tr>
<tr>
<td>9.0 - 10.0</td>
<td>0.7 ± 0.2</td>
<td>0.8 ± 0.2</td>
<td>97</td>
<td>1.0 - 2.0</td>
<td>1.0 - 2.0</td>
</tr>
<tr>
<td>10.0 - 11.0</td>
<td>0.8 ± 0.2</td>
<td>0.8 ± 0.2</td>
<td>94</td>
<td>0.6 - 1.0</td>
<td>0.6 - 1.0</td>
</tr>
<tr>
<td>11.0 - 12.0</td>
<td>0.9 ± 0.2</td>
<td>0.7 ± 0.2</td>
<td>91</td>
<td>0.1 - 0.6</td>
<td>0.1 - 0.6</td>
</tr>
</tbody>
</table>

Note: No stacked Mean S_{250}^{\odot} mJy
cyan bars. We caution the reader that each source may now have two different \( L_{250} \) estimates based on the best-fit temperatures in the \( z \) bin and in the \( L_{1.4} \) bin. However, we can see that the predicted \( L_{250} \) and SFR now seem much more reasonable, even in the highest bins. All values are well below the FIRC, except for the three suspected starburst galaxies and the \( 2\sigma \) upper limit of a suspected quasar.

The final assumed temperature in each bin is shown in Table 5.3. Most best fitting temperatures are greater than 20 K, confirming that this fixed temperature was not the most appropriate choice for our sample. The highest best fit temperatures (\( >40 \) K) occur in the highest bin for both binning schemes. There is a strong correlation between \( z \) and \( L_{1.4} \) in our sample, with most sources in the highest \( z \) bin also in the highest \( L_{1.4} \) bin (see Figure 5.11). This implies that radio galaxies in our sample at \( z > 2 \) have higher dust temperatures than their low-redshift counterparts. We will discuss this apparent correlation in section 5.5.2. Considering this, it is not surprising that the fixed temperature model fails to predict reasonable \( L_{250} \) values for these sources. This also explains the significant difference between the stacked values for the fixed and best-fitting temperature methods in the highest bins. Based on these results, we believe that the best-fitting \( T \) method provides the more realistic \( K \)-corrections.

### 5.4.4 Examining possible biases

We now examine the impact of starburst galaxies, quasars and unreliable photometric redshifts on the stacked \( L_{250} \) measurements. In all cases we use the best-fitting \( T \) model to calculate \( K \)-corrections. It is ideal to remove any starburst galaxies from our sample to ensure that we are studying the targeted physics. However, the top panel of Figure 5.12 shows that any remaining starbursts in our sample should only influence the stacked value in the lowest \( z \) and \( L_{1.4} \) bin and will not significantly alter our science conclusions.

The middle panel of Figure 5.12 shows that our stacked measurements are not ro-
Table 5.3: The mean 250/µm luminosities and star-formation rates derived by temperature-fitting within bins of 1.4 GHz luminosity and redshift.

<table>
<thead>
<tr>
<th>Bin</th>
<th>Parameter Range</th>
<th>No. stacked Entries</th>
<th>Best-fit $T$ (K)</th>
<th>Red $\chi^2$</th>
<th>$\log\left(\frac{\text{SFR}}{M_{\odot}/y} \right)$</th>
<th>$\log_{10}(L_{250}/\mu\text{m})$</th>
<th>$\log_{10}(L_{1.4}\mu\text{m})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z = 0.0$</td>
<td>$20 - 2.6$</td>
<td>1</td>
<td>51</td>
<td>1.0</td>
<td>$31.1 \pm 1.4$</td>
<td>$25.0 \pm 1.4$</td>
<td>$27.0 \pm 1.4$</td>
</tr>
<tr>
<td>$z = 0.6 - 1.0$</td>
<td>$24.3 - 26.8$</td>
<td>2</td>
<td>56</td>
<td>0.8</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
</tr>
<tr>
<td>$z = 0.6 - 1.0$</td>
<td>$26.8 - 28.0$</td>
<td>3</td>
<td>55</td>
<td>0.8</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
</tr>
<tr>
<td>$z = 0.6 - 1.0$</td>
<td>$26.3 - 26.8$</td>
<td>4</td>
<td>54</td>
<td>0.8</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
</tr>
<tr>
<td>$z = 0.6 - 1.0$</td>
<td>$26.3 - 26.8$</td>
<td>5</td>
<td>53</td>
<td>0.8</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
</tr>
<tr>
<td>$z = 0.6 - 1.0$</td>
<td>$26.3 - 26.8$</td>
<td>6</td>
<td>52</td>
<td>0.8</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
<td>$24.3 \pm 1.1$</td>
</tr>
</tbody>
</table>
bust to the influence of synchrotron contamination from quasars. While we have identified 9 percent of our sample as quasars, we should be mindful of the impact of any remaining quasars. These are likely to exist since LEPHARE uses a best-fit QSO template for 76 sources (30 percent of the sample). We do note, however, that our results are robust to whether or not we use all 76 objects in our quasar exclusion. Willott et al. (2000) calculate a quasar fraction of 37 percent for a sample of radio galaxies at $z > 1$ and with comparable radio luminosity to our sources. This is again considerably higher than our 6 percent quasar fraction at the same redshifts so we expect to be contaminated by a number of unidentified quasars.

Finally, the bottom panel of Figure 5.12 shows that the inclusion of the sources with unreliable redshifts does not affect the overall statistics. This confirms that (potentially) large errors on relatively few (19) sources do not influence the average results.

We would like to further gauge the importance of redshift errors for all sources (not just the few very unreliable redshifts) on the stacked $L_{250}$ since the bootstrapping method does not account for this. To do so, we add scatter to the redshifts using $z_{\text{new}} = z + z_{\text{error}} \times G$ where $G$ is randomly sampled from a Gaussian distribution with a mean of 0 and a standard deviation of 1. We then bin and stack using these perturbed redshifts and repeat this 100 times. The 1σ dispersion in the stacked measurement indicates the uncertainty in the results due to redshift errors. As shown in Figure 5.13, we find that redshift error does not dominate the total error on the stacked measurements.

5.5 Discussion of results

5.5.1 Comparison of star formation rates

Here we compare our derived SFRs to other studies as a reality check. The two $S_{250} > 5\sigma$ sources in our sample with the highest redshifts also correspond to the two
Figure 5.12: Comparison between the stacked $L_{250}$ using various sample cuts in bins of $L_{1.4}$ (left) and $z$ (right) using temperature fitting to the bins.
Figure 5.13: A comparison of the stacked $L_{250}$ with (red) and without (black) the introduction of random redshift errors in bins of $L_{1.4}$ (left) and $z$ (right). FIR luminosities have been calculated using the $T$ fitting model and quasars and objects on the FIRC have been excluded. Black error bars were generated via bootstrapping. Red error bars show the 1σ standard deviation in the 100 stacks produced with the inclusion of random redshift error.

with the highest $L_{1.4}$. These have $L_{250}$ values of $1.74 \times 10^{26}$ and $3.55 \times 10^{26}$ W Hz$^{-1}$ corresponding to estimated star-formation rates of 511 and 1075 M⊙ yr$^{-1}$, respectively. If their redshifts are correct (and they are not quasars) then they must be extreme starbursts hosting AGN. These are similar to the SFRs found by Barthel et al. (2012) and Seymour et al. (2012) for individual high redshift AGN with starburst activity. Respectively, they report SFRs of up to 770 M⊙ yr$^{-1}$ at $z = 2.474$ and 1390 M⊙ yr$^{-1}$ at $z = 2.156$. We do, in fact, suspect that the lower redshift source is a quasar based on its NIR colours. We have no indication of whether or not the higher redshift (and higher SFR) object is a quasar since it does not have a GAMA spectrum nor a $J$-band magnitude.

The average star formation rates we measure are in broad agreement with those of Seymour et al. (2011). These authors examined radio galaxies over a similar redshift ($0.4 < z < 3$) and radio luminosity range ($25 < \log(L_{1.4}) < 26.5$) using far-infrared data from Spitzer and the Herschel Multi-tiered Extragalactic Survey (HERMES). Their results and ours are both also consistent with the measurements of H10 at $z \lesssim 0.8$. 

173
5.5. DISCUSSION

Figure 5.14: Comparison between the stacked $L_{250}$ in bins of $L_{1.4}$ (left) and $z$ (right) using the source samples of this work and that of V13 with different $K$-correction models, as indicated in the legend. Values are plotted at the mean x-axis value within each bin. Error bars are derived using the bootstrapping method.

Since both our low redshift average SFRs and individual, high redshift SFRs agree with other measurements in the literature, we have confidence that the average SFRs of our high redshift sample are also realistic.

5.5.2 Comparison to the low-redshift population

To properly interpret our results within the context of star formation in radio galaxies, we require a comparison population of radio-quiet sources. However, the catalogues presently available contain very few such sources at high redshift. Creating a sufficiently large redshift-matched comparison sample to facilitate stacked detections would require significantly deeper NIR data and is beyond the scope of this work. Therefore, we have a degeneracy in our results; we cannot ascertain whether the trends we observe are due to the high-redshift of the sources (i.e. age-related), or due to intrinsic properties relating to the physical processes causing the radio emission, such as feedback mechanisms of the AGN.

We can therefore only compare the properties of high-redshift radio sources to the lower-redshift population. In Figure 5.14 we compare our stacked results to those of V13 using our chosen binning scheme for both samples and using both $K$-correction...
models. We are clearly sampling out to much higher redshifts than V13. We find
good agreement where our samples overlap at low redshift and see that the $L_{250}$ (and
therefore the SFR) of our objects continues to increase out to $z = 5$. Since we are
stacking and detecting signal below the noise, this trend is not simply attributable
to Malmquist bias i.e. the preferential detection of more luminous objects at high
redshift due to the H-ATLAS survey sensitivity limits.

At low $L_{1.4}$, we measure lower average $L_{250}$ than V13. The results of section 5.4.4
above suggest that this may be due to starburst galaxies remaining in the V13 sample.
Indeed, V13 use a less conservative FIRC cut of $q_{250} > 1.3$ as opposed to our $q_{250} >
1.1$ - see section 5.4.2 for an explanation of this choice. In addition, V13 are also
biased to brighter, more massive galaxies. At higher $L_{1.4}$, we measure much higher
average $L_{250}$ than V13, although we note that the V13 sample suffers from low number
statistics here with only 20 and 4 sources in the two highest bins, respectively. There
is also far more discrepancy between the two $K$-correction models at high $L_{1.4}$ in
our sample than there is for the V13 sample. The isothermal $K$-correction model is
more successful for the low redshift V13 sample and best fitting temperatures span
the relatively small range of 16$< T <$23 K. The fitted temperatures for our sample
span the larger range 16$< T <$40 K.

As seen in Figure 5.11, there is a strong correlation between redshift and $L_{1.4}$ in
our sample. Therefore, the excess $L_{250}$ we see at high $L_{1.4}$ compared to V13 is likely
caused by higher dust temperatures in the higher redshift objects in our sample. That
is, we are seeing an evolution in dust temperature that is not seen in the low redshift
population of radio galaxies. However, we caution that the temperatures we derive
here are not truly representative of the entire radio galaxy population. Rather, the $\chi^2$
fitting will be biased towards sources more luminous in the infrared. We also expect
a systematic increase in the temperature with redshift since objects in our sample
must be hotter to be luminous in the H-ATLAS bands and therefore to contribute to
the fitting process.

In Figure 5.11, the upper dotted line indicates the point above which HERGs begin to
dominate the AGN population \(10^{26} \text{W Hz}^{-1}\); Best & Heckman, 2012). It is evident that there is likely to be a large fraction of HERGs in our sample, in contrast to the lower redshift studies of H10 and V13 which are dominated by LERGs due to their particular selection criteria. H13 found that HERGs have distinctly higher \(L_{250}\), SFRs and dust temperatures than LERGs. This agrees with the scenario where dust in LERGs is in thermal equilibrium with the quiescent background stars, while HERGs have large amounts of active star formation which heats the dust.

It is possible that the redshift trends in the average \(L_{250}\) and SFR of our sample may be partly attributable to a change in the dominant radio galaxy population (from LERGs to HERGs) across the bins. However, it is unlikely that this is the sole explanation for the trend. H13 also see evidence for an increase in the \(L_{250}\) and dust temperature of HERGs over \(0 < z < 1\). This suggests that our results reflect a genuine evolution in the overall radio galaxy population, rather than simply a sample selection bias.

Our conclusion is that radio galaxies display increasing SFRs with increasing \(z\) and \(L_{1.4}\). This continues trends seen at lower redshift \((z < 0.8)\) by H10 and V13 and by Seymour et al. (2011) who also found a considerable trend in the SFR of radio galaxies with redshift - up to 150 times the local value by \(z \approx 3\). As discussed above, our low redshift results are consistent with these.

The challenge now is to decipher the primary driver behind this evolution. H10 and V13 found little difference between the FIR trends of radio-loud AGN and normal galaxies matched in redshift and \(K\) band luminosity at \(z < 0.8\). Furthermore, the ‘main sequence’ population of star-forming radio-quiet galaxies seem to display increasing far-infrared luminosities up to \(z \sim 2.5\) (e.g. Elbaz et al., 2011). Therefore, it seems likely that the evolution of the galaxy population as a whole has a strong effect on the SFR evolution we observe in our sample.

Indeed, H13 found an excess in the \(L_{1.4}\) of HERGs, compared to the radio-quiet population. Seymour et al. (2011), on the other hand, find no strong trends in the
SFR with $L_{1.4}$ in excess to the redshift dependence (although they are affected by low number statistics). However, they do note that they observe a stronger redshift evolution than seen for the infrared luminosity function, which traces general star formation. This could be attributable to a dependence on AGN activity, but may also be caused by increased merging events.

5.6 Conclusions

We have examined the far-infrared and star forming properties of 250 radio-loud AGN with $S_{1.4} > 10\text{ mJy}$ selected from the NVSS and FIRST radio continuum surveys. This selection is successful in providing a highly complete sample while excluding most star-forming (non-AGN) radio galaxies. We have used new, deep VIKING and INT near-infrared and optical imaging data to identify the host galaxies and derive photometric redshifts for a deeper population than previously possible. We have also obtained GAMA spectroscopic redshifts for 59 sources. We find that the photometric redshifts agree reasonably well with the available spectroscopic redshifts and can reproduce the known $K$-$z$ relation. We find that redshift error does not dominate the overall error in our results. We identified 21 quasars in our sample using their near-infrared colours or their optical spectral features, where available. We carefully excluded these sources from the analysis since our results are not robust to synchrotron contamination.

The final sample spans the range $0 \lesssim z \lesssim 5$ and $10^{23} \lesssim L_{1.4} \lesssim 10^{28} \text{ W Hz}^{-1}$. We extracted H-ATLAS SPIRE and PACS photometry for these objects. When converting these far-infrared fluxes into luminosities, we find that the $K$-correction is strongly dependent upon the assumed temperature of the dust. We cannot directly measure this temperature since few objects are detected in more than one H-ATLAS band, if any. Assuming an isothermal temperature for all sources gives unrealistic star formation rates. Instead, we bin by redshift and radio luminosity and find the best-fitting
temperature for each sub-sample. This provides star formation rates that are comparable to other values from the literature, at least at low redshift. We see evidence for evolution in the average dust temperature with redshift and $L_{1.4}$, although Malmquist bias affects these results.

We examine the average 250 $\mu$m luminosity by stacking within each bin. We find that the $L_{250}$ of strong radio galaxies, a proxy for the SFR, increases with both redshift and 1.4 GHz luminosity, continuing trends seen at lower redshift. These evolving trends could partially be attributable to a changing population of galaxies in our sample - from LERG-dominated at lower redshift, to predominantly HERGs at higher redshift. However, the majority of the SFR evolution is likely due to age-related evolution of the galaxy population as a whole. Other studies hint at a relationship between radio-loud activity and SFR, in ‘excess’ to the redshift dependence. However, we are currently unable to determine whether such an excess exists in our data, since we lack a high redshift, radio-quiet comparison population.

Future work will involve construction of such a redshift-matched radio-quiet sample. Expanding this work over the full 161 deg$^2$ H-ATLAS Phase 1 field will shortly be possible, as VIKING and INT data become available in these regions. This should expand the sample size ten-fold, permit finer binning at higher redshifts, and thus allow more detailed studies of the complex interplay between radio-loud AGN activity and star formation rates of the host galaxies.

Acknowledgments

The Herschel-ATLAS is a project with Herschel, which is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. The H-ATLAS website is http://www.h-atlas.org/. GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA
input catalogue is based on data taken from the SDSS and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KIDS, VISTA VIKING, WISE, HATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO and the Participating Institutions. The GAMA website is http://www.gama-survey.org/.
Chapter 6

Summary and future work

6.1 Thesis summary

This thesis develops and implements a statistical data analysis technique, referred to as ‘stacking’, to push the sensitivity limitations of observational studies of galaxy evolution. We show that stacking is a powerful tool for examining average properties of faint and/or distant galaxies. We apply this technique to the spectral axis of 21 cm radio data to examine the average H $\text{I}$ properties of galaxies. We also stack far-infrared images of radio-loud AGN to examine the average star forming properties of these systems. Here we summarise the technique and the conclusions of these analyses.

The stacking process and data quality considerations are presented in Chapter 2. Unlike conventional observational techniques, stacking does not require individual detections. Rather, target objects in the observed field are identified using an external source catalogue. Data at these source positions are extracted, aligned either along the spectral axis or in the image plane, and then combined using an appropriate weighted average. This can result in a strong statistical detection of the population.

We find that inaccuracies or incompleteness in the external catalogue can significantly impact the conclusions of a stacking analysis and must be carefully considered. For
example, accurate redshifts are required for H\textsc{i} spectral stacking. If the redshift errors are too large (greater than the median H\textsc{i} profiles of the individual galaxies) then the co-added spectrum will be artificially broadened and measured average quantities may be incorrect. Since redshift accuracy is usually required to $< 150\,\text{km}\,\text{s}^{-1}$, this precludes the use of photometric redshifts which generally have much larger errors. Therefore, accurate spectroscopic redshift data are required to facilitate spectral stacking analyses. We also find that the results of a stacking analysis are particularly sensitive to the influence of any signal in the data from sources other than the targeted astronomical objects. Such artificial signals present in the H\textsc{i} spectral data include RFI from satellites and standing wave response patterns induced by strong continuum sources. If these are not carefully excised, they may restrict the achievable sensitivity, contribute to poor baselines or entirely prevent a stacked detection.

In Chapter 3 we stack 21 cm data to examine the average H\textsc{i} properties of galaxies over a larger volume than usually possible. H\textsc{i} is the primary fuel for star formation and therefore plays an essential role in galaxy evolution studies. Constraining the cosmic H\textsc{i} mass density ($\Omega_{\text{H}_\text{i}}$) evolution is vital in understanding why the star formation rate density of the Universe declines rapidly below $z \sim 2 - 3$. However, $\Omega_{\text{H}_\text{i}}$ is not well measured beyond the local Universe, mostly due to telescope sensitivity limitations. A number of previous studies have surveyed relatively small sky areas and used either long integration times or the stacking technique to push the redshift boundary of H\textsc{i} detections. We show that the stacking technique can be applied to fairly shallow, yet wide-field data to make accurate measurements of $\Omega_{\text{H}_\text{i}}$ with minimal cosmic variance bias.

We apply the stacking technique to wide-field 21 cm data from HIPASS and also to higher-redshift data over a 42\,deg$^2$ SGP field collected with the Parkes radio telescope. Optical source positions and spectroscopic redshifts with $85\,\text{km}\,\text{s}^{-1}$ accuracy are provided by the 2dF Galaxy Redshift Survey (2dFGRS). We show that stacking 15,093 2dFGRS galaxies at $z < 0.04$ and 3,277 galaxies at $0.04 < z < 0.13$ produces strong statistical detections. The achieved sensitivity is equivalent to ob-
serving a single source for 37 d (~888 h). Thus, stacking is an efficient method of obtaining high-significance detections at intermediate redshift using relatively short (~52 h) integration times. We find no intrinsic limitation in the sensitivity that can be achieved via stacking, provided non-Gaussian noise components (such as RFI) are carefully removed.

While the large beamwidth of the Parkes telescope allows rapid survey speeds and therefore longer integration time per field, this also introduces considerable confusion (multiple sources in a beam) into the H gas data. Therefore, average H I masses may be greatly overestimated. When calculating H I mass-to-light ratios, this effect can be counteracted by artificially confusing the luminosities to match the extent of confusion in the H I data. These adjusted ratios are then used in combination with the luminosity density function to derive the cosmic H I mass density (\( \Omega_{\text{H I}} \)). We find \( \Omega_{\text{H I}} = (2.82^{+0.30}_{-0.59}) \times 10^{-4} h^{-1} \) at \( 0 < z < 0.04 \) and \( (3.19^{+0.43}_{-0.59}) \times 10^{-4} h^{-1} \) at \( 0.04 < z < 0.13 \). This is consistent with no evolution over the past \( \sim 1 h^{-1} \) Gyr. These results agree with other measurements in the literature over the given redshift interval, including those from the HIPASS, ALFALFA and AUDS surveys. This indicates consistency between the results of 21 cm stacking and direct detection methods.

The particular advantage of using the stacking method in this context is that it does not require direct detections and therefore allows measurements of \( \Omega_{\text{H I}} \) over much larger space volumes than otherwise possible: \( 3.7 \times 10^5 \) Mpc\(^3\) at \( z < 0.04 \) and \( 3.4 \times 10^5 \) Mpc\(^3\) at \( 0.04 < z < 0.13 \). Our \( \Omega_{\text{H I}} \) measurements are therefore far less susceptible to the influence of cosmic variance bias than previous studies, even for the large all-sky surveys. Finally, we show that stacking can reproduce observed trends in galaxy H I content with luminosity and colour. Specifically, we find that bluer galaxies (with \( b_J - r_F \approx 0.3 \)) have H I mass-to-\( b_J \) luminosity ratios \( \sim 12 \) times higher than redder galaxies (with \( b_J - r_F \approx 1.4 \)). We also find that low luminosity galaxies have higher mass-to-light ratios than high luminosity galaxies, following similar trends found in the literature.

In Chapter 4 we extend this work with an H I stacking analysis of Galaxy and Mass
Assemble (GAMA) galaxies. We focus on two equatorial 48 deg² GAMA fields with low redshift (z < 0.04) 21 cm data provided by HIPASS and new, higher redshift (0.04 < z < 0.14) HI data collected with the Parkes radio telescope. Severe RFI at low frequencies reduces the usable range of the data to z < 0.109. The GAMA redshift survey is 1.1 magnitudes deeper than the 2dFGRS. Therefore, GAMA probes further down the faint end of the luminosity function and provides more spectra per square degree available for stacking. GAMA source selection is also well-characterised. This allows us to carefully account for completeness and selection bias in our measurements. The multiwavelength nature of the GAMA campaign has enabled other authors to derive accurate stellar masses and stellar mass density functions. We capitalise on the availability of such information to compare various methods of deriving $\Omega_{HI}$, from a stacking analysis. This includes the difference between using HI mass-to-light ratios versus HI-to-stellar mass ratios, as well as the difference between various weighting and averaging strategies. We find good consistency between the different averaging methods and associated sample completeness corrections. We also find that HI mass-to-light ratios at z < 0.04 and HI-to-stellar mass ratios at 0.04 < z < 0.11 provide $\Omega_{HI}$ measurements consistent with the results of the 2dFGRS analysis and with other results in the literature. However, in a given redshift interval, the use of luminosity or stellar mass fractions give systematically different results. This can only partially be explained by cosmic variance.

Overall, we again conclude that little evolution in the HI content of galaxies is seen over the past 0.9 h⁻¹ Gyr. Since the cosmic star formation rate density does evolve considerably over this range, this suggests that significant replenishment of neutral gas must be occurring in galaxies. This is consistent with simulations and theoretical models of galaxy evolution which predict a balance between infalling gas and outflows, resulting in regulated gas masses for galaxies and little evolution of the global density of HI.

The star forming properties of radio galaxies are examined in Chapter 5. The complex relationship between AGN activity and star formation plays a key role in our
understanding of galaxy evolution. It is unknown how the feedback from the jets of radio-loud AGN interacts with the interstellar medium, enhances or hinders further star formation and thereby influences the evolutionary pathway of the host galaxy. Many previous studies have been restricted to small samples or have traced star formation at wavelengths contaminated by the AGN.

We use wide-field, far-infrared (FIR) data from the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS) to overcome these barriers. This data provides a much larger galaxy sample than previously achievable and at wavelengths that are excellent, uncontaminated probes for star formation. The additional availability of new, sensitive optical and near-infrared data over the 14 deg$^2$ Science Demonstration Phase (SDP) field allows the identification of counterparts to 250 NVSS/FIRST radio sources and the derivation of photometric redshifts for objects out to $z \sim 5$. This is a new regime which has not been examined in previous studies, including recent stacking analyses. Stacking the 250 $\mu$m H-ATLAS data overcomes Malmquist bias and allows statistical detections of FIR emission out to these high redshifts. We are therefore able to examine the average star formation rates (SFRs) of large numbers of radio galaxies over a wider redshift ($0 < z < 5$) and radio luminosity range ($23 < \log L_{1.4} < 28$) than previously possible.

Since individual galaxies in our sample are rarely detected in one or more H-ATLAS band, the main difficulty in calculating FIR luminosities and SFRs is determining dust temperatures and therefore the appropriate $K$-correction. Unlike similar samples of galaxies at low redshift, we find that average measured SFRs of the studied population are heavily dependent on the assumed dust temperature. By determining the best-fitting temperature within bins of redshift and 1.4 GHz luminosity, we find reasonable average SFRs at low redshift and individual high-redshift SFRs ($511$ and $1075 M_\odot$ yr$^{-1}$) consistent with other measurements in the literature. We find that the average SFR of the radio galaxy sample continues to increase with both redshift and 1.4 GHz luminosity. This may be partially explained by a transition from predominantly low-excitation radio galaxies in the sample at low redshift to mainly
high-excitation sources at higher redshift. However, the observed trends are most likely attributable to evolution of the overall galaxy population. We await the availability of a redshift-matched, radio-quiet comparison sample of galaxies to determine whether the data reveal any excess relationship between SFR and radio-loud AGN activity. Nonetheless, this work has demonstrated the important role of the stacking technique in studies of galaxy evolution and AGN physics.

6.2 Future work

The field of galaxy evolution is beginning to develop rapidly thanks to a suite of new and planned survey telescopes employing recent technological advances. The next major leap forward in this field will be provided by the SKA and the pathfinder radio telescopes including ASKAP, MeerKAT and APERTIF. These instruments will completely revolutionise the field of radio astronomy and can be used in synergy with cutting-edge multiwavelength facilities to potentially enable incredibly exciting new discoveries.

Development of the stacking technique will be vital for radio astronomy in the future – when statistical analyses of large galaxy populations will be required to make sense of the extremely large data sets. In the short term, progress can be made using the Compact Array Broad-band Backend (CABB) recently installed at ATCA. CABB has ‘zoom mode’ capability providing fine spectral resolution (32kHz/channel) over the entire broad, instantaneous bandwidth (> 180MHz). This, combined with the excellent spatial resolution (< 1 arcmin) of ATCA, will facilitate spectral stacking studies out to $z \sim 0.18$ with very little source confusion. For this purpose, we have already initiated a program of follow-up observations of the SGP and GAMA fields with ATCA. However, the small field-of-view of this instrument means that long integration times are required to survey the full regions to a sensitivity sufficient for achieving stacked detections.
The next generation of radio telescopes will be truly superb for H\textsc{i} stacking experiments. The advanced aperture and/or correlator technology will simultaneously provide a large field-of-view (and therefore fast survey speeds) as well as a fine angular resolution (minimising confusion). They will also have the capacity to operate at frequencies corresponding to high H\textsc{i} redshifts.

Future H\textsc{i} surveys planned for the pathfinder instruments have been designed to overlap with spectroscopic surveys. For example, DINGO will provide deep H\textsc{i} data over the GAMA fields out to $z \approx 0.4$. WALLABY\textsuperscript{1} will be an ASKAP all-sky H\textsc{i} survey with substantially better sensitivity and resolution than HIPASS. To complement this, the TAIPAN\textsuperscript{2} all-sky spectroscopic survey (Hopkins et al., in prep) is planned for the UK-Schmidt telescope. On APERTIF, the MDS\textsuperscript{3} (Verheijen et al. 2009) will partially target the sky region covered by the HETDEX\textsuperscript{4} spectroscopic redshift survey (Hill et al. 2004). The LADUMA\textsuperscript{5} survey (Holwerda et al. 2012) on MeerKAT will target the COSMOS field where a plethora of multiwavelength information is available.

The availability of such high quality H\textsc{i} data and ancilliary optical spectroscopic information will allow accurate stacking experiments to be conducted. Even though these instruments will already have the capacity to directly detect H\textsc{i} signatures at higher redshifts than ever before, stacking can always be used to extract information from beyond the sensitivity limits. Certainly, the results of this thesis show no intrinsic limitation in the sensitivity achievable via stacking. Direct and stacked detections with these pathfinder and future SKA surveys will provide advanced insights into the evolutionary patterns of neutral hydrogen gas out to unprecedented redshifts.

Use of the spectral stacking technique is not restricted to the 21 cm emission line, but should be applicable to any spectral signature. For example, stacking of CO emission

\begin{itemize}
  \item \textsuperscript{1}Wide-field ASKAP L-band Legacy All-sky Blind Survey (www.atnf.csiro.au/research/WALLABY)
  \item \textsuperscript{2}Transforming Astronomical Imaging-surveys through Polychromatic Analysis of Nebulae
  \item \textsuperscript{3}Medium-Deep Survey
  \item \textsuperscript{4}Hobby-Eberly Telescope Dark Energy Experiment
  \item \textsuperscript{5}Looking at the Distant Universe with the MeerKAT Array
\end{itemize}
lines could be used to extend observational studies of the molecular gas content of galaxies. While little evolution of $\Omega_{\text{HI}}$ is found over $0 < z < 1$, this is the range over which most of the stars in galaxies were formed (Hopkins & Beacom 2006). Since the molecular gas phase is the transition between neutral gas and star formation, it is possible that stronger evidence of evolution in the molecular gas density ($\Omega_{\text{mol}}$) may be seen. Very few observational constraints on $\Omega_{\text{mol}}$ exist at these high redshifts due to sensitivity limits of existing submillimetre telescopes, and only for individual galaxies with intense star-formation (e.g. Daddi et al. 2010, Aravena et al. 2012).

The recently-commissioned Atacama Large Millimeter/submillimeter Array (ALMA) boasts superior sensitivity and larger bandwidth than any preceding submillimetre instrument (Wootten & Thompson 2009). By applying the stacking technique to ALMA CO observations and combining this information with high redshift $\text{H} \, \text{I}$ measurements, it may be possible to track the co-evolution of atomic to molecular gas fractions. This could provide improved explanations for the variation in the cosmic star-formation rate.

The conclusions of the Herschel stacking analysis, presented in Chapter 5, are currently restricted by the lack of a redshift-matched, radio-quiet comparison sample. Such high-redshift, radio-quiet galaxies are rare and difficult to identify within the limits of the optical and near-infrared data in the SDP field. However, H-ATLAS, INT and VIKING data are now available over the full $161 \, \text{deg}^2$ Phase 1 region. Using this wide-field data, it will be possible to construct a sufficiently large comparison sample such that statistical FIR detections are achieved out to $z = 5$. In fact, the analysis of the radio-loud AGN can also be expanded over this wider area and should allow a ten-fold increase in the sample size. This will enable finer binning at high redshifts and the identification of any excess influence on the star formation rate by the radio-loud AGN activity. Finally, the accuracy of the results can be improved further if spectroscopy of all relevant objects can be obtained. This will not only provide accurate redshifts, but will allow the easy identification of contaminating quasars. Multi-object spectrographs capable of such surveys include the AutoFib2 Wide Field
Fibre Optical Spectrograph (AF2+WYFFOS) on the William Herschel Telescope and the Very Large Telescope’s Visible Multiobject Spectrograph (VIMOS). The combination of these improved, multiwavelength data sets will provide a powerful probe of the inter-relationship between AGN feedback, star formation and galaxy evolution in general.
References

Aird J. et al., 2010, Monthly Notices of the Royal Astronomical Society, 401, 2531
Baldry I. K. et al., 2010, Monthly Notices of the Royal Astronomical Society, 404, 86


Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101


Benson A. J., 2010, Physics Reports, 495, 33


Croom S. M. et al., 2009, Monthly Notices of the Royal Astronomical Society, 399, 1755


Daddi E. et al., 2010, The Astrophysical Journal, 713, 686
REFERENCES


Driver S. P. et al., 2011, Monthly Notices of the Royal Astronomical Society, 413, 971

Driver S. P. et al., 2009, Astronomy and Geophysics, 50, 050000


192
REFERENCES

Dye S. et al., 2010, Astronomy and Astrophysics, 518, L10

Eales S. et al., 2010, Publications of the Astronomical Society of the Pacific, 122, 499


Elbaz D. et al., 2011, Astronomy and Astrophysics, 533, A119


Griffin M. J. et al., 2010, Astronomy and Astrophysics, 518, L3

Haas M. et al., 2004, Astronomy and Astrophysics, 424, 531


Hardcastle M. J. et al., 2010, Monthly Notices of the Royal Astronomical Society, 409, 122


REFERENCES


Ibar E. et al., 2010, Monthly Notices of the Royal Astronomical Society, 409, 38

Ilbert O. et al., 2006, Astronomy and Astrophysics, 457, 841


Jarvis M. J. et al., 2010, Monthly Notices of the Royal Astronomical Society, 409, 92

Johnston S. et al., 2008, Experimental Astronomy, 22, 151


Karachentsev I. D., Makarov D. I., Karachentseva V. E., Melnik O. V., 2008, Astronomy Letters, 34, 832


Lah P. et al., 2007, Monthly Notices of the Royal Astronomical Society, 376, 1357
Lah P. et al., 2009, Monthly Notices of the Royal Astronomical Society, 399, 1447
Lawrence A. et al., 2007, Monthly Notices of the Royal Astronomical Society, 379, 1599


Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution

REFERENCES


Pascale E. et al., 2011, Monthly Notices of the Royal Astronomical Society, 415, 911


REFERENCES

Pilbratt G. L. et al., 2010, Astronomy and Astrophysics, 518, L1

Poglitsch A. et al., 2010, Astronomy and Astrophysics, 518, L2


Rawlings S., 2011, arXiv1105.6333


Robotham A. et al., 2010, Publications of the Astronomical Society of Australia, 27, 76


REFERENCES


Seymour N. et al., 2011, Monthly Notices of the Royal Astronomical Society, 413, 1777


Smith D. J. B. et al., 2012, Monthly Notices of the Royal Astronomical Society, 427, 703


Springel V. et al., 2005, Nature, 435, 629


REFERENCES


Wright E. L. et al., 2010, The Astronomical Journal, 140, 1868


Zwaan M. A. et al., 2003, The Astronomical Journal, 125, 2842

Appendix A

List of Acronyms

2dF: Two Degree Field
2dFGRS: 2dF Galaxy Redshift Survey
2PIGG: 2dFGRS Percolation-Inferred Galaxy Group
AAO: Australian Astronomical Observatory
AAT: Anglo-Australian Telescope
AGN: Active Galactic Nuclei
ALFA: Arecibo L-band Feed Array
ALFALFA: Arecibo Legacy Fast ALFA
ALMA: Atacama Large Millimeter/submillimeter Array
APERTIF: Aperture Tile In Focus
ASA: Astronomical Society of Australia
ASKAP: Australian Square Kilometre Array Pathfinder
ATCA: Australia Telescope Compact Array
ATNF: Australian National Telescope Facility
AUDS: Arecibo Ultra Deep Survey
CABB: Compact Array Broadband Backend
CASS: CSIRO Astronomy and Space Science
CHILES: COSMOS H I Large Extragalactic Survey
APPENDIX A. LIST OF ACRONYMS

COSMOS: Cosmological Evolution Survey

CSIRO: Commonwealth Scientific and Industrial Research Organisation

Dec: Declination

DINGO: Deep Investigation of Neutral Gas Origins

DLA: Damped Lyman Alpha

ETG: Early-Type Galaxy

FIR: Far infra-red

FIRC: FIR-radio correlation

FIRST: Faint Images of the Radio Sky at Twenty-Centimetres

FRI/II: Fanaroff- Riley class I/II

G09: GAMA 9\textsuperscript{h} field

G12: GAMA 12\textsuperscript{h} field

G15: GAMA 15\textsuperscript{h} field

GALEX: Galaxy Evolution Explorer

GAMA: Galaxy and Mass Assembly

GASS-1: GALEX Arecibo SDSS Survey

GBT: Green Bank Telescope

GMRT: Giant Metre-wave Radio Telescope

H-ATLAS: Herschel Astrophysical Terahertz Large Area Survey

HERG: High-excitation radio galaxy

HETDEX: Hobby-Eberly Telescope Dark Energy Experiment

HI: Neutral atomic hydrogen

HICAT: HIPASS Catalogue

HIMF: HI mass function

HIPASS: HI Parkes All Sky Survey

HOPCAT: HIPASS Optical Catalogue

HST: Hubble Space Telescope
ICRAR: International Centre for Radio Astronomy Research
IF: Intermediate frequency
ICM: Intracluster medium
IGM: Intergalactic medium
ISM: Interstellar medium
IMF: Initial Mass Function
INT: Isaac Newton Telescope
JVLA: Jansky Very Large Array
KAT: Karoo Array Telescope
K-S: Kolmogorov-Smirnov
LADUMA: Looking at the Distant Universe with the MeerKAT Array
LERG: Low-excitation radio galaxy
LOFAR: Low Frequency Array
MDS: Medium Deep Survey
MIRIAD: Multichannel Image Reconstruction, Image Analysis and Display
NIR: Near infra-red
NRAO: National Radio Astronomy Observatory
NVSS: NRAO VLA Sky Survey
PACS: Photodetector Array Camera and Spectrometer
PAH: Polycyclic Aromatic Hydrocarbon
PSF: Point-Spread Function
QSO: Quasi-Stellar Object
RA: Right Ascension
RFI: Radio Frequency Interference
SAM: Semi-analytical Model
SCUBA: Submillimetre Common User Bolometer Array
SED: Spectral Energy Distribution
APPENDIX A. LIST OF ACRONYMS

SDP: Science Demonstration Phase
SDSS: Sloan Digital Sky Survey
SFR: Star-formation rate
SGP: South Galactic Pole
SKA: Square Kilometre Array
SMBH: Supermassive Black Hole
S/N: Signal-to-noise ratio
SPH: Smoothed-Particle Hydrodynamics
SPIRE: Spectral and Photometric Imaging Receiver
TAIPAN: Transforming Astronomical Imaging-surveys through Polychromatic Analysis of Nebulae
UKIDSS-LAS: UKIRT Infrared Deep Sky Survey - Large Area Survey
UKIRT: United Kingdom Infrared Telescope
UV: Ultra-violet
VISTA: Visible and Infrared Survey Telescope for Astronomy
VIKING: VISTA Kilo-Degree Infrared Galaxy Survey
VLA: Very Large Array
VLT: Very Large Telescope
VST: VLT Survey Telescope
WALLABY: Wide-field ASKAP L-band Legacy All-sky Blind Survey
WFC: Wide Field Camera
WISE: Wide-field Infrared Survey Explorer
WMAP: Wilkinson Microwave Anisotropy Probe
WSRT: Westerbork Synthesis Radio Telescope
Appendix B

Luminosity and stellar mass density functions and the completeness extrapolation factor

In this appendix we present the luminosity and stellar mass density functions derived from GAMA data by various authors and used in the analysis of Chapter 4. From these, we calculate the completeness extrapolation factors required for the confusion correction applied in Chapter 4.

B.1 Luminosity density

The \( r \) band luminosity density function can be described by a Schechter function:

\[
\rho(L_r) dL_r = \phi^* \left( \frac{L_r}{L^*_r} \right)^{\alpha+1} e^{-\left( \frac{L_r}{L^*_r} \right)} \frac{dL_r}{L^*_r}.
\]  

(B.1)

Table 4 of Driver et al. (2012) defines the Schechter parameters over the range \( 0.013 < z < 0.1 \) as \( \phi^* = (1.24^{+0.04}_{-0.03}) \times 10^{-2} \, h^3 \text{Mpc}^{-3} \), \( \alpha = -1.12^{+0.01}_{-0.01} \) and \( L_r = 10^{0.4(M_{r,\odot} - M^*_r)} \)

where \( M^*_r = -20.86^{+0.04}_{-0.02} + 5 \log h \) and \( M_{r,\odot} = 4.71 \) (as defined in Hill et al., 2010).
APPENDIX B. COMPLETENESS FACTORS  B.1. LUMINOSITY DENSITY

As in Driver et al. (2012), the total luminosity density (the integral of the luminosity density function over $L_r$) can be defined as:

$$\rho_{L_r} = \int_0^{\infty} \rho(L_r)dL_r = \phi^* L_r \Gamma(\alpha + 2), \quad (B.2)$$

where the complete gamma function is:

$$\Gamma(a) = \int_0^{\infty} t^{a-1}e^{-t}dt. \quad (B.3)$$

Inserting the Schecter function parameters of Driver et al. (2012), equation B.2 gives $\rho_{L_r} = (2.27^{0.02}\pm0.01) \times 10^8 L_\odot h \text{Mpc}^{-3}$. However, Driver et al. (2012) caution that this function does not well-constrain the upturn in the faint end of the luminosity function. Hence they recommend a ‘summation’ approach to calculating the total luminosity density, resulting in a $\sim 1\%$ difference: $\rho_{L_r} = (2.29^{+0.06}_{-0.06}) \times 10^8 L_\odot h \text{Mpc}^{-3}$.

In Chapter 4, the galaxy luminosities are artificially confused via Equation 4.5 to mimic the confusion in the 21cm data. This equation requires a completeness extrapolation factor, $f_{L_r}$, to account for galaxies below the GAMA sensitivity limits. For each galaxy, this is defined as:

$$f_{L_r} = \frac{\rho_{L_r}}{\rho_{L_r,\text{lim}}}, \quad (B.4)$$

where $L_{r,\text{lim}}$ is the luminosity to which GAMA is complete at the redshift of the galaxy in question. $\rho_{L_r,\text{lim}}$ is the definite integral of the luminosity density function between 0 and $L_{\text{lim}}$:  

209
\[ \rho_{L_r,\text{lim}} = \int_0^{L_{\text{lim}}} \rho(L_r) dL_r = \phi^* L_r (\Gamma(\alpha + 2, \frac{L_{\text{lim}}}{L_r^*}) ) \tag{B.5} \]

where the incomplete gamma function is:

\[ \Gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt. \tag{B.6} \]

Therefore, the completeness extrapolation factor becomes:

\[ f_{L_r} = \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 2, \frac{L_{\text{lim}}}{L_r^*})} \tag{B.7} \]

### B.2 Stellar mass density

The stellar mass density function is defined by Baldry et al. (2012) as a double Schechter function:

\[ \rho(M) dM = M^* \left[ \phi_1^* \left( \frac{M}{M^*} \right)^{\alpha_1+1} + \phi_2^* \left( \frac{M}{M^*} \right)^{\alpha_2+1} \right] e^{-M/M^*} \frac{dM}{M^*} \tag{B.8} \]

where \( \log M^* = 10.35 \pm 0.05 + 2 \log(h) \), \( \phi_1^* = (11.5 \pm 0.99) \times 10^{-3} h^3 \), \( \phi_2^* = (2.33 \pm 0.67) \times 10^{-3} h^3 \), \( \alpha_1 = -0.35 \pm 0.18 \) and \( \alpha_2 = -1.47 \pm 0.05 \).

The total stellar mass density is defined as:

\[ \]
\[ \rho_M = \int_0^\infty \rho(M)dM \]
\[ = \phi_1^* M^* \Gamma(\alpha_1 + 2) + \phi_2^* M^* \Gamma(\alpha_2 + 2) \quad (B.9) \]

and is equal to \( \rho_M = (3.2 \pm 0.5) \times 10^8 h M_\odot \text{Mpc}^{-3} \) (Baldry et al., 2012).

As in Equation B.4, the stellar mass completeness extrapolation factor for galaxies below the survey limits can be defined as:

\[ f_M = \frac{\rho_M}{\rho_{M,\text{lim}}} \quad (B.10) \]

where \( M_{\text{lim}} \) is the value of the GAMA stellar mass completeness function at the redshift of the galaxy in question. This function is defined by Robotham et al., (in prep) and is shown in Figure 4.3 of Chapter 4. \( \rho_{M,\text{lim}} \) is the integral of the stellar mass density function between 0 and \( M_{\text{lim}} \):

\[ \rho_{M,\text{lim}} = \int_0^{M_{\text{lim}}} \rho(M)dM \]
\[ = \phi_1^* M^* (\Gamma(\alpha_1 + 2, M_{\text{lim}}/M^*) + \phi_2^* M^* (\Gamma(\alpha_2 + 2, M_{\text{lim}}/M^*)) \quad (B.11) \]

Therefore, the completeness extrapolation factor becomes:

\[ f_M = \frac{\phi_1^* M^* (\Gamma(\alpha_1 + 2, M_{\text{lim}}/M^*) + \phi_2^* M^* (\Gamma(\alpha_2 + 2, M_{\text{lim}}/M^*))}{\phi_1^* M^* (\Gamma(\alpha_1 + 2, M_{\text{lim}}/M^*) + \phi_2^* M^* (\Gamma(\alpha_2 + 2, M_{\text{lim}}/M^*))} \quad (B.12) \]
Appendix C

Radio galaxy catalogue

This appendix presents the data catalogue for all NVSS/FIRST radio galaxies used in the analysis of Chapter 5. Only selected parameters are shown in the table below. From left to right, these columns are: The object identifier number (for comparison with the images in Appendix D), the right ascension and declination, the source of the best object position (see below), the best available redshift, the NVSS 1.4 GHz flux (mJy), and a flag to identify quasars (see below).

In each case, the position has been defined by: V = VIKING source catalogue, F = FIRST source catalogue, N = NVSS source catalogue, M = manually.

The key to the quasar flags is as follows: 0 = No classification available, 1 = Not a quasar (determined via GAMA spectra), 2 = Confirmed quasar (determined via GAMA spectra), 3 = Suspected quasar (identified via NIR colours).
<table>
<thead>
<tr>
<th>ID</th>
<th>RA</th>
<th>Dec</th>
<th>Position</th>
<th>z</th>
<th>1.4 GHz flux</th>
<th>Q flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.0</td>
<td>08:57:01.77</td>
<td>+01:31:31.0</td>
<td>V</td>
<td>0.27332</td>
<td>96.3</td>
<td>1</td>
</tr>
<tr>
<td>25.0</td>
<td>08:57:21.43</td>
<td>+01:11:12.8</td>
<td>M</td>
<td>2.5667</td>
<td>18.8</td>
<td>3</td>
</tr>
<tr>
<td>26.0</td>
<td>08:57:24.22</td>
<td>+01:22:58.7</td>
<td>V</td>
<td>2.5581</td>
<td>11.3</td>
<td>0</td>
</tr>
<tr>
<td>34.0</td>
<td>08:57:49.80</td>
<td>+01:35:30.4</td>
<td>V</td>
<td>0.28124</td>
<td>88.5</td>
<td>1</td>
</tr>
<tr>
<td>35.0</td>
<td>08:57:50.56</td>
<td>+01:37:57.1</td>
<td>V</td>
<td>0.27442</td>
<td>10.2</td>
<td>1</td>
</tr>
<tr>
<td>42.0</td>
<td>08:58:04.94</td>
<td>+01:09:13.5</td>
<td>V</td>
<td>0.07119</td>
<td>20.2</td>
<td>1</td>
</tr>
<tr>
<td>45.0</td>
<td>08:58:07.61</td>
<td>+00:34:30.8</td>
<td>V</td>
<td>1.2876</td>
<td>23.9</td>
<td>0</td>
</tr>
<tr>
<td>60.0</td>
<td>08:58:34.25</td>
<td>+00:49:57.5</td>
<td>V</td>
<td>0.6676</td>
<td>23.8</td>
<td>1</td>
</tr>
<tr>
<td>68.0</td>
<td>08:58:49.36</td>
<td>+00:52:24.0</td>
<td>V</td>
<td>0.5819</td>
<td>22.5</td>
<td>0</td>
</tr>
<tr>
<td>69.0</td>
<td>08:58:49.70</td>
<td>+00:40:13.0</td>
<td>V</td>
<td>0.8972</td>
<td>33.4</td>
<td>0</td>
</tr>
<tr>
<td>70.0</td>
<td>08:58:48.10</td>
<td>+01:38:24.1</td>
<td>F</td>
<td>3.2783</td>
<td>10.4</td>
<td>0</td>
</tr>
<tr>
<td>70.1</td>
<td>08:58:50.71</td>
<td>+01:37:57.5</td>
<td>F</td>
<td>3.655</td>
<td>18.7</td>
<td>3</td>
</tr>
<tr>
<td>80.0</td>
<td>08:59:09.76</td>
<td>+01:35:19.5</td>
<td>V</td>
<td>1.1711</td>
<td>40.1</td>
<td>0</td>
</tr>
<tr>
<td>84.0</td>
<td>08:59:14.18</td>
<td>+00:41:36.4</td>
<td>F</td>
<td>2.6258</td>
<td>12.9</td>
<td>0</td>
</tr>
<tr>
<td>85.1</td>
<td>08:59:15.40</td>
<td>+01:02:26.4</td>
<td>V</td>
<td>1.882</td>
<td>28.5</td>
<td>0</td>
</tr>
<tr>
<td>88.0</td>
<td>08:59:20.56</td>
<td>+00:47:12.3</td>
<td>V</td>
<td>0.464</td>
<td>38.5</td>
<td>0</td>
</tr>
<tr>
<td>92.0</td>
<td>08:59:33.49</td>
<td>-00:09:00.8</td>
<td>V</td>
<td>1.65316</td>
<td>31.1</td>
<td>2</td>
</tr>
<tr>
<td>107.0</td>
<td>08:59:57.07</td>
<td>-00:03:32.0</td>
<td>F</td>
<td>4.7402</td>
<td>10.6</td>
<td>0</td>
</tr>
<tr>
<td>112.0</td>
<td>09:00:04.02</td>
<td>+01:35:56.0</td>
<td>V</td>
<td>2.7082</td>
<td>85</td>
<td>0</td>
</tr>
<tr>
<td>113.0</td>
<td>09:00:04.52</td>
<td>-00:25:48.5</td>
<td>V</td>
<td>0.61106</td>
<td>14.6</td>
<td>1</td>
</tr>
<tr>
<td>114.0</td>
<td>09:00:05.06</td>
<td>+00:04:46.8</td>
<td>V</td>
<td>0.05386</td>
<td>16.3</td>
<td>1</td>
</tr>
<tr>
<td>114.1</td>
<td>09:00:04.67</td>
<td>+00:03:32.1</td>
<td>V</td>
<td>0.26226</td>
<td>39.7</td>
<td>1</td>
</tr>
<tr>
<td>119.0</td>
<td>09:00:16.10</td>
<td>+00:13:46.5</td>
<td>V</td>
<td>2.0464</td>
<td>24.1</td>
<td>0</td>
</tr>
<tr>
<td>119.1</td>
<td>09:00:14.67</td>
<td>+00:13:05.4</td>
<td>V</td>
<td>1.8475</td>
<td>15.2</td>
<td>0</td>
</tr>
<tr>
<td>120.0</td>
<td>09:00:16.26</td>
<td>-00:25:16.9</td>
<td>V</td>
<td>0.6425</td>
<td>30.7</td>
<td>0</td>
</tr>
<tr>
<td>123.0</td>
<td>09:00:19.32</td>
<td>+00:27:02.6</td>
<td>N</td>
<td>1.5796</td>
<td>10.3</td>
<td>0</td>
</tr>
<tr>
<td>135.0</td>
<td>09:00:47.81</td>
<td>+00:28:40.5</td>
<td>N</td>
<td>4.5241</td>
<td>11.9</td>
<td>0</td>
</tr>
<tr>
<td>136.0</td>
<td>09:00:49.41</td>
<td>+01:02:03.4</td>
<td>V</td>
<td>0.5081</td>
<td>11.5</td>
<td>0</td>
</tr>
<tr>
<td>145.0</td>
<td>09:01:04.08</td>
<td>+02:02:01.1</td>
<td>F</td>
<td>1.6159</td>
<td>16.9</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>RA</td>
<td>Dec</td>
<td>Position</td>
<td>z</td>
<td>1.4 GHz flux</td>
<td>Q flag</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>--------</td>
<td>----------</td>
<td>-------</td>
<td>--------------</td>
<td>--------</td>
</tr>
<tr>
<td>155.0</td>
<td>09:01:15.70</td>
<td>+00:36:56.8</td>
<td>V</td>
<td>3.4535</td>
<td>12.4</td>
<td>0</td>
</tr>
<tr>
<td>159.0</td>
<td>09:01:18.94</td>
<td>-00:07:56.5</td>
<td>V</td>
<td>1.3203</td>
<td>30.4</td>
<td>0</td>
</tr>
<tr>
<td>164.0</td>
<td>09:01:25.85</td>
<td>-00:37:03.1</td>
<td>V</td>
<td>1.90795</td>
<td>328.7</td>
<td>2</td>
</tr>
<tr>
<td>166.0</td>
<td>09:01:26.45</td>
<td>+01:03:42.4</td>
<td>V</td>
<td>1.0891</td>
<td>34.7</td>
<td>0</td>
</tr>
<tr>
<td>171.0</td>
<td>09:01:30.14</td>
<td>-00:02:18.9</td>
<td>V</td>
<td>0.25458</td>
<td>62.7</td>
<td>1</td>
</tr>
<tr>
<td>185.0</td>
<td>09:01:39.55</td>
<td>+01:07:28.3</td>
<td>V</td>
<td>0.7608</td>
<td>38.7</td>
<td>0</td>
</tr>
<tr>
<td>187.0</td>
<td>09:01:40.59</td>
<td>-01:21:15.7</td>
<td>V</td>
<td>0.1223</td>
<td>14.5</td>
<td>0</td>
</tr>
<tr>
<td>188.0</td>
<td>09:01:41.50</td>
<td>+00:58:57.2</td>
<td>V</td>
<td>0.8237</td>
<td>11.3</td>
<td>0</td>
</tr>
<tr>
<td>189.0</td>
<td>09:01:41.92</td>
<td>+01:21:06.4</td>
<td>V</td>
<td>4.4106</td>
<td>37.7</td>
<td>0</td>
</tr>
<tr>
<td>196.0</td>
<td>09:01:50.56</td>
<td>+00:30:55.0</td>
<td>V</td>
<td>1.66848</td>
<td>41.7</td>
<td>2</td>
</tr>
<tr>
<td>197.0</td>
<td>09:01:50.47</td>
<td>+00:54:52.9</td>
<td>V</td>
<td>4.0566</td>
<td>12.1</td>
<td>0</td>
</tr>
<tr>
<td>202.0</td>
<td>09:01:53.17</td>
<td>-01:22:05.3</td>
<td>V</td>
<td>0.08</td>
<td>36.1</td>
<td>0</td>
</tr>
<tr>
<td>204.0</td>
<td>09:01:55.35</td>
<td>-00:33:34.6</td>
<td>V</td>
<td>0.5741</td>
<td>285.6</td>
<td>0</td>
</tr>
<tr>
<td>205.0</td>
<td>09:01:56.58</td>
<td>+01:43:45.3</td>
<td>V</td>
<td>0.5851</td>
<td>16.8</td>
<td>0</td>
</tr>
<tr>
<td>209.0</td>
<td>09:01:58.77</td>
<td>+01:38:33.4</td>
<td>V</td>
<td>0.35386</td>
<td>164.9</td>
<td>1</td>
</tr>
<tr>
<td>211.0</td>
<td>09:02:00.29</td>
<td>+00:31:14.8</td>
<td>V</td>
<td>1.6841</td>
<td>52.1</td>
<td>0</td>
</tr>
<tr>
<td>212.0</td>
<td>09:02:02.27</td>
<td>-01:12:04.6</td>
<td>V</td>
<td>0.8956</td>
<td>14.7</td>
<td>0</td>
</tr>
<tr>
<td>213.0</td>
<td>09:02:03.27</td>
<td>-00:49:15.9</td>
<td>V</td>
<td>2.836</td>
<td>129.3</td>
<td>0</td>
</tr>
<tr>
<td>217.0</td>
<td>09:02:06.71</td>
<td>+00:00:19.2</td>
<td>V</td>
<td>1.6182</td>
<td>225.8</td>
<td>0</td>
</tr>
<tr>
<td>219.0</td>
<td>09:02:09.66</td>
<td>-01:29:55.6</td>
<td>V</td>
<td>1.0935</td>
<td>184.4</td>
<td>0</td>
</tr>
<tr>
<td>221.0</td>
<td>09:02:11.22</td>
<td>-00:13:44.2</td>
<td>V</td>
<td>0.19627</td>
<td>19.3</td>
<td>1</td>
</tr>
<tr>
<td>226.0</td>
<td>09:02:14.88</td>
<td>+01:48:05.0</td>
<td>V</td>
<td>0.8728</td>
<td>17.6</td>
<td>0</td>
</tr>
<tr>
<td>233.0</td>
<td>09:02:20.31</td>
<td>-00:57:00.4</td>
<td>V</td>
<td>0.2386</td>
<td>14.3</td>
<td>0</td>
</tr>
<tr>
<td>236.0</td>
<td>09:02:22.80</td>
<td>-01:29:38.4</td>
<td>V</td>
<td>0.5279</td>
<td>13.6</td>
<td>0</td>
</tr>
<tr>
<td>243.0</td>
<td>09:02:30.77</td>
<td>-01:19:31.5</td>
<td>V</td>
<td>2.252</td>
<td>58.5</td>
<td>0</td>
</tr>
<tr>
<td>248.0</td>
<td>09:02:32.54</td>
<td>+01:00:18.9</td>
<td>V</td>
<td>2.0485</td>
<td>29.9</td>
<td>0</td>
</tr>
<tr>
<td>250.0</td>
<td>09:02:34.13</td>
<td>-01:11:25.0</td>
<td>F</td>
<td>2.5154</td>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td>251.0</td>
<td>09:02:34.10</td>
<td>-00:21:21.7</td>
<td>V</td>
<td>0.8499</td>
<td>103.4</td>
<td>0</td>
</tr>
<tr>
<td>253.0</td>
<td>09:02:35.67</td>
<td>-00:31:56.4</td>
<td>V</td>
<td>2.5942</td>
<td>21.9</td>
<td>0</td>
</tr>
</tbody>
</table>
## APPENDIX C. RADIO GALAXY CATALOGUE

<table>
<thead>
<tr>
<th>ID</th>
<th>RA</th>
<th>Dec</th>
<th>Position</th>
<th>$z$</th>
<th>1.4 GHz flux</th>
<th>Q flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>255.0</td>
<td>09:02:37.33</td>
<td>+01:01:36.1</td>
<td>V</td>
<td>0.6895</td>
<td>14.5</td>
<td>3</td>
</tr>
<tr>
<td>258.0</td>
<td>09:02:42.56</td>
<td>-00:32:16.0</td>
<td>V</td>
<td>0.8528</td>
<td>20.1</td>
<td>0</td>
</tr>
<tr>
<td>262.0</td>
<td>09:02:45.88</td>
<td>-01:08:29.1</td>
<td>V</td>
<td>1.193</td>
<td>24.3</td>
<td>0</td>
</tr>
<tr>
<td>263.0</td>
<td>09:02:46.24</td>
<td>+00:40:54.9</td>
<td>V</td>
<td>3.2936</td>
<td>20.8</td>
<td>0</td>
</tr>
<tr>
<td>273.0</td>
<td>09:02:57.38</td>
<td>+00:09:08.3</td>
<td>F</td>
<td>1.5761</td>
<td>56.7</td>
<td>0</td>
</tr>
<tr>
<td>275.0</td>
<td>09:02:57.93</td>
<td>-00:20:22.3</td>
<td>V</td>
<td>1.6696</td>
<td>15.7</td>
<td>0</td>
</tr>
<tr>
<td>278.0</td>
<td>09:03:02.96</td>
<td>+01:26:41.8</td>
<td>V</td>
<td>0.5385</td>
<td>26.9</td>
<td>0</td>
</tr>
<tr>
<td>280.0</td>
<td>09:03:07.86</td>
<td>+02:11:52.2</td>
<td>V</td>
<td>0.32924</td>
<td>24.2</td>
<td>1</td>
</tr>
<tr>
<td>284.0</td>
<td>09:03:10.15</td>
<td>-00:45:13.5</td>
<td>F</td>
<td>2.3225</td>
<td>18.5</td>
<td>0</td>
</tr>
<tr>
<td>291.0</td>
<td>09:03:17.23</td>
<td>-00:07:59.1</td>
<td>V</td>
<td>0.29734</td>
<td>11.5</td>
<td>1</td>
</tr>
<tr>
<td>293.1</td>
<td>09:03:18.06</td>
<td>-01:20:05.1</td>
<td>M</td>
<td>2.9952</td>
<td>13.3</td>
<td>0</td>
</tr>
<tr>
<td>298.0</td>
<td>09:03:22.48</td>
<td>-01:52:42.3</td>
<td>V</td>
<td>1.3117</td>
<td>16.7</td>
<td>0</td>
</tr>
<tr>
<td>306.0</td>
<td>09:03:31.59</td>
<td>+01:08:47.6</td>
<td>V</td>
<td>0.52224</td>
<td>63.1</td>
<td>1</td>
</tr>
<tr>
<td>310.0</td>
<td>09:03:33.32</td>
<td>-01:26:30.8</td>
<td>M</td>
<td>3.4707</td>
<td>13.1</td>
<td>0</td>
</tr>
<tr>
<td>311.0</td>
<td>09:03:33.51</td>
<td>-00:46:39.7</td>
<td>V</td>
<td>0.41426</td>
<td>51.7</td>
<td>1</td>
</tr>
<tr>
<td>313.0</td>
<td>09:03:35.32</td>
<td>-00:25:32.5</td>
<td>V</td>
<td>1.5995</td>
<td>13.7</td>
<td>0</td>
</tr>
<tr>
<td>314.0</td>
<td>09:03:36.80</td>
<td>+01:26:20.7</td>
<td>V</td>
<td>0.5949</td>
<td>13.6</td>
<td>0</td>
</tr>
<tr>
<td>317.0</td>
<td>09:03:41.86</td>
<td>+00:53:12.8</td>
<td>V</td>
<td>0.8701</td>
<td>20.6</td>
<td>0</td>
</tr>
<tr>
<td>319.0</td>
<td>09:03:42.40</td>
<td>+00:10:07.6</td>
<td>V</td>
<td>0.9627</td>
<td>69.5</td>
<td>0</td>
</tr>
<tr>
<td>323.0</td>
<td>09:03:44.96</td>
<td>+01:38:38.9</td>
<td>V</td>
<td>3.5504</td>
<td>56.4</td>
<td>0</td>
</tr>
<tr>
<td>328.0</td>
<td>09:03:48.13</td>
<td>-00:17:11.2</td>
<td>V</td>
<td>0.8857</td>
<td>68</td>
<td>0</td>
</tr>
<tr>
<td>329.0</td>
<td>09:03:48.38</td>
<td>+01:31:12.7</td>
<td>V</td>
<td>1.80619</td>
<td>12.9</td>
<td>2</td>
</tr>
<tr>
<td>332.0</td>
<td>09:03:50.89</td>
<td>-01:11:21.3</td>
<td>V</td>
<td>0.7567</td>
<td>23.7</td>
<td>3</td>
</tr>
<tr>
<td>337.0</td>
<td>09:03:55.08</td>
<td>-00:36:08.3</td>
<td>V</td>
<td>0.7992</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>338.0</td>
<td>09:03:56.38</td>
<td>-01:51:26.3</td>
<td>V</td>
<td>0.974</td>
<td>13.3</td>
<td>0</td>
</tr>
<tr>
<td>340.0</td>
<td>09:03:57.85</td>
<td>-01:23:55.5</td>
<td>V</td>
<td>0.8766</td>
<td>172.8</td>
<td>0</td>
</tr>
<tr>
<td>342.0</td>
<td>09:03:59.58</td>
<td>-01:40:33.9</td>
<td>V</td>
<td>2.1084</td>
<td>23.4</td>
<td>0</td>
</tr>
<tr>
<td>346.0</td>
<td>09:04:00.83</td>
<td>+00:38:19.7</td>
<td>V</td>
<td>0.8415</td>
<td>73.2</td>
<td>0</td>
</tr>
<tr>
<td>348.0</td>
<td>09:04:01.04</td>
<td>+01:27:29.3</td>
<td>V</td>
<td>0.05347</td>
<td>19.4</td>
<td>1</td>
</tr>
</tbody>
</table>
APPENDIX C. RADIO GALAXY CATALOGUE

<table>
<thead>
<tr>
<th>ID</th>
<th>RA</th>
<th>Dec</th>
<th>Position</th>
<th>z</th>
<th>1.4 GHz flux</th>
<th>Q flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>357.0</td>
<td>09:04:07.50</td>
<td>+00:24:10.0</td>
<td>V</td>
<td>4.6884</td>
<td>30.5</td>
<td>0</td>
</tr>
<tr>
<td>360.0</td>
<td>09:04:10.94</td>
<td>+00:26:05.4</td>
<td>V</td>
<td>0.3816</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>362.0</td>
<td>09:04:14.10</td>
<td>-00:21:44.9</td>
<td>V</td>
<td>0.35295</td>
<td>42.3</td>
<td>1</td>
</tr>
<tr>
<td>363.0</td>
<td>09:04:16.56</td>
<td>-00:52:46.1</td>
<td>M</td>
<td>3.4806</td>
<td>45.3</td>
<td>0</td>
</tr>
<tr>
<td>366.0</td>
<td>09:04:16.36</td>
<td>+01:13:35.8</td>
<td>V</td>
<td>0.2286</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>374.0</td>
<td>09:04:20.42</td>
<td>-01:06:14.5</td>
<td>F</td>
<td>1.3323</td>
<td>11.9</td>
<td>0</td>
</tr>
<tr>
<td>375.0</td>
<td>09:04:20.47</td>
<td>+02:18:35.5</td>
<td>V</td>
<td>0.45407</td>
<td>11.1</td>
<td>1</td>
</tr>
<tr>
<td>376.0</td>
<td>09:04:20.60</td>
<td>+00:47:00.8</td>
<td>V</td>
<td>2.3196</td>
<td>17.6</td>
<td>0</td>
</tr>
<tr>
<td>378.0</td>
<td>09:04:21.87</td>
<td>-01:32:06.6</td>
<td>V</td>
<td>0.8884</td>
<td>18.9</td>
<td>0</td>
</tr>
<tr>
<td>385.0</td>
<td>09:04:25.60</td>
<td>+01:20:15.9</td>
<td>V</td>
<td>0.7525</td>
<td>76</td>
<td>0</td>
</tr>
<tr>
<td>387.0</td>
<td>09:04:28.55</td>
<td>+00:33:07.2</td>
<td>N</td>
<td>1.1889</td>
<td>33.6</td>
<td>0</td>
</tr>
<tr>
<td>392.0</td>
<td>09:04:31.04</td>
<td>+01:00:47.0</td>
<td>V</td>
<td>0.31391</td>
<td>25.7</td>
<td>1</td>
</tr>
<tr>
<td>396.0</td>
<td>09:04:35.45</td>
<td>-01:32:29.4</td>
<td>V</td>
<td>1.8162</td>
<td>103.2</td>
<td>0</td>
</tr>
<tr>
<td>398.0</td>
<td>09:04:36.61</td>
<td>-01:46:52.9</td>
<td>V</td>
<td>1.2696</td>
<td>36.2</td>
<td>0</td>
</tr>
<tr>
<td>400.0</td>
<td>09:04:40.18</td>
<td>-01:38:43.9</td>
<td>V</td>
<td>3.6577</td>
<td>98.7</td>
<td>0</td>
</tr>
<tr>
<td>405.0</td>
<td>09:04:46.00</td>
<td>+02:08:43.0</td>
<td>V</td>
<td>0.8827</td>
<td>329.6</td>
<td>0</td>
</tr>
<tr>
<td>410.0</td>
<td>09:04:48.72</td>
<td>+02:09:53.0</td>
<td>F</td>
<td>1.8872</td>
<td>15.4</td>
<td>0</td>
</tr>
<tr>
<td>415.0</td>
<td>09:04:56.26</td>
<td>-00:44:41.5</td>
<td>V</td>
<td>2.0076</td>
<td>84</td>
<td>0</td>
</tr>
<tr>
<td>429.0</td>
<td>09:05:09.10</td>
<td>-01:29:32.3</td>
<td>V</td>
<td>0.8467</td>
<td>16.1</td>
<td>0</td>
</tr>
<tr>
<td>433.0</td>
<td>09:05:11.09</td>
<td>-01:02:19.8</td>
<td>V</td>
<td>2.6172</td>
<td>37.8</td>
<td>0</td>
</tr>
<tr>
<td>439.0</td>
<td>09:05:20.16</td>
<td>+01:31:03.7</td>
<td>V</td>
<td>0.6866</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>444.0</td>
<td>09:05:25.59</td>
<td>+02:09:53.9</td>
<td>V</td>
<td>4.2882</td>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>456.0</td>
<td>09:05:37.22</td>
<td>+00:51:42.3</td>
<td>M</td>
<td>0.5836</td>
<td>33.8</td>
<td>0</td>
</tr>
<tr>
<td>457.0</td>
<td>09:05:38.52</td>
<td>+00:42:37.0</td>
<td>M</td>
<td>1.8359</td>
<td>37.5</td>
<td>0</td>
</tr>
<tr>
<td>458.0</td>
<td>09:05:37.14</td>
<td>-01:01:27.5</td>
<td>V</td>
<td>1.0614</td>
<td>14.4</td>
<td>0</td>
</tr>
<tr>
<td>462.0</td>
<td>09:05:39.58</td>
<td>-01:05:15.9</td>
<td>F</td>
<td>3.8426</td>
<td>10.4</td>
<td>0</td>
</tr>
<tr>
<td>466.0</td>
<td>09:05:42.20</td>
<td>-01:18:50.8</td>
<td>V</td>
<td>0.8615</td>
<td>14.7</td>
<td>0</td>
</tr>
<tr>
<td>467.0</td>
<td>09:05:43.27</td>
<td>+01:52:07.4</td>
<td>F</td>
<td>4.9613</td>
<td>13.8</td>
<td>0</td>
</tr>
<tr>
<td>469.0</td>
<td>09:05:46.99</td>
<td>-01:12:43.7</td>
<td>V</td>
<td>0.9878</td>
<td>43.3</td>
<td>0</td>
</tr>
</tbody>
</table>
APPENDIX C. RADIO GALAXY CATALOGUE

<table>
<thead>
<tr>
<th>ID</th>
<th>RA</th>
<th>Dec</th>
<th>Position</th>
<th>z</th>
<th>1.4 GHz flux</th>
<th>Q flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>470.0</td>
<td>09:05:46.94</td>
<td>-00:27:15.7</td>
<td>V</td>
<td>0.7738</td>
<td>25.2</td>
<td>0</td>
</tr>
<tr>
<td>472.0</td>
<td>09:05:48.35</td>
<td>+00:17:23.3</td>
<td>V</td>
<td>4.7543</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>485.0</td>
<td>09:05:58.80</td>
<td>+00:20:20.6</td>
<td>F</td>
<td>0.8706</td>
<td>15.7</td>
<td>0</td>
</tr>
<tr>
<td>489.0</td>
<td>09:06:01.15</td>
<td>-01:04:11.4</td>
<td>V</td>
<td>2.574</td>
<td>13.1</td>
<td>0</td>
</tr>
<tr>
<td>492.0</td>
<td>09:06:05.58</td>
<td>-00:04:12.1</td>
<td>V</td>
<td>2.0782</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>494.0</td>
<td>09:06:07.10</td>
<td>-00:00:36.2</td>
<td>V</td>
<td>1.2044</td>
<td>28.3</td>
<td>0</td>
</tr>
<tr>
<td>498.0</td>
<td>09:06:11.19</td>
<td>+01:19:55.3</td>
<td>V</td>
<td>0.60587</td>
<td>10.8</td>
<td>1</td>
</tr>
<tr>
<td>499.0</td>
<td>09:06:12.16</td>
<td>-00:36:30.9</td>
<td>V</td>
<td>1.5523</td>
<td>18.5</td>
<td>0</td>
</tr>
<tr>
<td>502.0</td>
<td>09:06:14.79</td>
<td>+01:54:02.2</td>
<td>V</td>
<td>0.31301</td>
<td>11.3</td>
<td>1</td>
</tr>
<tr>
<td>521.0</td>
<td>09:06:27.49</td>
<td>+00:52:50.4</td>
<td>V</td>
<td>2.74007</td>
<td>16.8</td>
<td>2</td>
</tr>
<tr>
<td>523.0</td>
<td>09:06:28.63</td>
<td>-01:27:43.8</td>
<td>V</td>
<td>4.5569</td>
<td>14.9</td>
<td>0</td>
</tr>
<tr>
<td>525.0</td>
<td>09:06:31.14</td>
<td>+01:08:25.7</td>
<td>V</td>
<td>2.5312</td>
<td>28.1</td>
<td>0</td>
</tr>
<tr>
<td>535.0</td>
<td>09:06:39.65</td>
<td>+00:02:35.4</td>
<td>F</td>
<td>0.32</td>
<td>13.6</td>
<td>3</td>
</tr>
<tr>
<td>537.0</td>
<td>09:06:41.90</td>
<td>+02:14:50.3</td>
<td>F</td>
<td>0.5712</td>
<td>75.3</td>
<td>0</td>
</tr>
<tr>
<td>539.0</td>
<td>09:06:43.58</td>
<td>-00:36:28.2</td>
<td>M</td>
<td>0.8579</td>
<td>37.5</td>
<td>0</td>
</tr>
<tr>
<td>542.0</td>
<td>09:06:44.27</td>
<td>-00:23:23.1</td>
<td>V</td>
<td>0.9816</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>544.0</td>
<td>09:06:47.88</td>
<td>-00:43:51.3</td>
<td>V</td>
<td>0.6751</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>545.0</td>
<td>09:06:48.34</td>
<td>+02:29:17.8</td>
<td>V</td>
<td>1.7752</td>
<td>18.7</td>
<td>0</td>
</tr>
<tr>
<td>548.0</td>
<td>09:06:52.23</td>
<td>+01:11:24.9</td>
<td>V</td>
<td>0.7857</td>
<td>15.4</td>
<td>0</td>
</tr>
<tr>
<td>550.0</td>
<td>09:06:54.00</td>
<td>+01:21:26.1</td>
<td>F</td>
<td>0.778</td>
<td>18.8</td>
<td>0</td>
</tr>
<tr>
<td>551.0</td>
<td>09:06:54.52</td>
<td>-01:20:31.1</td>
<td>V</td>
<td>2.6169</td>
<td>27.7</td>
<td>0</td>
</tr>
<tr>
<td>554.0</td>
<td>09:06:55.88</td>
<td>+00:14:46.5</td>
<td>V</td>
<td>1.1444</td>
<td>63.8</td>
<td>3</td>
</tr>
<tr>
<td>556.0</td>
<td>09:06:56.41</td>
<td>+02:35:02.4</td>
<td>V</td>
<td>1.3911</td>
<td>18.9</td>
<td>0</td>
</tr>
<tr>
<td>557.0</td>
<td>09:06:56.72</td>
<td>+00:31:24.1</td>
<td>V</td>
<td>0.5731</td>
<td>24.2</td>
<td>0</td>
</tr>
<tr>
<td>559.0</td>
<td>09:06:57.36</td>
<td>-00:42:39.0</td>
<td>V</td>
<td>1.2122</td>
<td>33.8</td>
<td>0</td>
</tr>
<tr>
<td>564.0</td>
<td>09:07:01.58</td>
<td>+01:17:45.8</td>
<td>V</td>
<td>1.17092</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>565.0</td>
<td>09:07:04.31</td>
<td>+02:34:39.1</td>
<td>V</td>
<td>1.0575</td>
<td>34.6</td>
<td>0</td>
</tr>
<tr>
<td>568.0</td>
<td>09:07:05.70</td>
<td>+02:13:23.1</td>
<td>N</td>
<td>3.4949</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>575.0</td>
<td>09:07:12.95</td>
<td>-00:53:31.5</td>
<td>V</td>
<td>1.1415</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>RA</td>
<td>Dec</td>
<td>Position</td>
<td>z</td>
<td>1.4 GHz flux</td>
<td>Q flag</td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td>------------</td>
<td>----------</td>
<td>---------</td>
<td>--------------</td>
<td>--------</td>
</tr>
<tr>
<td>579.0</td>
<td>09:07:20.72</td>
<td>+01:04:58.0</td>
<td>V</td>
<td>0.60243</td>
<td>10.4</td>
<td>1</td>
</tr>
<tr>
<td>581.0</td>
<td>09:07:21.10</td>
<td>-00:14:24.2</td>
<td>V</td>
<td>0.4941</td>
<td>10.3</td>
<td>0</td>
</tr>
<tr>
<td>589.0</td>
<td>09:07:29.85</td>
<td>+01:50:22.9</td>
<td>V</td>
<td>0.5879</td>
<td>13.2</td>
<td>0</td>
</tr>
<tr>
<td>590.0</td>
<td>09:07:30.20</td>
<td>+00:55:18.9</td>
<td>F</td>
<td>1.2154</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>595.0</td>
<td>09:07:34.76</td>
<td>+01:25:04.4</td>
<td>V</td>
<td>0.10243</td>
<td>12.8</td>
<td>1</td>
</tr>
<tr>
<td>597.0</td>
<td>09:07:36.56</td>
<td>+00:43:41.5</td>
<td>V</td>
<td>1.227</td>
<td>10.6</td>
<td>0</td>
</tr>
<tr>
<td>613.0</td>
<td>09:07:52.34</td>
<td>+02:47:52.9</td>
<td>V</td>
<td>2.1173</td>
<td>12.9</td>
<td>0</td>
</tr>
<tr>
<td>623.0</td>
<td>09:08:06.27</td>
<td>+01:19:55.5</td>
<td>V</td>
<td>0.65956</td>
<td>67.1</td>
<td>1</td>
</tr>
<tr>
<td>625.0</td>
<td>09:08:07.54</td>
<td>+02:06:10.5</td>
<td>V</td>
<td>1.6208</td>
<td>17.9</td>
<td>0</td>
</tr>
<tr>
<td>626.0</td>
<td>09:08:08.81</td>
<td>+00:35:00.6</td>
<td>V</td>
<td>1.705</td>
<td>13.7</td>
<td>3</td>
</tr>
<tr>
<td>630.1</td>
<td>09:08:17.17</td>
<td>-01:17:37.0</td>
<td>F</td>
<td>0.3663</td>
<td>280.6</td>
<td>0</td>
</tr>
<tr>
<td>632.0</td>
<td>09:08:18.20</td>
<td>-00:54:04.9</td>
<td>V</td>
<td>1.518</td>
<td>128.3</td>
<td>0</td>
</tr>
<tr>
<td>639.0</td>
<td>09:08:21.98</td>
<td>+00:33:00.9</td>
<td>F</td>
<td>3.4668</td>
<td>32.3</td>
<td>0</td>
</tr>
<tr>
<td>644.0</td>
<td>09:08:27.48</td>
<td>+02:02:41.0</td>
<td>V</td>
<td>1.3634</td>
<td>26.9</td>
<td>0</td>
</tr>
<tr>
<td>651.0</td>
<td>09:08:33.97</td>
<td>+00:57:15.6</td>
<td>V</td>
<td>0.73914</td>
<td>21.2</td>
<td>1</td>
</tr>
<tr>
<td>653.0</td>
<td>09:08:34.50</td>
<td>-01:02:31.9</td>
<td>V</td>
<td>0.3624</td>
<td>12.6</td>
<td>3</td>
</tr>
<tr>
<td>655.0</td>
<td>09:08:35.43</td>
<td>+01:55:16.5</td>
<td>V</td>
<td>0.6747</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>657.0</td>
<td>09:08:37.20</td>
<td>-01:09:00.8</td>
<td>V</td>
<td>1.0997</td>
<td>47.5</td>
<td>0</td>
</tr>
<tr>
<td>665.0</td>
<td>09:08:42.78</td>
<td>-01:23:27.0</td>
<td>V</td>
<td>1.9748</td>
<td>175.6</td>
<td>0</td>
</tr>
<tr>
<td>667.0</td>
<td>09:08:44.99</td>
<td>+01:59:45.7</td>
<td>V</td>
<td>2.0744</td>
<td>18.5</td>
<td>0</td>
</tr>
<tr>
<td>669.0</td>
<td>09:08:49.56</td>
<td>-01:03:47.4</td>
<td>V</td>
<td>2.4061</td>
<td>17.2</td>
<td>3</td>
</tr>
<tr>
<td>677.0</td>
<td>09:08:59.45</td>
<td>+00:27:26.5</td>
<td>V</td>
<td>0.7668</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>685.0</td>
<td>09:09:06.55</td>
<td>-01:02:26.4</td>
<td>F</td>
<td>0.5859</td>
<td>28.7</td>
<td>0</td>
</tr>
<tr>
<td>687.0</td>
<td>09:09:08.10</td>
<td>+00:54:40.5</td>
<td>V</td>
<td>1.2147</td>
<td>570.4</td>
<td>0</td>
</tr>
<tr>
<td>690.0</td>
<td>09:09:10.09</td>
<td>+01:21:35.6</td>
<td>V</td>
<td>1.02554</td>
<td>708</td>
<td>2</td>
</tr>
<tr>
<td>690.1</td>
<td>09:09:11.00</td>
<td>+01:21:36.5</td>
<td>V</td>
<td>1.5926</td>
<td>51.7</td>
<td>3</td>
</tr>
<tr>
<td>693.0</td>
<td>09:09:11.20</td>
<td>+00:23:54.6</td>
<td>V</td>
<td>0.06965</td>
<td>13.1</td>
<td>1</td>
</tr>
<tr>
<td>695.0</td>
<td>09:09:13.10</td>
<td>-00:59:36.0</td>
<td>V</td>
<td>1.5905</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>697.0</td>
<td>09:09:13.79</td>
<td>-00:12:20.2</td>
<td>V</td>
<td>0.30853</td>
<td>75.5</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix C. Radio Galaxy Catalogue

<table>
<thead>
<tr>
<th>ID</th>
<th>RA</th>
<th>Dec</th>
<th>Position</th>
<th>z</th>
<th>1.4 GHz flux</th>
<th>Q flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>701.0</td>
<td>09:09:19.75</td>
<td>+00:00:25.4</td>
<td>V</td>
<td>0.28054</td>
<td>17.6</td>
<td>1</td>
</tr>
<tr>
<td>707.0</td>
<td>09:09:27.55</td>
<td>-01:10:52.9</td>
<td>V</td>
<td>1.6701</td>
<td>48.6</td>
<td>0</td>
</tr>
<tr>
<td>713.0</td>
<td>09:09:31.55</td>
<td>-01:12:33.5</td>
<td>V</td>
<td>0.4216</td>
<td>21.5</td>
<td>0</td>
</tr>
<tr>
<td>714.0</td>
<td>09:09:31.03</td>
<td>+00:22:22.5</td>
<td>V</td>
<td>0.9199</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>715.0</td>
<td>09:09:31.62</td>
<td>+02:14:57.6</td>
<td>N</td>
<td>2.4673</td>
<td>37</td>
<td>0</td>
</tr>
<tr>
<td>716.0</td>
<td>09:09:37.71</td>
<td>+01:19:41.6</td>
<td>V</td>
<td>0.8848</td>
<td>132.8</td>
<td>0</td>
</tr>
<tr>
<td>718.0</td>
<td>09:09:37.47</td>
<td>-00:27:55.1</td>
<td>V</td>
<td>3.2187</td>
<td>11.4</td>
<td>0</td>
</tr>
<tr>
<td>721.0</td>
<td>09:09:39.25</td>
<td>-00:40:17.8</td>
<td>V</td>
<td>0.5385</td>
<td>15.2</td>
<td>0</td>
</tr>
<tr>
<td>722.0</td>
<td>09:09:39.43</td>
<td>-00:55:56.0</td>
<td>V</td>
<td>1.8584</td>
<td>12.2</td>
<td>0</td>
</tr>
<tr>
<td>723.0</td>
<td>09:09:39.86</td>
<td>+02:00:05.4</td>
<td>V</td>
<td>0.7405</td>
<td>156.6</td>
<td>0</td>
</tr>
<tr>
<td>725.0</td>
<td>09:09:41.99</td>
<td>+00:47:12.7</td>
<td>V</td>
<td>0.5648</td>
<td>363</td>
<td>0</td>
</tr>
<tr>
<td>726.0</td>
<td>09:09:42.22</td>
<td>+00:04:50.2</td>
<td>F</td>
<td>0.9896</td>
<td>13.9</td>
<td>0</td>
</tr>
<tr>
<td>731.0</td>
<td>09:09:47.66</td>
<td>-00:56:16.0</td>
<td>F</td>
<td>1.3598</td>
<td>18.24</td>
<td>0</td>
</tr>
<tr>
<td>733.0</td>
<td>09:09:49.02</td>
<td>+01:57:18.3</td>
<td>V</td>
<td>0.7147</td>
<td>15.7</td>
<td>1</td>
</tr>
<tr>
<td>737.0</td>
<td>09:09:53.16</td>
<td>+00:36:06.2</td>
<td>V</td>
<td>1.0667</td>
<td>68.1</td>
<td>0</td>
</tr>
<tr>
<td>738.0</td>
<td>09:09:53.37</td>
<td>-00:12:01.6</td>
<td>V</td>
<td>3.7012</td>
<td>49.2</td>
<td>0</td>
</tr>
<tr>
<td>743.0</td>
<td>09:09:57.08</td>
<td>+01:51:55.3</td>
<td>V</td>
<td>0.28222</td>
<td>15.8</td>
<td>1</td>
</tr>
<tr>
<td>744.0</td>
<td>09:09:57.48</td>
<td>-00:39:40.1</td>
<td>F</td>
<td>0.9826</td>
<td>34.4</td>
<td>0</td>
</tr>
<tr>
<td>745.0</td>
<td>09:09:59.47</td>
<td>-00:55:51.7</td>
<td>V</td>
<td>0.6967</td>
<td>10.6</td>
<td>0</td>
</tr>
<tr>
<td>751.0</td>
<td>09:10:06.38</td>
<td>+01:51:50.1</td>
<td>V</td>
<td>1.9117</td>
<td>17.2</td>
<td>0</td>
</tr>
<tr>
<td>753.0</td>
<td>09:10:07.55</td>
<td>-01:11:48.3</td>
<td>F</td>
<td>0.5393</td>
<td>37.1</td>
<td>0</td>
</tr>
<tr>
<td>757.0</td>
<td>09:10:11.55</td>
<td>+01:39:03.9</td>
<td>V</td>
<td>0.28411</td>
<td>110.4</td>
<td>1</td>
</tr>
<tr>
<td>758.0</td>
<td>09:10:11.27</td>
<td>+01:01:39.5</td>
<td>V</td>
<td>1.0219</td>
<td>225.9</td>
<td>0</td>
</tr>
<tr>
<td>763.0</td>
<td>09:10:18.23</td>
<td>-00:42:05.5</td>
<td>V</td>
<td>1.1449</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>767.0</td>
<td>09:10:22.88</td>
<td>+00:19:35.3</td>
<td>V</td>
<td>0.7855</td>
<td>318.9</td>
<td>0</td>
</tr>
<tr>
<td>768.0</td>
<td>09:10:23.56</td>
<td>+00:15:41.9</td>
<td>V</td>
<td>2.1329</td>
<td>18.3</td>
<td>0</td>
</tr>
<tr>
<td>771.0</td>
<td>09:10:27.25</td>
<td>-00:26:09.1</td>
<td>V</td>
<td>1.0498</td>
<td>34.5</td>
<td>0</td>
</tr>
<tr>
<td>772.0</td>
<td>09:10:27.54</td>
<td>-00:44:24.5</td>
<td>V</td>
<td>1.088</td>
<td>119.9</td>
<td>0</td>
</tr>
<tr>
<td>777.0</td>
<td>09:10:34.04</td>
<td>+01:54:27.9</td>
<td>V</td>
<td>1.0321</td>
<td>21.6</td>
<td>1</td>
</tr>
<tr>
<td>ID</td>
<td>RA (°)</td>
<td>Dec (°)</td>
<td>Position</td>
<td>z</td>
<td>1.4 GHz flux</td>
<td>Q flag</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>-----</td>
<td>--------------</td>
<td>--------</td>
</tr>
<tr>
<td>782.0</td>
<td>09:10:39.12</td>
<td>+00:03:13.5</td>
<td>F</td>
<td>4.3943</td>
<td>23.2</td>
<td>0</td>
</tr>
<tr>
<td>783.0</td>
<td>09:10:39.54</td>
<td>-00:58:15.2</td>
<td>V</td>
<td>4.0105</td>
<td>80.1</td>
<td>0</td>
</tr>
<tr>
<td>787.0</td>
<td>09:10:45.60</td>
<td>-00:07:07.0</td>
<td>M</td>
<td>4.5911</td>
<td>13.1</td>
<td>0</td>
</tr>
<tr>
<td>790.0</td>
<td>09:10:50.79</td>
<td>+01:22:12.9</td>
<td>V</td>
<td>0.33381</td>
<td>51.5</td>
<td>1</td>
</tr>
<tr>
<td>796.0</td>
<td>09:10:59.24</td>
<td>-01:07:24.7</td>
<td>V</td>
<td>0.08</td>
<td>12.1</td>
<td>0</td>
</tr>
<tr>
<td>799.0</td>
<td>09:11:05.68</td>
<td>+01:36:22.8</td>
<td>V</td>
<td>0.49185</td>
<td>13.2</td>
<td>1</td>
</tr>
<tr>
<td>803.0</td>
<td>09:11:14.98</td>
<td>+01:33:53.9</td>
<td>V</td>
<td>0.3135</td>
<td>15.8</td>
<td>1</td>
</tr>
<tr>
<td>808.0</td>
<td>09:11:18.56</td>
<td>-00:21:00.3</td>
<td>N</td>
<td>3.5394</td>
<td>23.9</td>
<td>0</td>
</tr>
<tr>
<td>809.0</td>
<td>09:11:19.65</td>
<td>+01:10:09.6</td>
<td>V</td>
<td>0.71285</td>
<td>40.5</td>
<td>1</td>
</tr>
<tr>
<td>811.0</td>
<td>09:11:22.89</td>
<td>+00:23:53.3</td>
<td>V</td>
<td>0.16364</td>
<td>12.1</td>
<td>1</td>
</tr>
<tr>
<td>812.0</td>
<td>09:11:24.02</td>
<td>+01:49:42.3</td>
<td>V</td>
<td>0.9834</td>
<td>13.5</td>
<td>0</td>
</tr>
<tr>
<td>820.0</td>
<td>09:11:32.76</td>
<td>-01:04:47.3</td>
<td>V</td>
<td>1.943</td>
<td>17.9</td>
<td>0</td>
</tr>
<tr>
<td>825.0</td>
<td>09:11:40.46</td>
<td>+01:03:09.0</td>
<td>V</td>
<td>2.6449</td>
<td>45.1</td>
<td>0</td>
</tr>
<tr>
<td>829.0</td>
<td>09:11:41.41</td>
<td>+00:11:02.0</td>
<td>V</td>
<td>0.7514</td>
<td>479.5</td>
<td>1</td>
</tr>
<tr>
<td>832.0</td>
<td>09:11:46.85</td>
<td>-00:14:45.5</td>
<td>V</td>
<td>0.24617</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>833.0</td>
<td>09:11:46.75</td>
<td>-01:00:02.8</td>
<td>V</td>
<td>0.6871</td>
<td>145.2</td>
<td>0</td>
</tr>
<tr>
<td>839.0</td>
<td>09:11:54.35</td>
<td>+01:10:41.6</td>
<td>V</td>
<td>0.1587</td>
<td>28.2</td>
<td>0</td>
</tr>
<tr>
<td>840.0</td>
<td>09:11:54.83</td>
<td>+01:08:55.5</td>
<td>V</td>
<td>0.16397</td>
<td>21.6</td>
<td>1</td>
</tr>
<tr>
<td>851.0</td>
<td>09:12:09.60</td>
<td>-00:09:54.9</td>
<td>V</td>
<td>2.26538</td>
<td>37.6</td>
<td>2</td>
</tr>
<tr>
<td>852.0</td>
<td>09:12:10.66</td>
<td>+01:16:49.1</td>
<td>V</td>
<td>0.38593</td>
<td>11.8</td>
<td>1</td>
</tr>
<tr>
<td>860.0</td>
<td>09:12:17.46</td>
<td>-01:02:32.9</td>
<td>V</td>
<td>0.1547</td>
<td>60.7</td>
<td>0</td>
</tr>
<tr>
<td>861.0</td>
<td>09:12:18.78</td>
<td>-00:04:01.1</td>
<td>V</td>
<td>4.7436</td>
<td>15.6</td>
<td>0</td>
</tr>
<tr>
<td>862.0</td>
<td>09:12:19.28</td>
<td>-00:35:46.5</td>
<td>V</td>
<td>1.0262</td>
<td>39.7</td>
<td>0</td>
</tr>
<tr>
<td>865.0</td>
<td>09:12:23.09</td>
<td>+01:17:52.8</td>
<td>V</td>
<td>1.0898</td>
<td>50.4</td>
<td>0</td>
</tr>
<tr>
<td>867.0</td>
<td>09:12:25.20</td>
<td>-00:53:42.7</td>
<td>F</td>
<td>0.9889</td>
<td>23.3</td>
<td>0</td>
</tr>
<tr>
<td>872.0</td>
<td>09:12:31.37</td>
<td>-00:38:18.0</td>
<td>V</td>
<td>0.5798</td>
<td>56.6</td>
<td>0</td>
</tr>
<tr>
<td>874.0</td>
<td>09:12:32.96</td>
<td>-00:50:21.0</td>
<td>V</td>
<td>0.47609</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>876.0</td>
<td>09:12:33.23</td>
<td>+01:12:41.2</td>
<td>V</td>
<td>2.2125</td>
<td>53.6</td>
<td>0</td>
</tr>
<tr>
<td>879.0</td>
<td>09:12:38.53</td>
<td>+00:26:44.7</td>
<td>V</td>
<td>0.40479</td>
<td>25.6</td>
<td>2</td>
</tr>
</tbody>
</table>
## APPENDIX C. RADIO GALAXY CATALOGUE

<table>
<thead>
<tr>
<th>ID</th>
<th>RA</th>
<th>Dec</th>
<th>Position</th>
<th>z</th>
<th>1.4 GHz flux</th>
<th>Q flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>886.0</td>
<td>09:12:48.85</td>
<td>+00:33:45.7</td>
<td>V</td>
<td>1.6682</td>
<td>15.3</td>
<td>0</td>
</tr>
<tr>
<td>889.0</td>
<td>09:12:55.98</td>
<td>-00:40:09.8</td>
<td>V</td>
<td>0.82866</td>
<td>29.2</td>
<td>2</td>
</tr>
<tr>
<td>893.0</td>
<td>09:13:06.62</td>
<td>-00:46:46.1</td>
<td>V</td>
<td>0.25481</td>
<td>12.6</td>
<td>1</td>
</tr>
<tr>
<td>896.0</td>
<td>09:13:15.67</td>
<td>+00:44:44.5</td>
<td>V</td>
<td>0.22502</td>
<td>44.7</td>
<td>1</td>
</tr>
<tr>
<td>905.0</td>
<td>09:13:25.06</td>
<td>+00:30:40.9</td>
<td>V</td>
<td>0.6686</td>
<td>30.6</td>
<td>0</td>
</tr>
<tr>
<td>912.0</td>
<td>09:13:33.67</td>
<td>-00:42:51.0</td>
<td>V</td>
<td>0.42625</td>
<td>106.2</td>
<td>2</td>
</tr>
<tr>
<td>924.0</td>
<td>09:13:48.80</td>
<td>-00:39:22.0</td>
<td>V</td>
<td>0.9235</td>
<td>10.9</td>
<td>0</td>
</tr>
<tr>
<td>925.0</td>
<td>09:13:49.98</td>
<td>+00:16:32.2</td>
<td>V</td>
<td>0.4849</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>926.0</td>
<td>09:13:50.71</td>
<td>+00:34:02.2</td>
<td>V</td>
<td>0.38212</td>
<td>26.4</td>
<td>1</td>
</tr>
<tr>
<td>927.0</td>
<td>09:13:52.18</td>
<td>-00:45:35.0</td>
<td>V</td>
<td>0.29501</td>
<td>15.2</td>
<td>1</td>
</tr>
<tr>
<td>931.0</td>
<td>09:13:59.13</td>
<td>+00:13:15.9</td>
<td>V</td>
<td>1.043</td>
<td>30.1</td>
<td>0</td>
</tr>
<tr>
<td>935.0</td>
<td>09:14:03.12</td>
<td>+00:07:46.8</td>
<td>F</td>
<td>4.1846</td>
<td>38.1</td>
<td>0</td>
</tr>
<tr>
<td>936.0</td>
<td>09:14:05.70</td>
<td>-00:22:22.8</td>
<td>N</td>
<td>0.973</td>
<td>11.2</td>
<td>0</td>
</tr>
<tr>
<td>940.0</td>
<td>09:14:08.88</td>
<td>-00:44:54.6</td>
<td>V</td>
<td>1.6288</td>
<td>18.9</td>
<td>0</td>
</tr>
<tr>
<td>947.0</td>
<td>09:14:17.59</td>
<td>-00:09:46.9</td>
<td>V</td>
<td>1.3399</td>
<td>29.9</td>
<td>0</td>
</tr>
<tr>
<td>953.0</td>
<td>09:14:24.43</td>
<td>-00:29:24.8</td>
<td>V</td>
<td>0.6424</td>
<td>27.7</td>
<td>0</td>
</tr>
<tr>
<td>965.0</td>
<td>09:14:46.13</td>
<td>-00:24:06.9</td>
<td>V</td>
<td>1.8283</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>969.0</td>
<td>09:14:56.97</td>
<td>-00:22:40.9</td>
<td>V</td>
<td>0.4591</td>
<td>11.5</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix D

Images of radio galaxies

This appendix shows image cutouts for all radio galaxies examined in Chapter 5 and presented in the source catalogue in Appendix C. Each image is labelled with an ID, corresponding the value in the first column of the catalogue.

The background of each cutout shows the $K_s$ band VIKING image. The colour scale is in counts and ranges from the maximum to minimum of each image. The majority of cutouts are $1 \times 1$ arcmin$^2$ in size, except for very extended sources which are $3 \times 3$ arcmin$^2$.

NVSS 1.4 GHz flux is indicated by the black contours. They are shown at the 0.001, 0.0013, 0.0018, 0.0025, 0.0028, 0.003, 0.0035, 0.004, 0.005, 0.006, 0.0075, 0.01, 0.015, 0.025, 0.045, 0.075, 0.12, 0.2, 0.35, 0.6, 0.9, 1.5 and 2.3 Jy levels. The FIRST 1.4 GHz flux is indicated by the magenta contours and are shown at the 0.0003, 0.0004, 0.0005, 0.0007, 0.001, 0.0015, 0.002, 0.0025, 0.0035, 0.005, 0.007, 0.01, 0.015, 0.025, 0.045, 0.075, 0.12, 0.2, 0.35, 0.6, 0.9, 1.5 and 2.3 Jy levels.

Red plus symbols and blue crosses show NVSS and FIRST source catalogue positions, respectively. Green circles show detections in the VIKING catalogue. Note that some are clearly false detections caused by artefacts or very strong, extended sources in the VIKING data (such as in cutout ID 68). Black diamonds indicate the VIKING
APPENDIX D. IMAGES OF RADIO GALAXIES

detection which is most likely the near-infrared counterpart to the radio source (based on source position proximity). Orange squares show any manually-defined positions. See Chapter 5 for an explanation as to why manual positions are required in some cases. The value of ‘Centre’, at the top of each image, indicates which of these positions are recorded in the final source catalogue. These follow the key given in Appendix C.
APPENDIX D. IMAGES OF RADIO GALAXIES
APPENDIX D. IMAGES OF RADIO GALAXIES
APPENDIX D. IMAGES OF RADIO GALAXIES
APPENDIX D. IMAGES OF RADIO GALAXIES

ID : 502, Centre : V

ID : 494, Centre : F

ID : 525, Centre : V

ID : 535, Centre : V

ID : 539, Centre : V

ID : 542, Centre : V

ID : 498, Centre : V
APPENDIX D. IMAGES OF RADIO GALAXIES
APPENDIX D. IMAGES OF RADIO GALAXIES

ID : 651
ID : 669
ID : 632
ID : 653
ID : 665
ID : 685

237
APPENDIX D. IMAGES OF RADIO GALAXIES
APPENDIX D. IMAGES OF RADIO GALAXIES

ID, Name, Corp., V

RA

Dec

ID, Name, Corp., V

RA

Dec

ID, Name, Corp., V

RA

Dec

ID, Name, Corp., V

RA

Dec
APPENDIX D. IMAGES OF RADIO GALAXIES