

UNIVERSITY *of*  
TASMANIA  

---

AUSTRALIA

THIS PAGE INTENTIONALLY LEFT BLANK

# Multi-frequency Matching, Classification, and Cosmic Evolution of Radio Galaxy Populations

By

Jesse Allen Swan

*Bachelor of Science with First Class Honours  
University of Tasmania, 2013*

A THESIS SUBMITTED IN PARTIAL SATISFACTION  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY  
IN  
ASTROPHYSICS  
IN THE  
SCHOOL OF NATURAL SCIENCES  
OF THE  
UNIVERSITY OF TASMANIA  
HOBART, AUSTRALIA

SUPERVISING COMMITTEE:

DOCTOR STANISLAV SHABALA  
PROFESSOR RAY NORRIS  
ASSOCIATE PROFESSOR ANDREW COLE

OCTOBER 2018



UNIVERSITY of  
TASMANIA

©2018 – JESSE ALLEN SWAN  
ALL RIGHTS RESERVED.

## DECLARATION

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in this thesis, and to the best of the my knowledge and belief no material previously published or written by another person except where due acknowledgment is made in the text of the thesis.

Signed:

Jesse Allen Swan  
Sunday 28<sup>th</sup> October, 2018

THIS PAGE INTENTIONALLY LEFT BLANK

# Multi-frequency Matching, Classification, and Cosmic Evolution of Radio Galaxy Populations

## ABSTRACT

This thesis combines the third data release of the Australia Telescope Large Area Survey (ATLAS) with complimentary DES photometry and OzDES targeted spectroscopy. ATLAS 1.4 GHz is a deep ( $< 20 \mu\text{Jy}/\text{beam}$ ) survey of the combined  $6.3 \text{ deg}^2$  sky area of CDFS and ELAIS-S1.

Cross-matching of 4903 ATLAS sources to photometric and spectroscopic counterparts is performed visually using custom interactive software: MCVCM. I compare the results from this visual matching to results from simple coordinate-based nearest-neighbour matching. The visually constructed catalogues are substantially more robust due to a better accounting of chance alignments and dense field mismatches. Both cross-matching techniques are performed directly from radio-to-optical, and indirectly by matching radio sources to an infrared-optical paired catalogue. There is a definite improvement for both visual and nearest-neighbour cross-matching when performed via SWIRE infrared intermediary catalogue. I use the visual catalogue to build a reliable sample of ATLAS sources matched to SWIRE infrared photometry, DES optical photometry, and OzDES spectra. This catalogue was developed carefully and is intended to for use as a comparison and training catalogue for the development of intelligent automated cross-matching algorithms

Galaxies are segregated into star-forming and AGN populations using a combination of visual spectral classification, morphology, multi-frequency diagnostics, and optical spectral line diagnostics. For these populations, 1.4 GHz radio luminosity functions (RLFs) are constructed incrementally out to a redshift  $z < 1.0$ . These results improve upon ATLAS-DR1 RLFs and agree with established literature results. Evidence of the space-density evolution of radio source is apparent for both populations towards higher redshifts, indicating increased star-formation and AGN luminosity in earlier cosmic epochs. I construct RLFs for both visual galaxy classification and a combined classification based on a novel numerical combination of classification diagnostics. The combined classification provides higher reliability and completeness resulting in less confounded and deeper RLF. Cross-matching techniques capable of identifying extended and confounded radio sources combined with a robust galaxy classification scheme are therefore essential for quantifying the evolution of radio galaxy populations.

---

THIS PAGE INTENTIONALLY LEFT BLANK

TO A YOUNG BOY DREAMING OF EXPLORING SPACE  
TO A YOUNG BOY DREAMING OF PUNCHING FIGHTER JETS THROUGH CLOUDS  
A YOUNG BOY DREAMING OF SWORDS AND MAGIC AND SECRET PLACES  
DREAMING OF BUYING HIS MUM AND DAD A NEW HOUSE  
TO A YOUNG BOY, FULL OF ANGST, WISHING THERE WAS MORE TO HIS LIFE

KEEP THAT HUNGER ALIVE, MATE  
KEEP HUNTING FOR MORE.

THIS PAGE INTENTIONALLY LEFT BLANK

# Contents

<b>0</b>	<b>INTRODUCTION</b>	<b>1</b>
0.1	Galaxies . . . . .	1
0.2	Radio Galaxies . . . . .	3
0.3	The Core of a Galaxy . . . . .	5
0.4	Active Galactic Nuclei . . . . .	5
0.5	Elusive AGN . . . . .	8
<b>1</b>	<b>ATLAS SOURCE IDENTIFICATION AND MULTI-FREQUENCY COUNTERPARTS</b>	<b>11</b>
1.1	Radio Data: ATLAS and the ATCA . . . . .	11
1.2	Optical Data: DES Photometry & OzDES Spectroscopy . . . . .	12
1.3	Infrared Data: SWIRE . . . . .	13
1.4	Cross-Matching Multi-Frequency Catalogues . . . . .	13
1.5	Comparison of Cross-matching Techniques . . . . .	14
1.6	Blind Nearest-Neighbour Cross-Matching . . . . .	19
1.7	Multi-Catalogue Visual Cross-Matching (MCVCM) . . . . .	22
1.8	Final ATLAS radio catalogue . . . . .	29
<b>2</b>	<b>DIAGNOSTIC POPULATION SEGREGATION OF ATLAS RADIO-GALAXIES</b>	<b>32</b>
2.1	Discerning Galaxy Characteristics . . . . .	32
2.2	Categorising Galaxies . . . . .	33
2.3	OzDES Redshifts . . . . .	35
2.4	Spectral Fitting With pPXF . . . . .	35
2.5	Spectral Diagnostic Classification . . . . .	36
2.6	Visual Spectral Classification . . . . .	36
2.7	Combining Classifications . . . . .	41
2.8	Radio–Infrared Correlation . . . . .	43
2.9	Final ATLAS Radio Population Classifications . . . . .	45
<b>3</b>	<b>EVOLUTION OF ATLAS RADIO-GALAXY POPULATIONS</b>	<b>48</b>
3.1	Radio Luminosity Functions . . . . .	49
3.2	Constructing Radio Luminosity Functions . . . . .	50
3.3	Radio-Luminosity Function Pipeline . . . . .	53
3.4	ATLAS Source Statistics and Restrictions . . . . .	54

3.5	The Complete ATLAS Radio Luminosity Function . . . . .	58
3.6	Split Radio Luminosity Functions . . . . .	59
3.7	Final Thoughts . . . . .	67
4	DISCUSSION AND SUMMARY	69
APPENDIX A MCVCM: A PYTHON PROGRAM FOR MULTI-CATALOGUE VISUAL CROSS-		
	MATCHING	72
A.1	Foreword . . . . .	72
A.2	Description . . . . .	72
A.3	Cutout Assembly . . . . .	73
A.4	User Interaction . . . . .	73
A.5	Version Control . . . . .	78
A.6	Catalogue Management . . . . .	79
A.7	Continuing Development . . . . .	79
APPENDIX B MEASURING CROSS-MATCHING CONTAMINATION VIA RANDOM PER-		
	TURBATION VERSUS CATALOGUE OFFSET	81
APPENDIX C DATA AND FIGURES		83
REFERENCES		100

# Acknowledgments

Visualisation and analyses throughout this thesis were produced using custom [Python](#) scripts. These scripts rely heavily on the pre-existing Python packages: [numpy](#), [matplotlib](#), [scipy](#), [astropy](#), [astroml](#), [iPython](#), & [scikit-learn](#). Spectral line fitting utilised Michele Cappellari's Python implementation of [pPXF](#). Nearest-neighbour positional cross-matching of catalogues was conducted in [TOPCAT](#). Full field image mosaics used Caltech University's [montage](#).

This thesis was written entirely in [Sublime Text](#) and compiled using [LaTeXTools](#). Almost all code was also written in Sublime Text with the eventual (albeit, far too late) switch to the much more feature-rich [PyCharm](#). Spelling mistakes and grammatical errors in this thesis were eradicated with [Grammarly](#). Many hours were saved during the assembly of the bibliography by use of Michele Vallisneri's [Compu-salon ADS database scraper](#).

Advice and insights were provided by my supervisors: Doctor Stas Shabala, Professor Ray Norris, and Associate Professor Andrew Cole; as well as by ATLAS team members: Rob Sharp, Doctor Nicholas Seymour, and Doctor Amy Kimball. The insight provided by the anonymous examiners of this thesis were invaluable, as they provided the final trimming required to click this piece of work smoothly into the infinite academic jig-saw.

Programming advice was curated from questions and answers on [Stack-Overflow](#). The Author would in particular like to thank Stack Overflow user [ImportanceOfBeingEarnest](#) for his seemingly infinite knowledge and passion for helping others.

With the exceptions of the acknowledgements above: the writing of this thesis, the scientific results herein, and all supporting code required for analysis was conducted or produced solely by me, *The Author*.

I am grateful for the financial support provided by the Australian Postgraduate Award, and especially the scholarship provided by The Mathematics and Physics department under the University Of Tasmania's Elite Research Scholarship program.

All non-professional acknowledgements, I have given in person. I will always be grateful for those whose company I have enjoyed over these four long years.

# Figures listing

1	Hubble galaxy classification sequence . . . . .	3
2	Illustration of the Unified Model for AGN . . . . .	7
3	ATLAS 1.4 GHz continuum maps of CDFS and ELAIS-S1 . . . . .	10
1.1	Photometric completeness of DES photometry . . . . .	12
1.2	MCVCM examples of common causes for incorrect cross-matches . . . . .	15
1.3	MCVCM examples of incorrect multi-component sources . . . . .	16
1.4	MCVCM examples of blended radio components . . . . .	17
1.5	Nearest-Neighbour cross-matching radius (DES–OzDES) . . . . .	21
1.6	Nearest-Neighbour cross-matching radius (ATLAS–DES & ATLAS–SWIRE) . . . . .	23
1.7	Optical cutout across three tiles . . . . .	26
1.8	Overlap of ATLAS, DES, and SWIRE . . . . .	27
1.9	Comparison of separation distance for techniques and data-paths . . . . .	29
1.10	ATLAS $m_G$ , $m_R$ , $m_I$ , $m_Z$ spectroscopic completeness . . . . .	31
2.1	pPXF fitting and spectrum for source 2940027955 . . . . .	37
2.2	Visual Classification: Star-Forming and Unknown . . . . .	38
2.3	Visual Classification: Ellipticals . . . . .	39
2.4	Visual Classification: Emission Line AGN . . . . .	40
2.5	BPT: visual classifications . . . . .	41
2.6	Lamareille blue-diagram: visual classification . . . . .	41
2.7	Lamareille blue-diagram: classifications from BPT . . . . .	42
2.8	Lamareille blue-diagram: Combined classification . . . . .	44
2.9	BPT: Combined classification . . . . .	44
2.10	$S_{1.4}$ vs. $S_{24,3,6,4,5,5,8,8,0\mu m}$ : combined classification . . . . .	45
2.11	$S_{24,3,6,4,5,5,8,8,0\mu m}/S_{1.4}$ vs. $\log(1+z)$ : combined classification . . . . .	46
2.12	$S_{24\mu m}/S_{1.4}$ vs. $\log(1+z)$ : manual re-classification . . . . .	46
2.13	ATLAS redshifts: star-forming galaxies and AGN . . . . .	47
3.1	Representation of a redshift limited volume . . . . .	51
3.2	Representation of a radio limited volume . . . . .	52
3.3	Reconstruction of BH12 RLF: examples of automatic binning . . . . .	54
3.4	ATLAS 1.4 GHz luminosity versus redshift . . . . .	55
3.5	Histogram of 1.4 GHz flux separated into classifications . . . . .	56

3.6	ATLAS $m_G, m_R, m_I, m_Z$ spectroscopic completeness . . . . .	57
3.7	Spectroscopic success rate of $G, R, I, Z$ . . . . .	58
3.8	ATLAS RLF of all sources: $0.005 < z < 0.26, 0.26 < z < 0.58, 0.58 < z < 1.0$ . . . . .	60
3.9	ATLAS RLF of star-forming galaxies: $z < 0.26, 0.26 < z < 0.58, 0.58 < z < 1.0$ . . . . .	63
3.10	ATLAS RLF of AGN: $z < 0.26, 0.26 < z < 0.58, 0.58 < z < 1.0$ . . . . .	64
3.11	ATLAS RLF of unknown sources: $z < 0.26, 0.26 < z < 0.58, 0.58 < z < 1.0$ . . . . .	66
3.12	ATLAS RLF of unknown sources & AGN: $0.26 < z < 0.58, 0.58 < z < 1.0$ . . . . .	67
3.13	High- $z$ ATLAS RLF of unknown sources & AGN: $z > 1.0$ . . . . .	68
A.1	Examples of various states during an MCVCM cross-matching instance . . . . .	80
B.1	Nearest-Neighbour cross-matching radius perturbation demonstration . . . . .	82
C.1	Recalculated ATLAS DR1 RLF . . . . .	83
C.2	MCVCM cutouts of $S_{24\mu\text{m}}$ selected AGN (1–9 of 23) . . . . .	85
C.3	MCVCM cutouts of $S_{24\mu\text{m}}$ selected AGN (10–18 of 23) . . . . .	86
C.4	MCVCM cutouts of $S_{24\mu\text{m}}$ selected AGN (19–23 of 23) . . . . .	87
C.5	Spectra of $S_{24\mu\text{m}}$ selected AGN (1–10 of 38) . . . . .	88
C.6	Spectra of $S_{24\mu\text{m}}$ selected AGN (11–20 of 38) . . . . .	89
C.7	Spectra of $S_{24\mu\text{m}}$ selected AGN (21–30 of 38) . . . . .	90
C.8	Spectra of $S_{24\mu\text{m}}$ selected AGN (31–38 of 38) . . . . .	91

*As we got farther and farther away, the Earth diminished in size.  
Finally, it shrank to the size of a marble, the most beautiful marble  
you can imagine ... seeing this has to change a man*

– James Irwin

# 0

## Introduction

### 0.1 GALAXIES

Edwin Hubble's measurements of Cepheid variable stars in nearby dwarf galaxy NGC 6822 was the first step in unfolding the true extent of our Universe ([Hubble 1925](#)). Before these results, galaxies were thought to be nebulae within our small ( $\sim 300\,000$  light year diameter) 'Universe of stars'; a Universe that we now know as the Milky Way Galaxy which is neither distinct nor unique amongst a Universe of galaxies. Galaxies are vast quantities of gravitationally bound dust, gas, and stars embedded within a much larger dark matter halo ([Navarro et al. 1996, 1997](#)). Most are host to a central super-massive black hole (SMBH) the mass of which is intimately tied to wider galactic properties.

#### 0.1.1 SYMPTOMATIC GALAXY CLASSES

In cold dark matter (CDM) cosmology the first galaxies formed in higher density filaments of the primordial Universe. These filaments were formed via the propagation of quantum fluctuations throughout this physical state, and these filaments are evident in the temperature variations in the cosmic microwave background (e.g., [Eisenstein et al. 2005](#); [Planck Collaboration et al. 2014c](#)). In these filaments dark matter and baryonic gas and dust coalesced under grav-

ity as the gas cooled and collapsed stars began to form about local centres of mass, forming the first galaxies. In these dense cosmic filaments, orbiting galaxies lost energy via dynamical friction merging into larger galaxies as the universe aged.

Separating these galaxies into classes is ubiquitously the first step in evaluating their properties. As scientific instruments are advanced we are able to discover increasingly detailed difference between galaxies within pre-defined classes, making classifications more complex and simultaneously improving our understanding of galaxy populations. The simplest and most obvious form of classification is via morphology, and the original and most famous of these is the Hubble Sequence (or colloquially, the Hubble Tuning Fork; [Hubble 1936](#); see Figure 1). This classification scheme breaks galaxies into three classes: smooth and featureless ellipticals, spirals, and irregulars (a later addition to the scheme). Within each of these classes, there are subclasses that focus on elongation, or the presence of a spiral bar, for example.

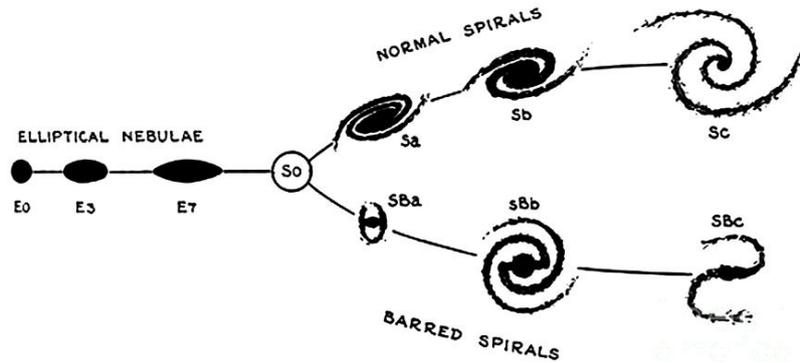
The differences between morphologically selected galaxies provide an excellent basis for population research. Spiral galaxies contain vast amounts of star formation, and their luminosity is dominated by these young, bright stars. They exist primarily in sparse clusters or cluster outskirts \* and tend to have much smaller masses than elliptical galaxies. The more massive elliptical galaxies have large featureless haloes and are dominated by an older stellar population due to a lack of ongoing star formation; they also tend to reside within denser cluster environments than spirals. Understanding the many differences between these galaxies and the environment that shapes them allows for the boot-strapping of more complex or controversial classification schemes that serve even further to advance our understanding of galaxy formation and evolution. Increasingly detailed observations and simulations of galaxies and galaxy populations fortify that understanding and allow new classification schemes to be tested.

### 0.1.2 MODELLING THE UNIVERSE

$\Lambda$ -CDM cosmology is the most broadly accepted cosmological model for the fundamental reason that it is thus far the best fit to our observations of the Universe. Particularly the 2011 Nobel Prize for Physics was awarded to Adam Riess, Brian Schmidt, and Saul Perlmutter for their discovery of an accelerating expanding Universe ([Riess et al. 1998](#); [Perlmutter et al. 1999](#)). The cosmological constant  $\Lambda$  predicts this accelerating expansion on the back of a cold dark matter (CDM) cosmological model. The CDM model estimates that roughly 5 per cent of energy is in the form of baryonic matter, 23 per cent is weakly interacting dark matter, and the remaining 72 per cent is dark energy permeating the entirety of space and accelerating the expansion of the Universe.

---

\*Although “red-spirals” are known to exist in large cluster interiors (e.g., [Wolf et al. 2009](#))



**Figure 1:** “Sequence of Nebular Types” (Hubble 1936, Figure 1). Hubble, E. states that mixed galaxy types exist between the two spiral branches and that the ‘S0’ transition phase is hypothetical. The transitions between ellipticals are noted as smooth and continuous.

Cosmological models such as these build upon over a century of observational data and help to advance astronomy by providing observational astronomy with a plethora of provably falsifiable predictions. These predictions act to focus astronomy and astrophysics on the furthering of our understanding of the Universe in significant ways. Enormous galaxy surveys provide data on the low redshift Universe faster than current methodologies can process, and groundbreaking observational techniques coupled with incredible technology (such as adaptive optics and space telescopes) yield detailed information on some of the most distant objects in the universe (e.g., gravitational lensed galaxies: [Watson et al. 2015](#)).

## 0.2 RADIO GALAXIES

### 0.2.1 A RADIO SKY

The many designs of radio telescopes and receivers allows for vast and comprehensive information on extragalactic objects. In their simplest form, single-dish radio telescopes are cheap and simple. These telescopes can be built large (diameters in excess of tens of meters), they have little telluric attenuation, and using the phenomena of interferometry we can easily construct arrays of telescopes that provide sampling resolution many orders of magnitude greater than a single dish can achieve. Modern radio astronomy is advancing with the development of expensive and complicated purpose-built radio telescope arrays such as ALMA (Atacama Large Millimeter/submillimeter Array) and the SKA (Square Kilometer Array). The brightest and most energetic objects in the universe emit vast amounts of energy in the radio spectrum, allow-

ing these sources to be detected at great distances with single dish telescopes and resolved with increasing detail using Long Baseline Arrays (LBA).

Bandpasses of radio-frequencies are sensitive to different physical processes. High-frequency radio is dominated by large-scale dust emission from thermal ( $\sim 20$  K) gas in the disk of galaxies (e.g., Planck 353 GHz sky map: [Planck Collaboration et al. 2014a](#)). Unfortunately, this means that extragalactic sources coincident with the galactic plane of our Milky Way (MW) suffer from dramatic line-of-sight extinction. For frequencies on the order of tens of gigahertz, the total sky-brightness is dominated by the Cosmic Microwave Background (CMB) colloquially, ‘the after-glow of the Big Bang’, this light permeates all of space with a black body temperature of 2.73 K. Whilst the CMB broadly dominates the the Planck 30 GHz sky map ([Planck Collaboration et al. 2014b](#)), observations of galaxies in this frequency regime are dominated by thermal free-free emission, the tail of the 80 K galactic dust black-body spectrum, and synchrotron emission.

This thesis focuses primarily 1.4 GHz synchrotron radiation, decelerating electrons whose radiation can be approximated by a power law  $F_\nu \propto \nu^\alpha$  as they spiral along strong magnetic fields. These power-law spectra exhibit a break or turn-over due to (for example) free-free or synchrotron self-absorption ([Jaffe & Perola 1973](#); [O’Dea 1998](#)). The observed flux (for example) dominates the Haslam 408 MHz sky map ([Haslam et al. 1982](#)) but falls off rapidly at higher frequencies, the steepness of this decline being affected by a redshift and spectral ageing degeneracy (e.g., [Machalski et al. 2010](#), and references therein). Galaxies with sufficient star formation (SF) or those host to a bright active-galactic nucleus (AGN) are detectable at synchrotron radiation frequencies. In SF galaxies the synchrotron emission primarily originates from the acceleration of electrons by supernova shocks, and (far less strongly) the free-free emission from HII regions. In galaxies host to a powerful AGN the emission generated from powerful jets dwarfs that of the star-formation ( $P_{1.4} > 10^{24}$  W/Hz).

## 0.2.2 UNDERSTANDING THE BRIGHTEST RADIO GALAXIES

Early observations of radio galaxies revealed luminosities requiring energy inputs far beyond what is possible from star formation processes ([Begelman et al. 1984](#)). Even more puzzling was the rapid variability of the brightest of these sources (quasars) which could see 10 to 300 per cent changes in peak flux over several days, or months ([Smith & Hoffleit 1963](#); [Greenstein & Schmidt 1964](#); [Hunstead 1972](#)). This variability provided the key for obtaining the spatial scale from which these vast emissions originated, restricted by the strict condition that the time-scale of luminosity variability is greater than the light-crossing time (i.e.,  $\tau_{\text{variability}} > R/c$ ; [Mushotzky 1984](#)).

The best early model for explaining this behaviour was provided by [Salpeter \(1964\)](#) and

Zel'dovich & Novikov (1964) in which they mathematically demonstrated that observed jets could be reproduced by the accretion of gas onto a supermassive black hole (SMBH) on the order of  $10^6 - 10^{10} M_{\odot}$ . This phenomenon would reproduce the energy output (with up to 10 per cent rest-mass energy conversion of in-falling gas) as well as the rapid variability due to the activity necessarily constrained to within tens to hundreds of Schwarzschild radii.

Supermassive black holes exist at the centre of most galaxies with the probability increasing as a function of galaxy mass (e.g., Kormendy & Richstone 1995; Kormendy & Ho 2013). The Milky Way galaxy possesses a central SMBH (Sagittarius A\*) with mass  $\sim 4.1 \times 10^6 M_{\odot}$  (Gillessen et al. 2009; Reid & Brunthaler 2004). The presence of Sagittarius A\* has been independently confirmed by multiple techniques: orbital periodicity of nuclear stars (Schödel et al. 2002; Gillessen et al. 2009); polarimetry (Bower et al. 1999; Meyer et al. 2006); Doppler measurements of orbiting dust and gas (Levin & Beloborodov 2003); water megamasers associated with the accretion disk (Miyoshi et al. 1995), and M- $\sigma$  velocity dispersion relation (Ferrarese & Merritt 2000; Wjithe 2006). Sagittarius A\* is not considered active, the search for jets has revealed only small remnants of previous activity (Doeleman et al. 2008; Bordoloi et al. 2017).

### 0.3 THE CORE OF A GALAXY

#### 0.3.1 RADIO JETS

The mechanism for the production of radio jets took longer to establish than did the origin of the emission. In a 1971 *Nature* publication Rees (1971) first posited that pressure of low-frequency ( $\lesssim 1$  kHz) radiation<sup>†</sup> can produce a cavity within the central region, and if there is sufficient collimation of this radiation, jets of relativistic electrons are produced with high efficiency. This theory overcame the restriction of earlier theories which depended on unreasonably large ordered magnetic fields, and adiabatic energy losses of the expanding plasma to produce the collimated jets with a ballooning termination. Rees (1971) reproduces the structure and more believably produces twin jets found in double-lobed sources.

#### 0.4 ACTIVE GALACTIC NUCLEI

The classification of AGN relies heavily on spectral characteristics and falls broadly into seven categories; Seyfert 1, Seyfert 2, broad-line radio galaxies (BLRGs), narrow-line radio galaxies (NLRGs), and blazars. As can be seen in Figure 2.4 Seyfert 1, Seyfert 2, BLRGs, and NLRGs

---

<sup>†</sup>This low-frequency radiation originates from the theory that a compact spinning source with a sufficiently strong magnetic field will lose angular momentum through the emission of radiation at the spin frequency (e.g., Znajek 1977).

each have strong, narrow emission of collisionally excited forbidden emission lines, and recombination lines. Of these four classes, Seyfert 1 and BLRGs have heavily broadened contributions to some of these narrow emission lines from plasma velocities on the order  $10^4 - 10^{10}$  km s<sup>-1</sup>. Of the blazar subclasses, BL Lac are distinguished by flat optical spectra with little or no emission lines, whilst flat-spectrum radio quasars (FSRQ) have the same flat optical spectrum, but often display optical emission lines.

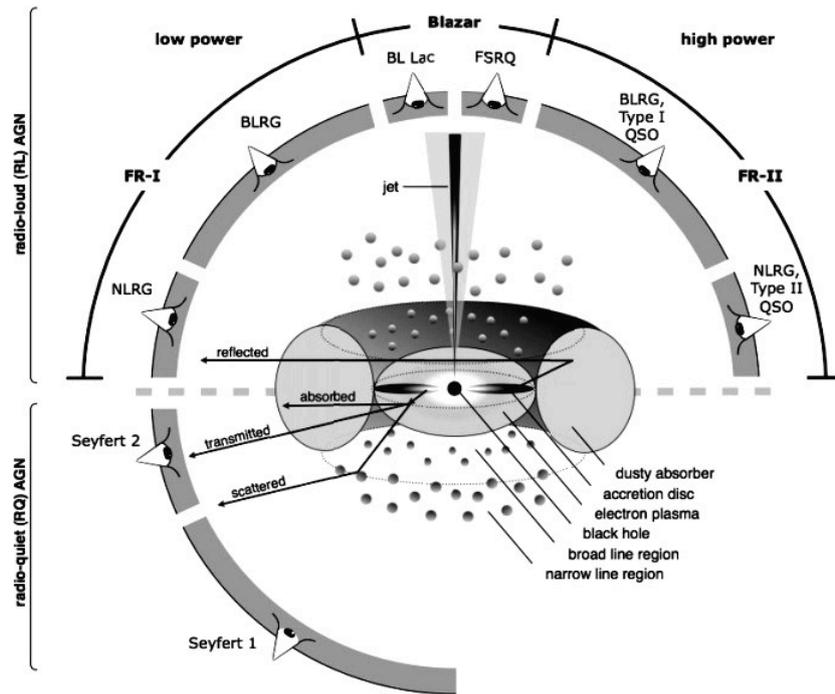
#### 0.4.1 THE UNIFIED MODEL OF AGN

The Unified Model (Antonucci 1993; Urry & Padovani 1995) explains these seemingly different AGN by starting with the accepted model of a galactic core and applying two simple physical principles: an obscuration model, and a beaming model. Stating that there should be no substantial difference between how the cores of active galaxies behave; the difference should only arise from the observation perspective and the energy output. The obscuration model involves an optically-thick dusty torus around the accretion disk and SMBH. This torus acts as a filter to the nuclear emission for particular orientations to our line-of-sight obscuring, scattering or absorbing and re-emitting to produce the AGN spectral variants. For example, BLRGs are reproduced when the high-velocity nuclear plasma is visible. The higher infrared–optical flux ratio of Seyfert-1 compared to Seyfert-2 galaxies is reproduced with partial obscuration of the optically bright nuclear plasma (see Figure 2).

#### 0.4.2 AGN ACCRETION MODES: COLD VERSUS HOT

The Unified model assumes that the emission mechanism of all AGN is the same, and consequently, any major observational differences are attributed to geometry. If this were true, all AGN would undergo very similar cosmic evolution. Best & Heckman (2012) demonstrated the independent evolution of high-excitation radio galaxies and low-excitation radio galaxies (hereafter HERGs and LERGS), the difference in evolution being attributed to different environments and accretion mechanism of the AGN.

The nuclear structure of high-excitation radio galaxies (also quasar-mode, or cold-mode) is similar to that depicted in the unified model (e.g., Figure 2); a physically small and optically bright accretion disk feeds material to the black hole (Antonucci 1993). This disk produces strongly ionising photons that collisionally excite strong optical and x-ray emission lines via the inverse Compton effect. This fast-moving plasma will produce broadened spectral emission lines that if visible, reproduces a Type-I AGN, or if obscured (by an optically thick dusty toroid) a Type-II AGN. The ‘cold’ fuel for this process arrives from merging of a gas-rich galaxy (Best et al. 2006; Hardcastle et al. 2007) a claim supported by the high rate of galaxy-galaxy interac-



**Figure 2:** Schematic representation of the Unified Model of AGN (Antonucci 1993; Urry & Padovani 1995), the various notations depict the expected classification based on an AGN viewed from that angle (Original graphic from Beckmann & Shrader 2012, source of the alterations depicting point of view are unknown).

tions and distorted morphology of this population (Smith & Heckman 1989).

Low-excitation radio galaxies (also radio-mode, or hot-mode) do not exhibit high-ionisation optical or x-ray emission characteristics (Laing et al. 1994; Best & Heckman 2012), nor do they display evidence of a nuclear accretion disk, or an infrared-bright dusty toroid (Müller et al. 2004; Whyson & Antonucci 2004), instead releasing energy in the form of kinetic jets (Gürkan et al. 2014). The fuel source for LERGs has been established as gas cooling on to the central black hole from the ‘hot’ halo or cluster environment (Best et al. 2006; Best & Heckman 2012) or the “chaotic accretion” of small fractions of a cold gas reservoir (see Gaspari et al. (2013)). LERGs are observed to have a higher mass-to-light ratio (ML ratio), redder stellar populations, and a more elliptical morphology than cold-mode HERGs (Kauffmann et al. 2003, 2008; Best & Heckman 2012).

#### 0.4.3 REGULATING STAR-FORMATION

Both HERGs and LERGs deposit large amounts of energy into their galaxy and surrounding cluster environment (Best et al. 2006). HERGs are shown to drive intragalactic winds that sup-

press and enhance star-formation via the heating of star-forming gas or the shock-triggering of star-formation in stable nebulae (Di Matteo et al. 2005; Hopkins et al. 2006) with an expected  $\sim 0.5$  Gyr lag (Schawinski et al. 2007). This Cold-mode accretion and the associated depositing of energy are important for establishing the correlations between black hole mass, and other stellar properties observed in the local universe (Silk & Rees 1998; Fabian 1999; Robertson et al. 2004).

Although the energy output from LERGs is approximately two orders of magnitude less than that of HERGs this energy is largely deposited locally, affecting star formation but also re-heating the hot halo gas feeding the AGN potentially generating a self-regulating system that limits galaxy growth (Best et al. 2006).

#### 0.4.4 EVIDENCE OF FEEDBACK

The introduction of AGN feedback from HERGs and LERGs into semi-analytic (SA) models (e.g., Croton et al. 2006; Shabala & Alexander 2009) and hydrodynamic simulations (e.g., Di Matteo et al. 2007; Sijacki et al. 2007) overcame many of the shortfalls of SA models compared to observational data such as metallicity and molecular contents, and run-away star formation (Nagashima et al. 2005; Benson et al. 2003). Direct observations of cold molecular gas entrained by the outflows of AGN further supports an AGN feedback system (Walter et al. 2002; Morganti et al. 2005; Bolatto et al. 2013).

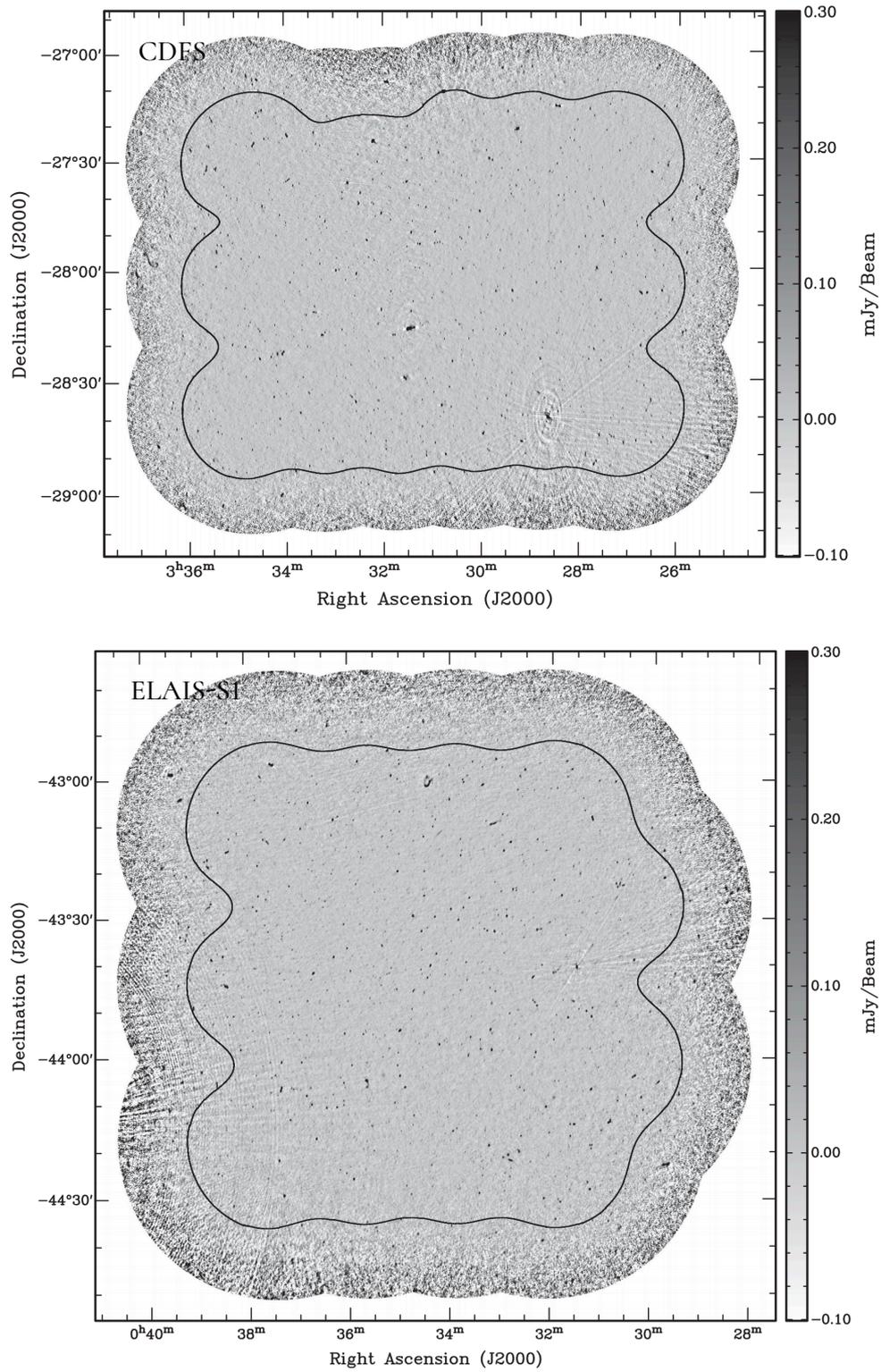
#### 0.5 ELUSIVE AGN

NGC 4945 is a famous example of a hard-to-identify AGN. The optical spectrum of the galaxy contains no indicators of an AGN (Lípari et al. 1997) and mid-infrared spectroscopy shows clear features of star-forming activity (Brandl et al. 2006); however it is the brightest x ray source in the sky at  $\sim 100$  keV and shows a Compton-thick AGN profile with a bright Fe  $K\alpha$  line (Puccetti et al. 2014); indisputably NGC 4945 is host to an AGN. The edge-on galaxy NGC 1448 tells a similar story: a high-resolution optical spectrum of the nucleus region shows no sign of an AGN, unresolved infrared emission is dominated by star formation, but an x-ray spectrum shows a clear Compton-thick AGN (Annunari et al. 2017).

NGC 4945 and NGC 1448 are both nearby edge-on galaxies at distances of 3.6 and 17.3 ( $\pm 2.3$ ) Mpc respectively. Even with this proximity typical diagnostics of optical spectroscopy and infrared colour-colour diagnostics are unable to classify these galaxies correctly. These examples highlight an issue with the classification of AGN in surveys, where it is not always possible to carefully examine multiple diagnostics or explore edge-cases. The small scale of the accretion disk permits relatively fast variability and emission of key diagnostic features are

obscured by the dusty environment around the accretion disk, especially if the nucleus resides within a spherical obscuration region (e.g., [Spoon et al. 2001](#)). The bulk of the host galaxy further acts to obscure or dilute the AGN. These issues become more apparent as the redshift of the sample grows, the spectra will be integrated across a large area of the host galaxy due to a fixed spectral fibre size. Evidently, the non-detection of an AGN does not rule out the presence of one. [Goulding & Alexander \(2009\)](#) found that from a sample of 17 galaxies identified as AGN using mid-infrared spectroscopy, only 7 were optically identified as AGN and these were each found to be face-on galaxies, this discovery indicating that obscuration of the galaxy core is very significant in optical identification – biasing samples that use only a single diagnostic.

In this thesis, I establish a reliable identification of ATLAS sources by visually cross-matching the radio source positions with photometric and infrared catalogues using custom software. I compare the results of this catalogue cross-matching to a blind-nearest neighbour cross-matching (Chapter 1). Using the visually cross-matched ATLAS catalogue I then segregate the ATLAS sources into galaxies dominated by star-formation or an AGN using visual spectral classification, spectral emission diagnostics, and radio-infrared diagnostics. I establish a simple means of combining these diagnostic classifications in an effort to minimise classification biases (Chapter 2). Finally, I construct the 1.4 GHz radio luminosity function (RLF) establishing the spatial density of ATLAS sources, and evaluate how this density changes with redshift. This is repeated for galaxies classified as AGN or SF to establish the independent evolution of these populations (Chapter 3), and conclude in Chapter 4.



**Figure 3:** ATLAS 1.4 GHz continuum maps of CDFS and ELAIS-S1 (figures are labelled accordingly). The solid contour denotes the bounding region of the ATLAS survey: sources outside of this region excluded from the ATLAS catalogue. Intensity of 1.4 GHz flux-density is indicated by the color-bar to the right of each plot. These graphics are sourced from [Franzen et al. \(2015\)](#).

*What does the student do? Chops wood, and carries buckets of water.*

*What does the master do? Chops wood, and carries buckets of water.*

– Chinese Proverb

# 1

## ATLAS Source Identification and Multi-Frequency Counterparts

### 1.1 RADIO DATA: ATLAS AND THE ATCA

The Australia Telescope Large Area Survey (ATLAS) is a 1.4 GHz radio survey of two fields: European Large Area ISO Survey (ELAIS) S1 field, and Chandra Deep Field South (CDFs) (Norris et al. 2006; Middelberg et al. 2008) using the Australia Telescope Compact Array (ATCA). These fields have a wealth of multi-frequency data available for comparison (e.g., WISE far-infrared (Wright et al. 2010), SWIRE mid-infrared (Surace & et al. 2012), SERVS near-infrared (Mauduit et al. 2012; Nyland et al. 2017), DES and SuperCOSMOS optical photometry (Flaugher 2005; Supercosmos Sky 2007), OzDES optical spectroscopy (Childress et al. 2017), and Chandra x-ray observations (Evans et al. 2014).

The ATLAS survey is currently the deepest radio survey of these fields \*, reaching as low as  $17 \mu\text{Jy beam}^{-1}$  with a spatial resolution of  $11'' \times 5''$  over the combined  $6.3 \text{ deg}^2$  sky area. From these observations, initial source-finding algorithms merged individual Gaussian emission components into a total of 5118 radio sources of which 169 were classified as multi-component sources (see Franzen et al. 2015 for a full description). Based on the per-source visual inspection

---

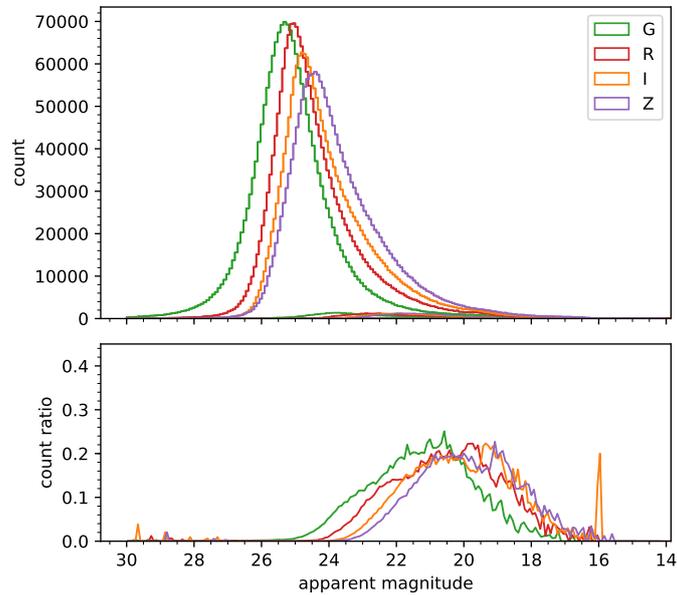
\*The Very Large Array has been used to survey six pointings (primary beam:  $2.8'' \times 1.6''$ ) in CDFS to a sensitivity of  $\sim 7 \mu\text{Jy beam}^{-1}$  from which they locate 883 radio sources (Miller et al. 2013).

in this chapter, I establish a different accounting of total source counts and multi-component sources.

## 1.2 OPTICAL DATA: DES PHOTOMETRY & OZDES SPECTROSCOPY

The Dark Energy Survey (DES; [Flaugher 2005](#)) utilises a purpose-built 570 Mega-pixel CCD camera installed on the Blanco 4 meter telescope at the Cerro Tololo Inter-American Observatory (CTIO; [Diehl et al. 2016](#)). The DES Data Release 1 (DR1; [Abbott et al. 2018](#)) photometric survey provides  $G R I Z$  photometric coverage of approximately 137 million sources over an area of 1800 square degrees. The median point-spread function spatial resolution of the  $I$ -band photometry used throughout this thesis is  $0.97''$  ([Drlica-Wagner et al. 2018](#)).

The Australian Dark Energy Survey (OzDES; [Yuan et al. 2015](#)) is a targeted spectroscopic follow-up of DES which uses the AAOmega spectrograph ([Smith et al. 2004](#)) on the 3.9 meter Anglo-Australian Telescope (AAT). Coupled with the Two Degree Field (2dF) 400-fibre multi-fibre positioning system ([Lewis et al. 2002](#)). Data Release 1 provides spectroscopy for 280 000 sources of which 244 000 have reliable spectroscopic redshifts ([Childress et al. 2017](#)). The top panel of Figure 1.1 shows histograms of source counts for each of the four DES photometric bands: reliably reaching magnitudes as faint as  $m_R < 25$ . The bottom panel of Figure 1.1 shows the per-bin ratio of DES sources with an OzDES redshift for each photometric band. The spectroscopic completeness is established and used in § 3. Throughout this thesis DES multi-band catalogue  $I$ -band astrometry is used for optical source positions, and source visualisation.



**Figure 1.1:** Source counts per photometric band ( $G$ ,  $R$ ,  $I$ ,  $Z$ ) for DES optical photometry. The bottom pane shows the fraction of sources per bin with an OzDES redshift measurement (top), and per-bin ratio of DES sources with available OzDES redshifts.

### 1.3 INFRARED DATA: SWIRE

Spitzer Wide-Area InfraRed Extragalactic Survey (SWIRE; [Lonsdale et al. 2003](#)) has detected 2 million IR-selected galaxies sources over  $\sim 49 \text{ deg}^2$  in six fields. The infrared photometry used in this thesis are MIPS 3.6 – 24  $\mu\text{m}$  channels. At 3.6  $\mu\text{m}$  SWIRE achieves  $5\sigma$  sensitivities of 5  $\mu\text{Jy}$  and a full-width at half-maximum (FWHM) resolution of 1.2", and at 24  $\mu\text{m}$   $5\sigma$  sensitivities of 200  $\mu\text{Jy}$  at spatial resolution of 5.5" ([Surace et al. 2005](#)). SWIRE source positions used herein are drawn from 3.6  $\mu\text{m}$  IRAC data which is complete to  $\sim 97$  per cent at 400  $\mu\text{Jy}$  ([Babbedge et al. 2006](#)).

### 1.4 CROSS-MATCHING MULTI-FREQUENCY CATALOGUES

Difficulties arise in the correct handling of multiple source catalogues when these catalogues originate from different instrumentation; especially when cross-matching optical or infrared source positions with radio source positions. Compared to the accurate astrometry and relatively high resolution of optical surveys, radio surveys use peak-finding algorithms to estimate positions of often diffuse or extended radio sources (e.g., [Franzen et al. 2015](#)). For isolated, compact radio galaxies this technique is sufficiently accurate; however, in dense fields, the radio sources are non-Gaussian and determining the centre of the radio emission becomes guess-work.

An uncertainty in radio source location is especially concerning when cross-matching in a dense optical or infrared field. It then becomes difficult to identify which source of many potential matches is the counterpart of a particular radio peak (e.g., Figure 1.2c). If the observational ‘depth’ of the radio and optical (or infrared) surveys are not well matched, then all issues with correct multi-frequency source association are exacerbated. Potential dense-field misassociation can be somewhat corrected if cross-matching is conducted manually by an expert, as she or he can comment and flag such sources for more careful statistical consideration. If however, automated positional cross-matching is used then poorly matched resolution and survey depth becomes a source for *systematic* error, preferentially misidentifying dense cluster sources.

#### 1.4.1 MATCHES AND MISMATCHES

Software such as TOPCAT<sup>†</sup> can very efficiently perform a sky-coordinate positional cross-matching between two catalogues. However, this easy-to-use software has no discretion for physical distribution or orientation of sources, nor can it provide confidence of match; for example in the event of a dense pairing field. This means we have to accept each match as correct whilst understanding that a large percentage of total matches will be incorrect. An incorrect match would constitute a match to a foreground or background galaxy in the optical (or infrared) catalogue unrelated to the radio source. Such a mismatch could arise from: a poorly located radio-flux peak (Figure 1.2a and Figure 1.4); a non-detection of an optical or infrared

---

<sup>†</sup>Tool for OPERations on Catalogues And Tables ([Taylor 2013](#)).

counterpart (Figure 1.2*b*); a single radio component encompassing a cluster of optical or infrared sources (Figure 1.2*c*); or isolated radio lobes of an extended AGN positionally matched to a coincident background galaxy (Figure 1.2*d*). For the remainder of this chapter, I will broadly refer to these incorrect matches as contaminants, spurious matches, or false matches.

In the case of sources where a host galaxy is misidentified, it begins to introduce a luminosity bias to our results. Due to instrumentation limits, a mismatch will favour a brighter (apparent magnitude) optical counterpart than the correct host. Because radio sources tend to be higher redshift than optical sources, a mismatched optical source is likely to be from a closer redshift causing the radio luminosity calculated from that redshift to be underestimated. If a given source belongs at a higher redshift but is mismatched to a lower redshift, then the luminosity calculated from the flux of this source is lower than the physical luminosity, and the source has been shifted away from its high-redshift population. Practically this has created an under-counting of bright, distant sources and an over-counting of nearby faint sources. Accumulating many of these errors over a large catalogue would result in a systematic luminosity error. With the possibility of vast differences in correct and chosen redshifts, and such mistakes being frequent, the result would be the fabrication of many faint radio sources.

Another issue with the automatic cross-matching method arises from the extended radio sources that make up a small fraction of the AGN in our sample. In blind nearest neighbour (BNN) cross-matching isolated lobes are almost always associated with a coincident, unrelated, optical source (e.g., Figure 1.2*d*). Without manual intervention or a sophisticated automatic algorithm (e.g., Fan et al. 2015) this is guaranteed because nearest-neighbour matching has no information to understand that these lobes are not genuine galaxies<sup>#</sup>. Further to the luminosity error mentioned above, this introduces an over-counting of true radio sources, and an underestimation of the combined flux of the correct radio host<sup>‡</sup>. Whilst there are relatively few of these sources in the ATLAS catalogue, these sources dominate the high-luminosity end of the radio luminosity function; as such, correctly calculating their luminosity and distance is essential for the correct accounting of the spatial densities of galaxies, especially at high redshift where the detection rate of faint sources has diminished.

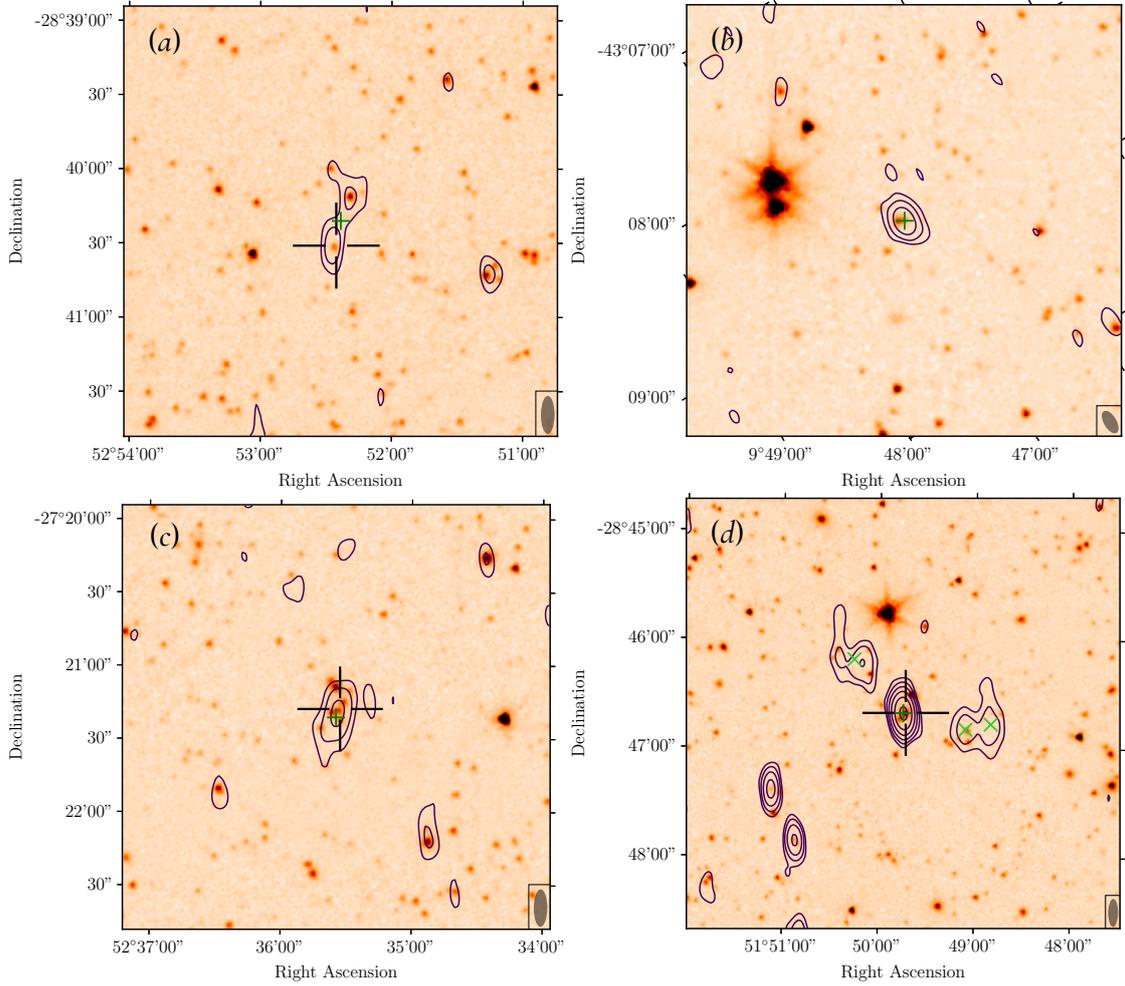
## 1.5 COMPARISON OF CROSS-MATCHING TECHNIQUES

The final two chapters of this thesis rely upon the availability of accurate redshifts and spectra for as many of our ATLAS sources as possible. Methodology aside, associating an ATLAS radio

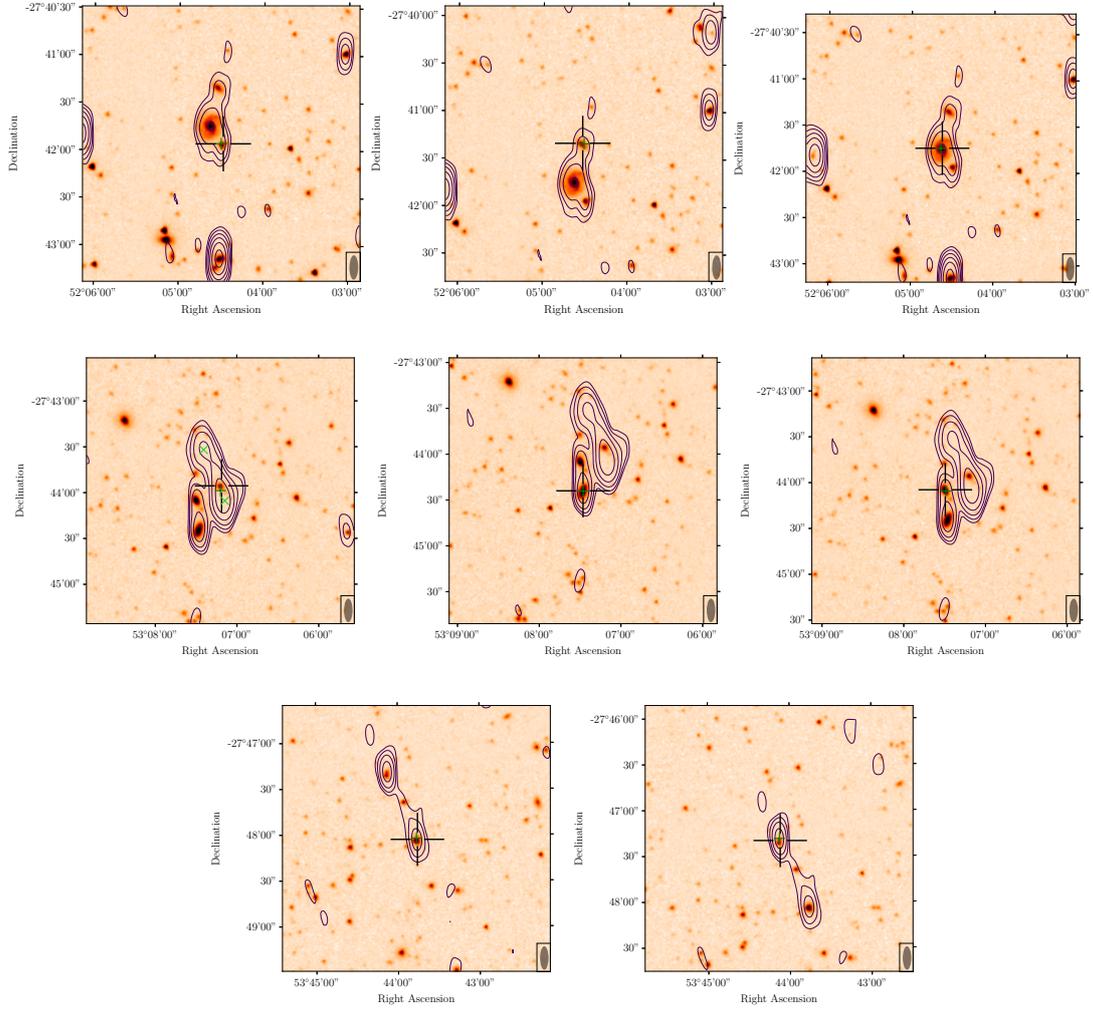
---

<sup>#</sup>There is an emerging research field that hopes to identify isolated lobes automatically, but it either relies upon AGN hot-spots of similar flux symmetrically centered about a radio core, or using Bayesian statistical weighting that encroaches upon supervised machine learning (e.g., Fan et al. 2015). Neither of these methods would be considered a *blind* nearest-neighbour cross-matching technique.

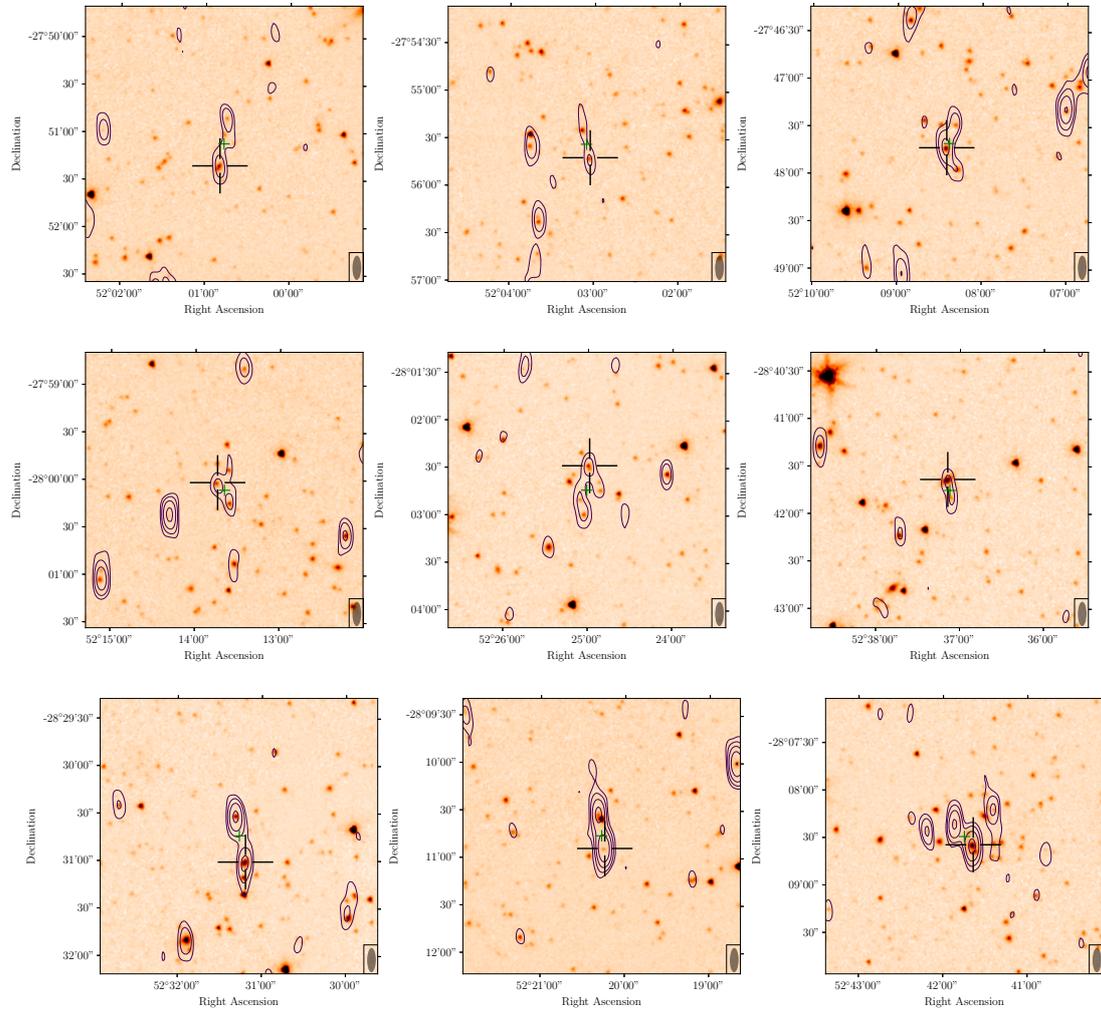
<sup>‡</sup>I the case of a typical double lobed AGN with a faint core and bright hot-spots this error would count one faint source, and two sources of approximately half of the total source flux; assuming that they are by chance associated with coincident sources at identical redshifts.



**Figure 1.2:** Four examples of scenarios that would result in an incorrect cross-match if the cross-matching were done by an automatic method without topographical information: figures were generated from MCVCM. In each panel 1.4 GHz, ATLAS radio continuum contours are overlaid onto SWIRE 3.6  $\mu\text{m}$  infrared emission, contour levels are  $2^n \times \langle \text{RMS} \rangle$  calculated for each cutout. The large black cross-hair shows the location of the user-selected infrared core, and the green crosses show the location of the selected radio component peaks. In the bottom right of each image is the ATCA beam for that field. Panel *a* (top-left) shows a blended radio source that has been incorrectly merged from two distinct radio sources. Panel *b* (top-right) shows a compact, isolated radio source without an infrared catalogue counterpart; the counterpart is visible but was likely removed from the SWIRE infrared catalogue due to obscuration from the nearby foreground star. MCVCM allows such sources to remain ‘unmatched’, but this is not guaranteed with BNN matching. Panel *c* (bottom-left) shows a single radio source encompassing a small cluster of infrared galaxies: these are inseparable. Panel *d* (bottom-right) shows a double-lobed AGN; the isolated radio lobes have been manually associated with the core, but this would not have been the case in a blind automated cross-matching system.



**Figure 1.3:** Three MCVCN examples of scenarios where the automatic source-finding algorithm (Franzen et al. 2015) incorrectly associated independent sources contained within an enveloping radio contour as single multi-component sources. In all three cases, MCVCN demarcations show the corrected radio-infrared associations (see Figure 1.2 for description of the plotting details). The top row shows a spiral galaxies with two possible satellites previously identified as a singular three component source. The middle plots shows an extended double-lobed AGN next to two infrared galaxies, previously identified as a five component source. The two bottom plots show two infrared galaxies incorrectly classified as a two component extended source.



**Figure 1.4:** Nine MCVCM examples of scenarios where the automatic source-finding algorithm (Franzen et al. 2015) merged the radio flux contribution from two or more sources into a single radio source, overestimating the flux of a single source, and under-counting the total sources – MCVCM is unable to remedy this without a catalogue of unfiltered radio continuum peaks. In all nine cases, MCVCM demarcations show the chosen radio-infrared associations (see Figure 1.2 for description of the plotting details).

source with an OzDES targeted spectrum can take two paths <sup>z</sup> (hereafter referred to as data-paths):

- ATLAS 1.4 GHz radio matched directly to DES optical photometry
- ATLAS 1.4 GHz radio matched to DES optical photometry via the SWIRE 3.6 $\mu$ m infrared catalogue as an intermediate step

The possible benefit of an infrared intermediary comes from the relationship between the infrared and 1.4 GHz radio emission discussed in § 0.3, along with the empirical evidence that > 90 per cent of optical galaxies are radio-quiet (Norris et al. 2006). For an optical survey limited to the same redshift range as a radio or infrared survey, there will be a higher density of detections in the optical survey. In practice, this causes an over-crowding effect when trying to match radio and optical catalogues that is often combated by performing a magnitude cut of the optical catalogue before matching to the radio catalogue (e.g., Sadler et al. 2007; Donoso et al. 2009; Pracy et al. 2016). This cut is made on the basis of an assumption that, at its core, assumes ‘bright things are bright’; that is, fainter optical sources are not likely hosts to a bright radio galaxy. I will show later in this section that there are sufficiently many ‘correct’ faint optical counterparts to make such a technique undesirable for an unbiased final catalogue. In this chapter, I aim to confirm that the cumulative position uncertainty between the radio, infrared, and optical catalogues will be smaller than the direct uncertainty between radio and optical catalogues. I also aim to generate a reliable multi-frequency catalogue by visually cross-matching sources and to compare this to a catalogue cross-matched using simple nearest-neighbour cross-matching.

There are several techniques used for cross-matching of catalogue pairs, some that will be covered or discussed in this thesis are:

- blind nearest-neighbour (BNN) positional matching <sup>el</sup>,
- manual matching of sources via visual catalogue inspection,
- custom machine learning, or Bayesian algorithms (e.g., Fan et al. 2015; Lukic et al. 2018),
- citizen scientist visual matching (e.g., Banfield et al. 2015).

The third and fourth techniques will be briefly discussed later in this chapter, but they are considered ‘intelligent’ automation and are beyond the scope of this work; these techniques require a pre-existing ‘golden standard’ cross-matched catalogue for training and comparison. This chapter aims to establish such a standard by conducting and evaluating the first two techniques (BNN and visual matching) for both of our data-paths.

---

<sup>z</sup>in all analyses I have performed a nearest-neighbour cross-matching of DES photometry to OzDES targeted spectroscopy, and assume that this cross-matching is sufficiently accurate not to affect our final ATLAS redshifts regardless of cross-matching technique or data-path.

<sup>el</sup>blind, in this sense, refers to the matching algorithm having no topographical or physical information about the sources beyond a right-ascension – declination coordinate

## 1.6 BLIND NEAREST-NEIGHBOUR CROSS-MATCHING

The cross-matching of two catalogues using a BNN technique only requires the specification of a matching radius. A radius that is too small compromises the completeness of the final catalogue, whilst with a sufficiently large matching radius, the catalogues can be matched to 100 per cent completion, but it becomes increasingly likely that those matches are spurious. In the following sections, I will discuss the reliability of cross-matching via BNN and establish an automated, repeatable method for choosing a cross-matching radius that maximises reliability.

### 1.6.1 EVALUATING BLIND NEAREST-NEIGHBOUR CROSS-MATCHING

To establish a standardised method for calculating a ‘best-fit’ cross-matching radius, I assess the increase in cross-match counts as cross-matching radius is increased. To quantify completion, this is performed for the cross-matching of our primary and secondary<sup>‡</sup> catalogues, whereas the number of false matches is quantified by cross-matching the primary catalogue to a randomised pairing catalogue. The random catalogue is generated by perturbing the right-ascension and declination of each source from the secondary catalogue by a random amount. This random perturbation ( $\delta_r$ ) is drawn from a range spanning  $-45'' < \delta_r < 45''$  (see Appendix § B for a full justification). The completeness,  $C$ , of each catalogue pairing is defined as the fraction of total ATLAS catalogue sources with a match, regardless of validity.  $C$  is calculated as a function of radius as:

$$C(r) = \frac{N(r)}{N_{\text{ATLAS}}} \quad (1.1)$$

For these catalogue pairings, a fraction of the total matches will be incorrect spurious matches or contaminants. It is effectively impossible to discover which of our matches between the primary and secondary catalogues are false, but the percentage of false matches can be estimated by performing the cross-matching between our primary catalogue and a randomised secondary catalogue.

The number of false matches will also increase as the cross-matching radius is increased, and the spurious fraction,  $S$ , is defined as:

$$S(r) = \frac{N_{\text{rand}}(r)}{N_{\text{ATLAS}}}. \quad (1.2)$$

The variable  $N_{\text{rand}}$  is the number of matches between the primary catalogue (ATLAS in this example) and the random catalogue described above. In both of these equations,  $N_{\text{ATLAS}}$  is the total number of sources in the ATLAS catalogue<sup>§</sup>.

<sup>‡</sup>In all cases the secondary catalogue is the catalogue with the greatest source density.

<sup>§</sup>Where necessary  $N_{\text{ATLAS}}$  is substituted for the appropriate source count from the catalogues being cross-matched.

In this chapter, I cross-match ATLAS 1.4 GHz with the SWIRE 3.6 $\mu$ m infrared catalogue, and ATLAS 1.4 GHz with DES photometric catalogue. I also cross-match DES with OzDES, and SWIRE with DES so that I can build a final ATLAS–DES–SWIRE–OzDES catalogue for both data-paths.

### 1.6.2 BLIND NEAREST-NEIGHBOUR CROSS-MATCHING: OzDES–DES

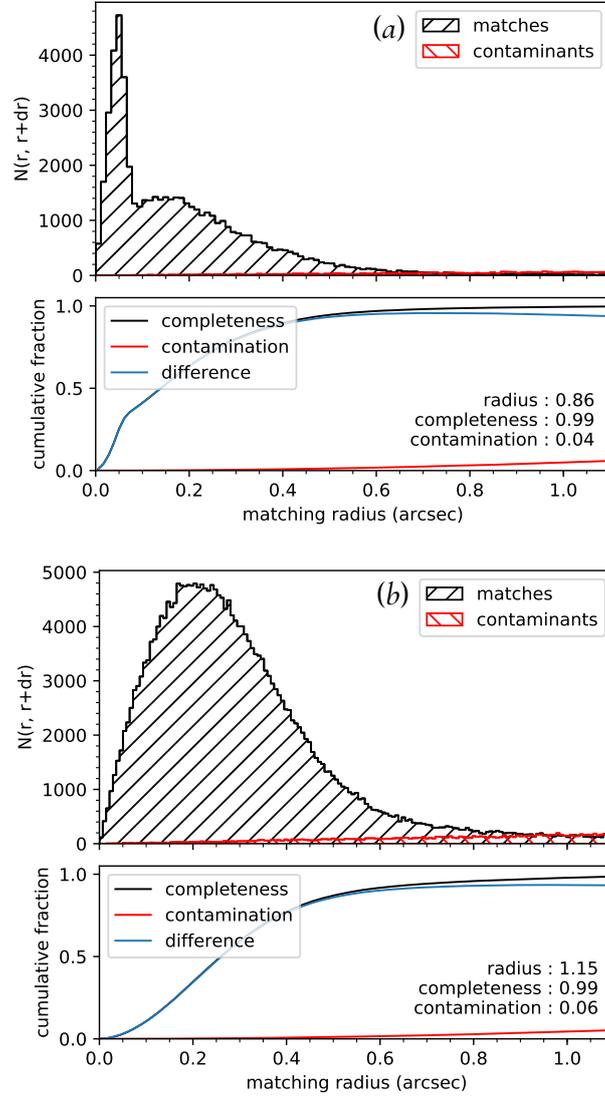
Because OzDES data are from targeted follow-up of DES photometric sources we expect to see reliable BNN cross-match pairing of these two catalogues. This result is confirmed by Figure 1.5: a visualisation of the number of matches obtained per matching radius bin for a blind nearest-neighbour search. For these data, we can obtain practically 100 per cent completeness with only 4 per cent contamination, the matching radius required to achieve this is just 0.86". As an aside, there is an interesting double-Gaussian feature (peaks: 0.05", 1.20") for the cross-matching of these catalogues. This arises from the different instrumentation used to obtain the DES photometry in the CDFS and ELAIS-S1 fields compared to those instruments used for other DES regions (Childress et al. 2017).

### 1.6.3 BLIND NEAREST-NEIGHBOUR CROSS-MATCHING: SWIRE–DES

Since I will also be cross-matching to DES photometry via the SWIRE infrared catalogue, I need to establish a similar cross-matching basis for the DES–SWIRE pairings. Figure 1.5 shows similar reliability as with DES–OzDES cross-matching:  $\sim$  100 per cent completeness with only 6 per cent contamination at a radius of 1.15"; this radius is similar to cross-matching radii used for pairing SWIRE infrared and other optical catalogues (Surace et al. 2005; Rowan-Robinson et al. 2008). These cross-matching results are sufficiently reliable that I assume an exact match of each ATLAS source to an OzDES redshift, where the appropriate matches are present in the subsequent catalogues.

### 1.6.4 BLIND NEAREST-NEIGHBOUR CROSS-MATCHING: ATLAS–SWIRE & ATLAS–DES

Now that I have established a reliable DES–OzDES, and SWIRE–OzDES redshift association to use with both of the data-paths, I can investigate the difference between BNN cross-matching and manual visual cross-matching of ATLAS. Because the precision of ATLAS source positions is much lower than either DES or SWIRE due to lower resolution and ATLAS source blending, I expect poor performance from a BNN technique. This is apparent upon inspection of Figure 1.6 which visualises BNN cross-matching in the same manner as Figure 1.5. This time instead of choosing an optimum cross-matching radius based on a radius at which we reach  $\sim$  100 per cent catalogue completeness, I have selected a radius for which we can maximise completeness below an (arbitrary) 10 per cent spurious fraction. For ATLAS–DES cross-matching we achieve only 39 per cent completeness before reaching a spurious fraction of 10 per cent (Figure 1.6), whereas cross-matching ATLAS–SWIRE this same technique can achieve 58 per cent



**Figure 1.5:** Top: OzDES-DES; Bottom: SWIRE-DES. Visualisation of the effect of increasing cross-matching radius for blind nearest neighbour (BNN) cross-matching technique. Catalogues being cross-matched are: OzDES targeted spectroscopic survey and the DES photometric catalogues (top); the SWIRE  $3.6 \mu\text{m}$  infrared and the DES photometric catalogues (bottom). In both panels, the red histogram of ‘contaminants’ involves the same cross-matching technique except with positions from one of the catalogues randomised to emulate dissociation (see § 1.6.1). The bottom pane of each panel is the cumulative sum of the results from the top pane, and the difference in this sum marked in blue. The values specified in this pane are: the cross-matching radius at which  $\sim 100$  per cent completeness is achieved, and the percentage of spurious matches at that radius.

**Table 1.1:** Comparison of the optimum radii for cross-matching our catalogues. The first column indicates the catalogues being cross-matched in the format '*primary - secondary*'. The '*Completeness Fraction*' column is the total fraction of sources with a cross-match (Equation (1.1)). The '*Spurious Fraction*' column is the fraction of matches between the primary catalogue and a randomly disassociated pairing catalogue generated from the secondary catalogue (Equation (1.2)).

	<i>Radius (arc seconds)</i>	<i>Completeness Fraction</i>	<i>Spurious Fraction</i>
<i>DES - OzDES</i>	0.86	0.99	0.04
<i>SWIRE - DES</i>	1.15	0.99	0.06
<i>ATLAS - DES</i>	2.28	0.39	0.10
<i>ATLAS - SWIRE</i>	3.86	0.58	0.10

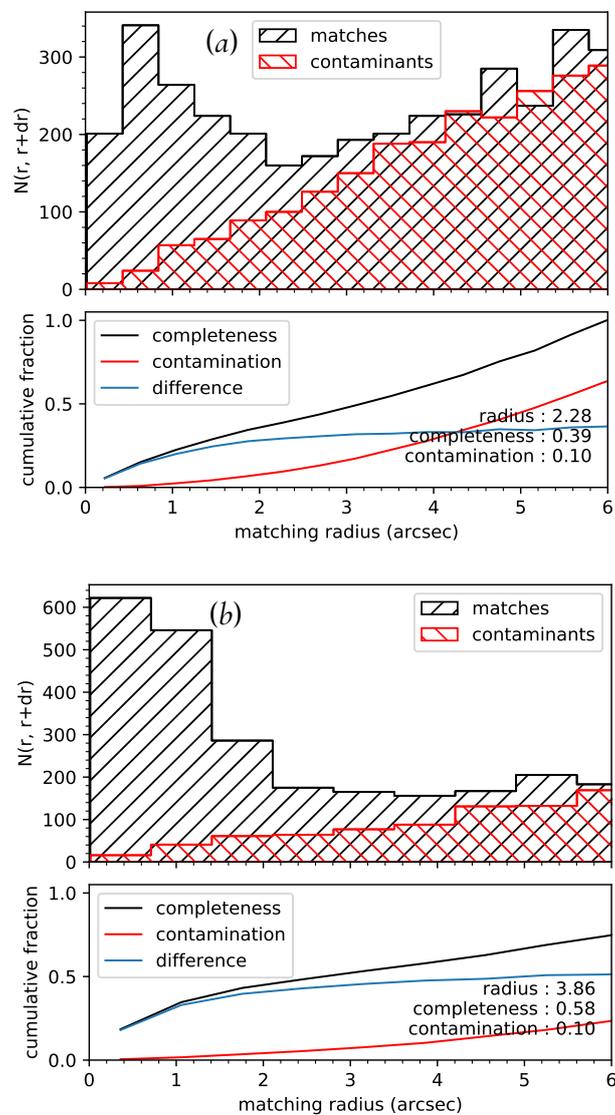
completeness. Attempting to increase completeness by increasing the matching radius would only serve to decrease the final catalogue reliability by increasing the spurious fraction. Table 1.1 summarises the optimum BNN cross-matching radii decided by this technique.

It is clear from these results that BNN cross-matching via either of these data-paths would result in a catalogue with both poor completeness and a high number of spurious matches. It is promising to see, however, that there is a definite improvement in catalogue completeness versus contamination when matching ATLAS to SWIRE compared with matching ATLAS to DES. Importantly, I have established that the reliability of cross-matching DES to OzDES and SWIRE to OzDES are nearly identical, so the improvement of using the SWIRE catalogue as an intermediate cross-matching step carries through to our final redshift acquisition. In the construction of the final BNN catalogue, I use the cross-matching radii established in this section.

## 1.7 MULTI-CATALOGUE VISUAL CROSS-MATCHING (MCVCM)

Visually inspecting each of our 5118 ATLAS radio sources and manually associating an infrared or optical counterpart is arduous and requires precise book-keeping. There is sufficient room for user error to justify the development of a semi-automated program to manage the back-end, freeing the expert to focus on the correct classification of sources. For the sake of expediency and repeatability, it is essential to have as much of this automated as possible. Developing a tool to achieve this task that is generalised enough to be usable for the broader astronomy community has been a large part of this thesis. Whilst there have been small software programs written by others to handle facets of this task, they mostly involved the generation of visual cutouts with the intention of manual book-keeping by the user; or the expert is restricted to a yes or no confirmation of a BNN match (e.g., [Norris et al. 2006, 2007](#)). Radio Galaxy Zoo (RGZ) and LOFAR galaxy zoo ([Williams et al. 2018](#)) use closed-source professionally developed software to perform this task for their sample. This interface is designed to allow their citizen scientists to make selections without worrying about the data handling. Since RGZ is unavailable for distribution, I have written MCVCM as an open-source Python program to emulate the RGZ interface.

I designed MCVCM to facilitate simple cross-matching of paired catalogues from any data



**Figure 1.6:** Top: ATLAS-DES; Bottom: ATLAS-SWIRE. Visualisation of the effect of increasing cross-matching radius for blind nearest neighbour (BNN) cross-matching technique. Catalogues being cross-matched are the ATLAS 1.4 GHz radio and the DES photometric catalogues (top); the ATLAS radio and the SWIRE 3.6  $\mu\text{m}$  infrared catalogues (bottom). In both panels, the red histogram of 'contaminants' involves the same cross-matching technique except with the positions from one of the catalogues randomised to emulate dissociation. The bottom pane of each panel is the cumulative sum of the results from the top pane, and the difference in this sum marked in blue. The values specified in this panel are: the cross-matching radius at which we have 10 per cent spurious matches and the completeness of source cross-matching at that radius

source. Provided that the images and catalogues are in standard `'fits'` \* format, the end user need only provide for each field:

- radio continuum image file,
- radio RMS image file,
- optical/infrared image file,
- radio positions catalogue (primary),
- optical/Infrared positions catalogue (secondary).

Once configured to read these files, MCVCM will produce an image cutout centred on the first source in the primary catalogue <sup>◊</sup>; this cutout will appear as radio contours overlaid onto an optical or infrared image <sup>‡</sup>. Across three phases, the user will select an infrared galaxy core (if there is an appropriate candidate) followed by any radio peaks associated with that core. The user performs this selection by clicking marks representing the right ascension and declination of the catalogue sources (see Figure A.1d Page 80). During cross-matching, the user can increase the field of view of the cutout to better gauge the environment. Optionally, they may also enter an integer confidence flag at any point ( $\in \{1, 2, 3, 4\}$ ), as well as a comment string for later reference. Once the user is satisfied and moves to the next source, MCVCM will append to a new catalogue file all information from the primary catalogue for the selected radio sources, the user flag and user comment, as well as a unique association string used to link all radio components to the infrared host selected from the secondary catalogue. The MCVCM software includes a suite of functionality to facilitate better cross-matching and to ensure no loss of progress in the event of a software or computer crash. See Appendix § A for a thorough breakdown of MCVCM, including instructions, examples, the web address of the online repository from which it can be accessed.

### 1.7.1 CONSIDERATIONS OF VISUAL CROSS-MATCHING

Selecting the correct infrared or optical counter-part for our ATLAS sources is vital for developing a rigorous cross-matched catalogue that can be used as a golden standard for training intelligent automated cross-matching systems. Suppose a radio component is erroneously associated with a brighter, low-redshift infrared or optical counterpart. In this case, the radio luminosity is underestimated (as the measured flux is now attributed to a closer source) and the number of nearby low-luminosity sources is artificially increased. Additionally, a bright radio

---

\*Flexible Image Transport System (Calabretta & Greisen 2002).

◊The primary catalogue is generally the catalogue with the fewest sources. In the case of this thesis that is the ATLAS 1.4 GHz component catalogue. This is not a strict requirement of MCVCM, but it does reduce the number of iterations for the user.

‡For the remainder of this thesis we will consider the MCVCM pairing catalogue and images to be 3.6  $\mu\text{m}$  SWIRE infrared, except where I am discussing the visual cross-matching of ATLAS-DES.

source is removed from a higher redshift bin. If instead, the user is to manually ignore such a source (that is, treat it as a non-detection) then the reliability of our catalogue increases. When I construct radio luminosity functions in § 3 volume corrections of the population statistics will be used to correct for these non-detections, making this the preferred option.

I found that during the visual cross-matching there is a temptation to try to match as many sources as possible, perhaps in an unconscious attempt to maximise matching completeness. However, it is essential when making these cross-match selections to not ‘select the closest’ source; careful consideration must be made for the topography. Occasionally there will be a radio source with an aligned infrared source that is visible in MCVCM, but has fallen below the detection limit of the SWIRE survey (e.g., Figure 1.2a). It is very likely that this undetected faint source is indeed the true host, and if one is to select a nearby source that is merely ‘close enough’, then we run into a problem of ‘false’ cross-identification (see § 1.4), it is instead better practice to leave this as an unmatched source.

Because the DES optical survey contains a higher density of sources than 3.6  $\mu\text{m}$  SWIRE infrared survey (80 000  $\text{deg}^{-2}$  versus 31 000  $\text{deg}^{-2}$ ) there are many more ‘chance alignments’ for a given ATLAS source when visually cross-matching to DES; therefore, the chance of making a false cross-identification is increased: even if one is careful to avoid them. As discussed in § 1.5, it is common practice when conducting BNN cross-matching to make a magnitude cut for the optical catalogue to reduce the number of potential candidates for a spurious match. However, one of the benefits of manually cross-matching is that it alleviates the need to impose potentially biasing magnitude assumptions upon the data. For this reason, I have retained all sources from each catalogue and made cross-matching selections purely based on the visual information.

I also expect that this excess of sources will increase the variability of final cross-matched catalogues if a single user or many users perform several cross-matching runs. Banfield et al. (2015) developed a ‘consensus algorithm’ that weighted contributions from a large user-base, unfortunately with only one complete visual cross-matching pass for both ATLAS–SWIRE, and ATLAS–DES<sup>11</sup> I am unable to quantify variability errors.

### 1.7.2 DIFFICULTIES WITH VISUAL CROSS-MATCHING OF ATLAS–DES

Manually cross-matching ATLAS against DES requires that the collection of  $1' \times 1'$  DES optical image tiles are mosaicked together<sup>12</sup>. Without mosaicking it would be very difficult to verify larger multi-component structures as they invariably exist across several tiles (e.g., Figure 1.7). Comparatively, once SWIRE image tiles are mosaicked they are full-field with no gaps. Once

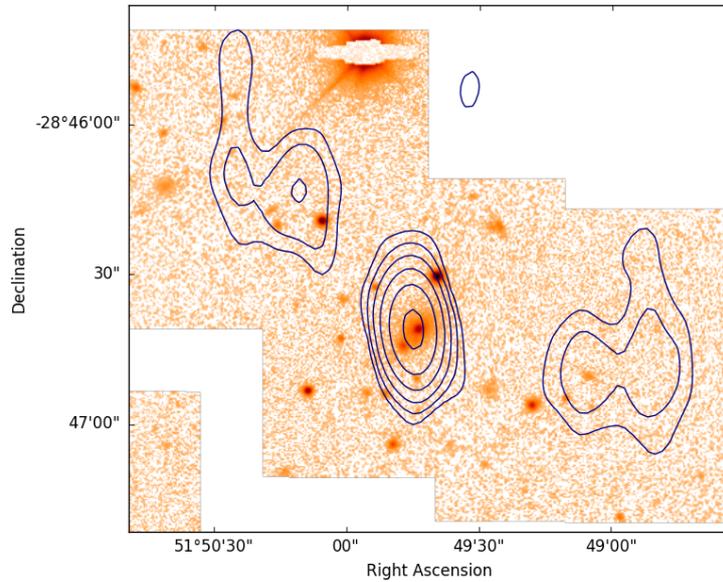
---

<sup>11</sup>Through the course of developing MCVCM I did perform many partial passes, but these were not included in the final analysis as I wanted to be certain that I was cross-matching with a consistent technique and error tolerance.

<sup>12</sup>Mosaicing is performed using montage (Jacob et al. 2010; Berriman et al. 2016, <http://montage.ipac.caltech.edu>).

the data is in the appropriate format I followed the procedure outlined in § 1.7 for the 5118 ATLAS radio components.

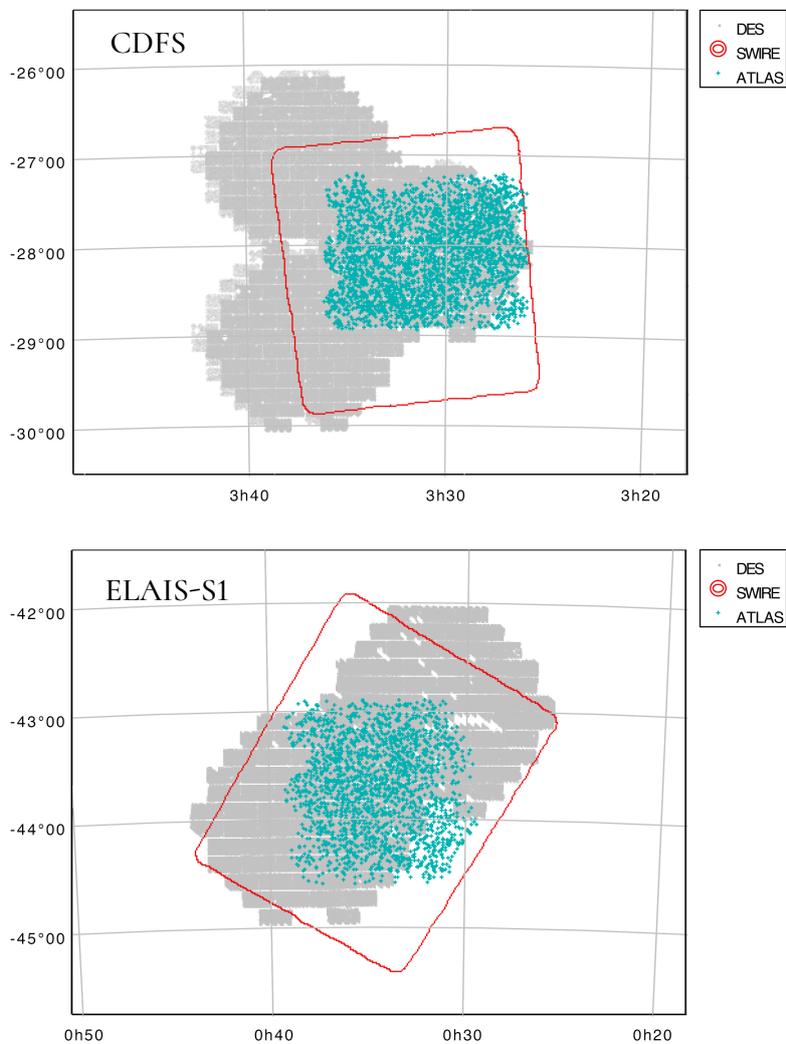
During MCVCM cross-matching of ATLAS–DES there were many cases when the density of the DES catalogue source detections made the selection of a likely optical galaxy difficult, in these cases, I deferred the selection to the brightest of the candidates. There is also a substantial difference in field overlap of SWIRE and DES with the ATLAS survey (see Figure 1.8). Many of the visual ATLAS–DES cross-matches had to be skipped as there were no available DES images or sources in these regions, this would compromise the performance of the ATLAS–DES cross-matching, compared to ATLAS–SWIRE for which the overlap is more complete.



**Figure 1.7:** Optical MCVCM cutout of ATLAS source CI0304. 1.4 GHz ATLAS radio continuum contours are overlaid onto DES ( $1' \times 1'$ ) mosaicked  $I$ -band tiles. Radio contour levels are  $2^n \times \langle \text{RMS} \rangle$  calculated for this field of view (see Figure 1.2(d) Page 15 for the  $3.6 \mu\text{m}$  infrared cutout of this source).

### 1.7.3 COMPARING CROSS-MATCHING TECHNIQUES

Comparing the results from the techniques can only be made internally as I have no means of determining the validity of all matches. Empirically, there are fewer spurious matches when BNNcross-matching via the SWIRE data-path, and there is a better accounting of blended and complicated sources when visually cross-matching. As discussed in § 1.4 it is expected that there will be a better pairing of ATLAS–SWIRE than ATLAS–DES due to better matching of survey depth, emission processes, and resolution. For these reasons, I will use the visual ATLAS–SWIRE cross-matching results as the standard for comparison of the other technique and data-path combinations.



**Figure 1.8:** Overlap of ATLAS, DES, and SWIRE for ELAIS-S1 and CDFS. The red line marks the bounding edge of SWIRE, the grey points are DES source detections, and the cyan points ATLAS source detections. There are 259 (13 per cent of 1989) ATLAS-ELAIS-S1 sources that are outside of the DES field, and 166 (6 per cent of 2914) ATLAS-CDFS sources outside of the DES field.

From the initial 5118 detections in the ATLAS catalogue MCVCM cross-matching (via DES and SWIRE) verifies 4903 individual galaxies of which 137 are extended or multi-component sources. Visual cross-matching against the SWIRE infrared catalogue results in an infrared counter-part for 95 per cent (4658) sources. In Table 1.2 I have treated these 4658 sources as correct for comparison of techniques and data-paths<sup>2</sup>. Table 1.2 compares how many of the matches from each of the other methods would result in an identical ATLAS–SWIRE–DES–OzDES pairing.

**Table 1.2:** Comparison of the quality of the four combinations of cross-matching techniques and data-paths: as indicated by the first two columns. The results from the visual cross-matching of ATLAS to SWIRE (shown in the first row) is used as our standard of comparison. The *Total* column is the number of ATLAS sources with a match via the associated method. The *‘Correct’* column is the total number of sources that have the same final infrared counterpart as *‘visual - SWIRE’*. The *Correct Fraction* column is this ratio of *‘Correct’* matches and the *Total* matches. The *Completeness* column is the ratio of *‘Correct’* to the total ATLAS source count (4903). Note that for the two BNN techniques I have already restricted the cross-matching radius to limit spurious matches to 10 per cent (see Figure 1.6) this will increase the reliability of the final catalogue at the expense of completeness.

<i>Data-path</i>	<i>Technique</i>	<i>Total</i>	<i>‘Correct’</i>	<i>Correct Fraction</i>	<i>Completeness</i>
<i>ATLAS–SWIRE</i>	Visual	4658	<b>4658</b>	-	0.95
<i>ATLAS–SWIRE</i>	BNN	2185	2135	0.98	0.44
<i>ATLAS–DES</i>	Visual	3905	2776	0.71	0.57
<i>ATLAS–DES</i>	BNN	1972	1870	0.95	0.38

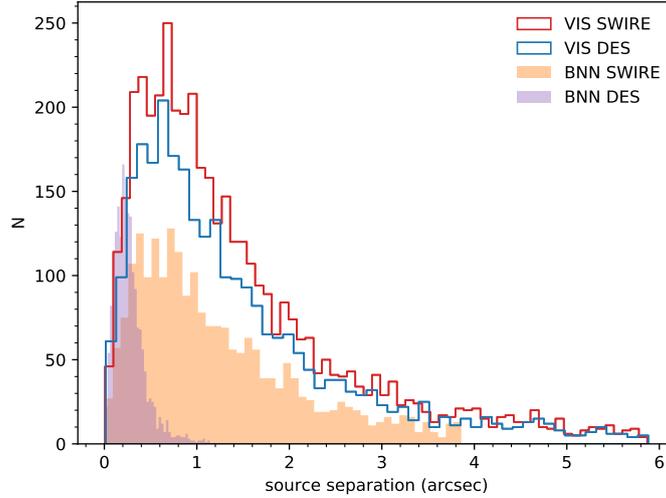
The first point to note is that regardless of the technique used, cross-matching via the 3.6  $\mu\text{m}$  SWIRE infrared catalogue yields a greater number of total matches, and a better *‘Correct Fraction’*, in the case of BNN.

Another notable point is that regardless of the data-path, visual cross-matching results in many more total matches. Of course, increasing the matching radius for BNN would increase the total number of matches for this technique, but as established in § 1.6 this would only serve to compromise the final sample with spurious matches. Ideally, visual cross-matching would result in 100 per cent reliability, but that is not the case.

Inspecting the correct fraction of visual matches for the ATLAS–DES cross-matching, almost 30 per cent of the matches resulted in an incorrect final SWIRE infrared core<sup>3</sup>. In § 1.5 I argued that having  $\sim 300$  per cent more sources per square degree in the DES survey compared to the SWIRE survey would result in many more spurious matches. Figure 1.9 is a visual demonstration of these effects; because the OzDES survey contains a much higher source density, and the data is high-resolution (compared to ATLAS) the BNN algorithm can find a match for each ATLAS source within a small search radius, regardless of validity. The MCVCM cross-matching (i.e., VIS SWIRE, and VIS OzDES) illustrates the *‘true distribution’* of source separations for

<sup>2</sup>I by no means assume that visual matching is 100 per cent reliable, I assume that it is more reliable than BNN and merely use the results as a basis for comparison

<sup>3</sup>Recall that the SWIRE–DES association is established in § 1.6.3



**Figure 1.9:** Comparison of separation between ATLAS 1.4 GHz source position and the respective SWIRE or OzDES positions for each data-path: visual cross-matching (MCVCM), or blind nearest-neighbour (BNN).

these data due to the nature of visual cross-matching (that is, the low-resolution of ATLAS 1.4 GHz data is accounted for during visual inspection). Comparing the source separations for SWIRE BNN to these ‘true distributions’ shows closer agreement than OzDES BNN, suggesting that better matched survey source density and resolution allows for the a better chance of a final BNN matched catalogue with few spurious matches.

It is worth noting that the order in which I did the MCVCM cross-matching may have affected the visual cross-matching results: I cross-matched ATLAS–DES, and later cross-matched ATLAS–SWIRE. It is likely that I acquired visual biases during the first cross-matching that I inadvertently applied to the second cross-matching. This may have marginally amplified the ratio of erroneous matches between data-paths, but would require several iterations of cross-matching by independent experts to quantify these biases.

## 1.8 FINAL ATLAS RADIO CATALOGUE

I have established that the cross-matching method that yields the highest completeness and accuracy of optical–radio source association involves an infrared intermediate step. This holds true irrespective of whether the cross-matching is done using automatic positional techniques, or through manual visual classification of each source. This could be a manifestation of galactic processes contributing to significant radio emission also producing significant infrared emission, but do not reliably produce significant optical emission. Alternatively, it could be largely due to the difference in observational depth of the SWIRE and DES catalogues. We are lucky to have radio data that is paired excellently in observational-depth and resolution (Franzen et al. 2015; Surace & et al. 2012) with the SWIRE infrared observations; were this not the case it is unlikely that we would see such impressive improvement when cross-matching via the SWIRE

catalogue compared to directly cross-matching to DES optical catalogue.

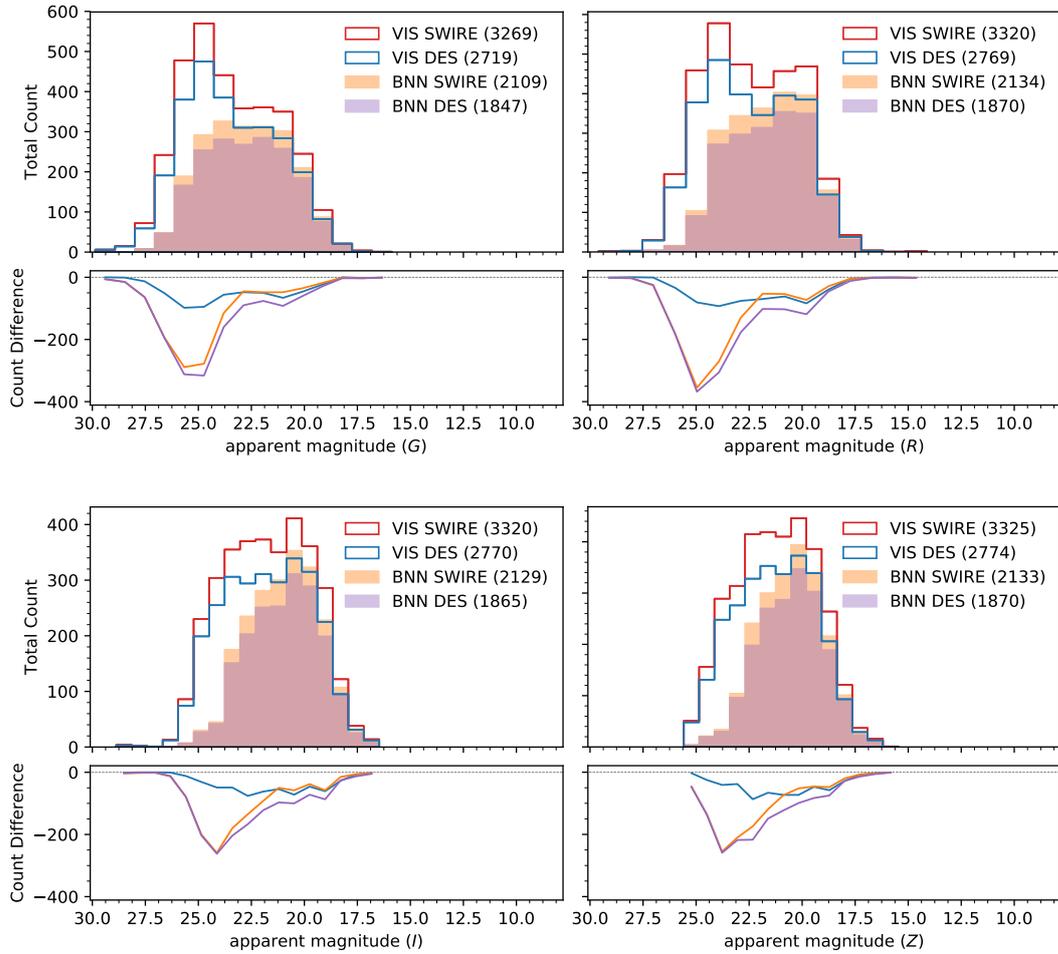
### 1.8.1 PHOTOMETRIC AND SPECTROSCOPIC COUNTERPARTS

For each of the DES photometric bands, Figure 1.10 shows the breakdown of matches to the ATLAS radio catalogue separated into the cross-matching techniques and data-paths: visual SWIRE–ATLAS; visual DES–ATLAS; BNN SWIRE–ATLAS; BNN DES–ATLAS. In all cases, visual ATLAS–SWIRE cross-matching performs the best which is likely due to the similar galactic emission processes producing radio and infrared emission that have been discussed above, as well as the lower surface density of SWIRE sources compared to DES sources decreasing the frequency of spurious matches.

### 1.8.2 CONCLUSION

The next generation of radio sky surveys will yield on the order of millions of sources as modern instrumentation improves data quality and quantity (Norris 2017). For the several thousand sources available in the ATLAS DR3 survey, cross-matching multi-frequency catalogues via manual source inspection is both practical and beneficial for the integrity of the data. For samples consisting of a few tens of thousands of sources, the resource commitment to such a task becomes questionable, and for catalogues with millions of sources, it becomes futile. To accommodate the next generation of large data various techniques are being developed to automate or delegate such time-intensive tasks. These techniques which include citizen science (Banfield et al. 2015) and machine learning (Lukic et al. 2018) require training data or comparison data to verify their legitimacy. Having now developed a sample of several thousand reliably cross-matched sources, we have a better groundwork on which to build this software.

I have demonstrated the value of visual cross-identification for both extended and compact sources, as well as the improvement in cross-matching a catalogue of deep 1.4 GHz radio data (ATLAS) to an optical spectroscopic catalogue (OzDES) via an infrared catalogue (SWIRE) instead of directly to the optical catalogue. I make use of this final multi-frequency catalogue in subsequent chapters.



**Figure 1.10:** Per-bin DES photometry source counts for the cross-matching methods discussed for each of the four DES photometric bands. The top pane of each panel shows the histogram of counts drawn from the 'Correct' sample established in Table 1.2. The bottom panels are the difference in counts (per bin) between the visual SWIRE-ATLAS cross-matching and the other methods.

*If people sat outside and looked at the stars each night, I'll bet they'd live a lot differently.*

– Calvin (& Hobbes)

# 2

## Diagnostic Population Segregation of ATLAS Radio-Galaxies

### 2.1 DISCERNING GALAXY CHARACTERISTICS

Absolute segregation of galaxies into sub-populations is inherently an unachievable goal. Galaxies do not occupy discrete population spaces for which there exists an obscure combination of observables that will allow us to unlock their real identities. Not only are galaxies the most complex systems in the Universe but they are the systems for which we have the least (and most contaminated) information. Electromagnetic radiation, from its various origins, is the only source of information for understanding these incomprehensibly vast celestial bodies. This light originates from a multitude of mechanisms and travels to our single point of reference over the course of aeons over which time and space it can have countless interactions obfuscating or obliterating the original information.

Because light has a finite and constant velocity we experience a look-back time and decreasing information quality of our observations as a function of distance; this concept translates well to our understanding of human history. Modern computers and record keeping are far superior to entirely paper-based systems of the early twentieth century, and this trend continues until we are making inferences about the Mesopotamians from broken pottery and charred bones. In both the terrestrial and the celestial case, the quality and abundance of our data de-

crease as our temporal delta increases.

Despite the shortcomings of light as a sole information source, it provides us with a wealth of knowledge through clever inferences based on the presence or lack of particular characteristic frequencies in a given environment. Despite absolute segregation of galaxy populations being wholly impossible, partial segregations of increasing complexity and reliability lead to a clearer understanding of galaxy formation and by extension, the history of our vast Universe.

## 2.2 CATEGORISING GALAXIES

Determining if star formation (SF) or an active galactic nucleus (AGN) is the dominant emission mechanism of a galaxy is commonly performed using optical spectral emission line diagnostics, infrared photometric diagnostics, optical photometric diagnostics, x-ray diagnostics, radio continuum flux-density, polarisation, Very Long Baseline Interferometry (VLBI) core detections or combinations of the above. Each of these techniques is sensitive to different populations and stellar ages, resulting in inherent assumptions if they are used in isolation. For example, x-ray diagnostics only identify  $\sim 50$  per cent of BPT AGN (Leahy et al. *submitted*), spectral diagnostics and infrared photometry are dependent on galaxy orientation altering line-of-sight extinction (Sanders et al. 1988), VLBI detections are a reliable confirmation of an AGN, but only 20 per cent of radio-loud AGN are visible in VLBI (this detection fraction is unrelated to sensitivity; Middelberg et al. 2013; Herrera Ruiz et al. 2017), and a non-detection does not guarantee an SF galaxy.

### 2.2.1 BPT DIAGRAM

The most broadly used spectral line diagnostic is the BPT diagram (Baldwin et al. 1981) which plots the emission line flux ratio of  $[\text{OIII}]\lambda 5007/\text{H}\beta$  against that of  $[\text{NII}]\lambda 6584/\text{H}\alpha$ . The population segregation relies on the difference in excitation energy of photons from the ultra-violet fields of young stellar populations, and that of the ionising regions surrounding an active galactic nucleus. Since  $[\text{NII}]$  and  $[\text{OIII}]$  are collisionally excited forbidden emission lines, and  $\text{H}\alpha$  and  $\text{H}\beta$  are recombination lines, there is a theoretical upper limit on the ratios  $[\text{OIII}]/\text{H}\beta$  and  $[\text{NII}]/\text{H}\alpha$  from excitation by star formation. Because AGN have a much higher upper limit on the energy of colliding photons, these ratios can increase well beyond the star-forming threshold derived by Kewley et al. (2001). Consequently, AGN will tend to dominate the upper right of the BPT diagram, with star-forming galaxies forming a locus in the lower-left.

Kewley et al. (2001) and Kauffmann et al. (2003) have developed the most complete and reliable division of galaxy classes. Kauffmann et al. (2003) produced an empirical AGN segregation using data from the Sloan Digital Sky Survey (SDSS; York et al. 2000), Kewley et al. (2001) theoretically derived a ‘maximum star-burst’ line based on the excitation energies outlined above. Galaxies that lie between the SF and AGN lines are referred to as ‘composite galaxies’; that is, their optical spectrum is assumed to be excited by both star-formation and AGN accretion. Strictly viewing these galaxies as exhibiting ‘blended’ emission is a misunderstanding of

the limitations of the BPT diagram. As instrumentation improves, and more robust diagnostics are established, it becomes clear that both strongly AGN and star-forming dominated galaxies can occupy this space.

The prominent short-coming of the BPT diagram is its redshift limitation. For a typical optical spectral window reaching  $\sim 9200 \text{ \AA}$ , [NII] $\lambda 6584$  is redshifted out of the spectrum beyond  $z > 0.4$ . Extensive work has gone towards the empirical and theoretical calibration of the BPT diagnostic diagram beyond  $z > 0.4$  using a combination of optical and infrared spectroscopy (e.g., Kewley et al. 2013). Unfortunately, the difficulty of being awarded sufficient observing time on an infrared and an optical telescope for the same project, the lower S/N of infrared spectroscopy, and the different instrument calibrations (required for merging spectra) prevents the use of these results from becoming ubiquitous. Additionally, high-redshift galaxies tend to be dusty and optically-thick, removing the utility of BPT and other spectral-line diagnostics as a classification tool. In galaxies such as these, which are frequent at high redshifts, line-diagnostics only measure the emission from the outer-most shell which is generally star-forming even in AGN (Assef et al. 2015; see § 0.5 for a discussion and examples).

### 2.2.2 ALTERNATIVE OPTICAL DIAGNOSTICS

In an attempt to bypass the redshift limitations of the BPT diagnostic diagram, significant work has gone into producing variations or alternatives (see Zhang & Hao (2018) for a comprehensive summary). These diagnostics vary the recipe of emission characteristics by using bluer spectral emission lines, photometric colours, or stellar properties. Lamareille (2010) replace the redshift limiting [NII]/H $\alpha$  ratio with [OII] $\lambda\lambda 3726, 3729$ /H $\beta\lambda 4861$ . This combination of ‘blue’ emission lines (hereafter blue-diagram) boasts 99 per cent star-forming galaxy agreement with BPT; however, it does not reliably segregate SF galaxies from Seyfert galaxies or BPT composites and is best used in conjunction with additional diagnostics, such as the  $D_n[4000]$  4000  $\text{\AA}$  continuum break (Stasińska et al. 2006; Marocco et al. 2011). The Mass-Extinction diagnostic developed by Juneau et al. (2011) uses stellar mass and [OIII]/H $\beta$ , this diagnostic was demonstrated on a sample of galaxies across  $0.3 < z < 0.9$ . Yan et al. (2011b) switch [NII]/H $\alpha$  for rest-frame  $U - B$  focusing on the metallicity–colour relation of AGN and star-forming galaxies. Both of these diagnostics achieve  $> 95$  per cent agreement with BPT AGN; however, the use of H $\beta\lambda 4861$  limits these diagnostics to  $z < 0.9$ .

Trouille et al. (2011) developed a diagnostic that focuses on the metallicity disparity of star-forming and AGN dominated galaxies (hereafter, TBT diagram), utilising  $g - z$  colours in combination with [NeIII]/[OII] as segregation. [NeIII] is an excellent tracer of high-excitation AGN, indicating an abundance of highly ionised gas more strongly than [OII], and metal-poor AGN will appear redder in  $g - z$  than bulge-dominated star-forming galaxies (Yan et al. 2011a). Importantly, [NeIII] $\lambda 3869$  and [OII] $\lambda\lambda 3726, 3729$  are close in wavelength, providing diagnostic coverage to  $z < 1.3$ . TBT was tested on a highly spectroscopically complete X-ray selected set of AGN (OPTX; Trouille et al. 2008, 2009; Trouille & Barger 2010) and achieved 98.7 per cent BPT-AGN, and BPT-SF agreement.

Mid-infrared (MIR) classification wedges (e.g., [Lacy et al. 2004](#); [Stern et al. 2005](#); [Lacy et al. 2007](#)) are used to identify obscured AGN by tracing AGN-heated dust (e.g., [Fiore et al. 2008, 2009](#); [Assef et al. 2013](#)). Inevitably this selection suffers heavily from contamination by star-forming galaxies as it traces the re-emitted radiation absorbed from other wavelengths ([Sajina et al. 2005](#)). [Donley et al. \(2012\)](#) developed a complex Spitzer IRAC criteria for selection of AGN using MIR selection wedges, and infrared power-law selection, then filtering out high-redshift star-forming galaxies via  $B$ -,  $z$ -, and  $K$ -band photometry ([Daddi et al. 2004](#)), distant red-galaxy identification ([Franx et al. 2003](#)), Lyman-break galaxy selection (e.g., [Steidel et al. 2003](#)), and submillimeter galaxy selection ([Menéndez-Delmestre et al. 2009](#)).

In this chapter, I perform classifications based upon the visual inspection of the available OzDES spectra and compare these to classifications using BPT and [Lamareille \(2010\)](#) blue-diagram. This combined classification will be compared with the radio-far-infrared correlation as a final consistency check. I construct a simple linear combination of each classification method weighted by the distance to their respective population division lines. I establish population segregation of SF and AGN galaxies from the ATLAS sample for use in the remainder of this thesis.

### 2.3 OZDES REDSHIFTS

OzDES data primarily consists of spectroscopic redshifts from the OzDES Fibre Programme but also includes redshifts from the literature where available. Table 2.1 shows a summary of the redshifts and spectra contained in the ATLAS subset of OzDES DR1. Each of these surveys employed a unique flag for the quality of a redshift measurement. I standardised these quality flags based upon the qualitative descriptions from the relevant data releases; this standardisation is outlined in Table C.1 (Page 84; moving forward this combined flag is referred to as  $cflag$ ). Throughout this thesis I limit sources  $cflag \geq 2$  to increase the reliability of our final populations.

### 2.4 SPECTRAL FITTING WITH pPXF

The OzDES spectra were analysed with the 2017 Python implementation of Penalized Pixel-Fitting (pPXF; [Cappellari & Emsellem 2004](#); [Cappellari 2017](#)) using the incorporated MILES stellar population models ([Vazdekis et al. 2010](#)). pPXF requires that the spectrum being fitted is of the same resolution and wavelength span as the stellar models. The Python package `pysynphot` ([STScI Development Team 2013](#)) can accurately perform this rebinning of each spectrum to these specifications whilst preserving flux. Once rebinned, each spectrum was passed through a modified pPXF pipeline and the emission-line fluxes recorded. Figure 2.1 shows an example of a spectrum fitted by pPXF.

**Table 2.1:** Summary of redshift and spectra contributions to the ATLAS subset of OzDES DR1 from various instrumentation and surveys. This table summarises the number of redshifts from each source, the number of optical spectra, and the median redshift of these subsets.

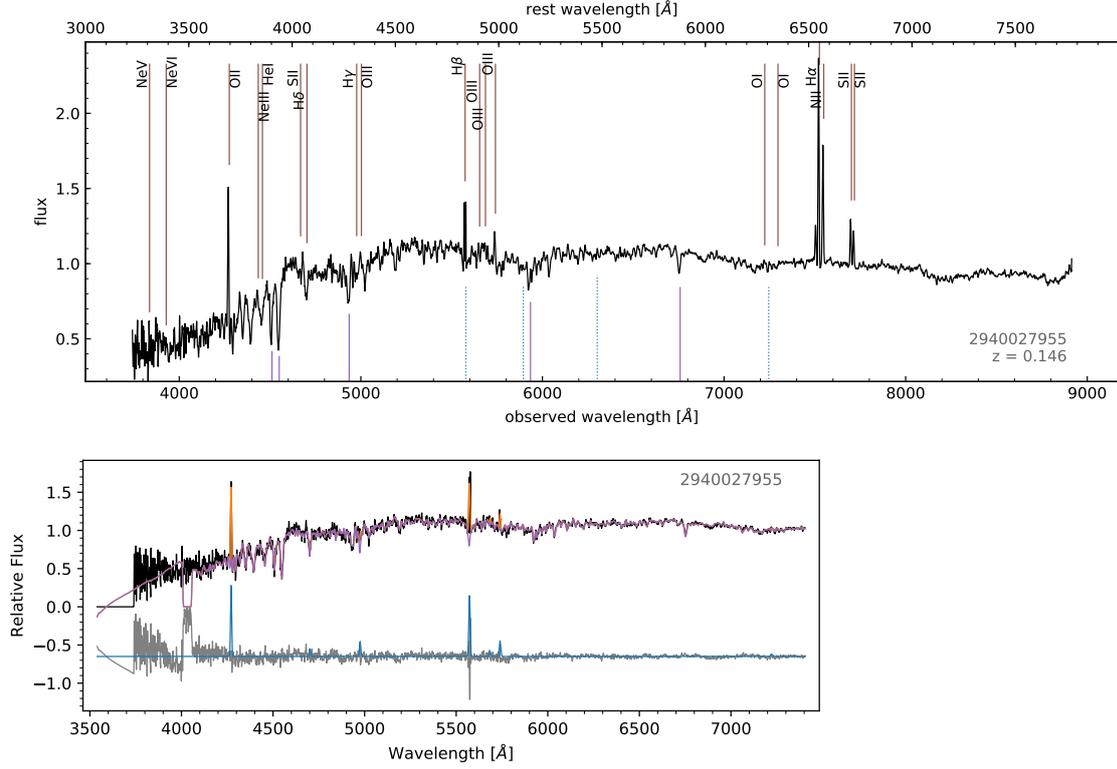
<i>Source</i>	<i>Redshift Count</i>	<i>Spectra Count</i>	<i>Median Redshift</i>
2dFGRS	81	34	0.14
6dF	16	0	0.054
ACES	30	0	0.64
ATLAS	589	122	0.32
DES_AAOmega	912	765	0.59
NED	4	4	1.4
NOAO_0334	10	0	0.30
NOAO_0522	34	0	0.25
PRIMUS	196	0	0.57
SpARCS	1	0	1.1
VVDS_CDFS	5	0	0.21
Total	1851	925	0.41

## 2.5 SPECTRAL DIAGNOSTIC CLASSIFICATION

Recall from § 2.2.1 that classification with BPT requires the four emission lines [NII], [OII],  $H\alpha$ , and  $H\beta$ . For the OzDES wavelength range ( $\lambda \in \{3300, 9200\}$  Å) this restricts the maximum redshift to  $z < 0.4$  before the [NII] is redshifted out of this spectral window. The rebinning required that OzDES spectra be truncated to  $\lambda \in \{3300, 7400\}$  Å to match the MILES continuum wavelength span; unfortunately, this unavoidable truncation severely reduces the number of sources that pPXF can fit. Figure 2.1 shows the difference between the entire OzDES spectrum, and the rebinned spectrum fitted in pPXF for a source at  $z = 0.146$ . With a maximum wavelength of 7400 Å, BPT diagnostics are restricted to  $z < 0.13$ ; this includes only 63 (7 per cent) sources, of which 55 (6 per cent) were successfully fitted by pPXF. The blue-diagram is similarly restricted to  $z < 0.52$ ; yielding 439 (47 per cent) sources, with 329 (36 per cent) successfully fitted by pPXF.

## 2.6 VISUAL SPECTRAL CLASSIFICATION

To compensate for the restrictions imposed by the stellar models used for the pPXF spectral fitting, I endeavoured to construct a set of visual classifications for the 925 ATLAS sources with a redshift and an optical spectrum. To do this, I wrote a light-weight interactive Python script that iterated through each spectrum and allows the user to select from several classification options: Elliptical (E), Elliptical with strong [OII] $\lambda\lambda 3726, 3729$  (E+[OII]), AGN emission line (AGNe), AGN broad line (AGNb), Star-forming (SF), or unknown (UNK); these classifications follow the nomenclature outlined in Mao et al. (2012). The user may also specify a confidence

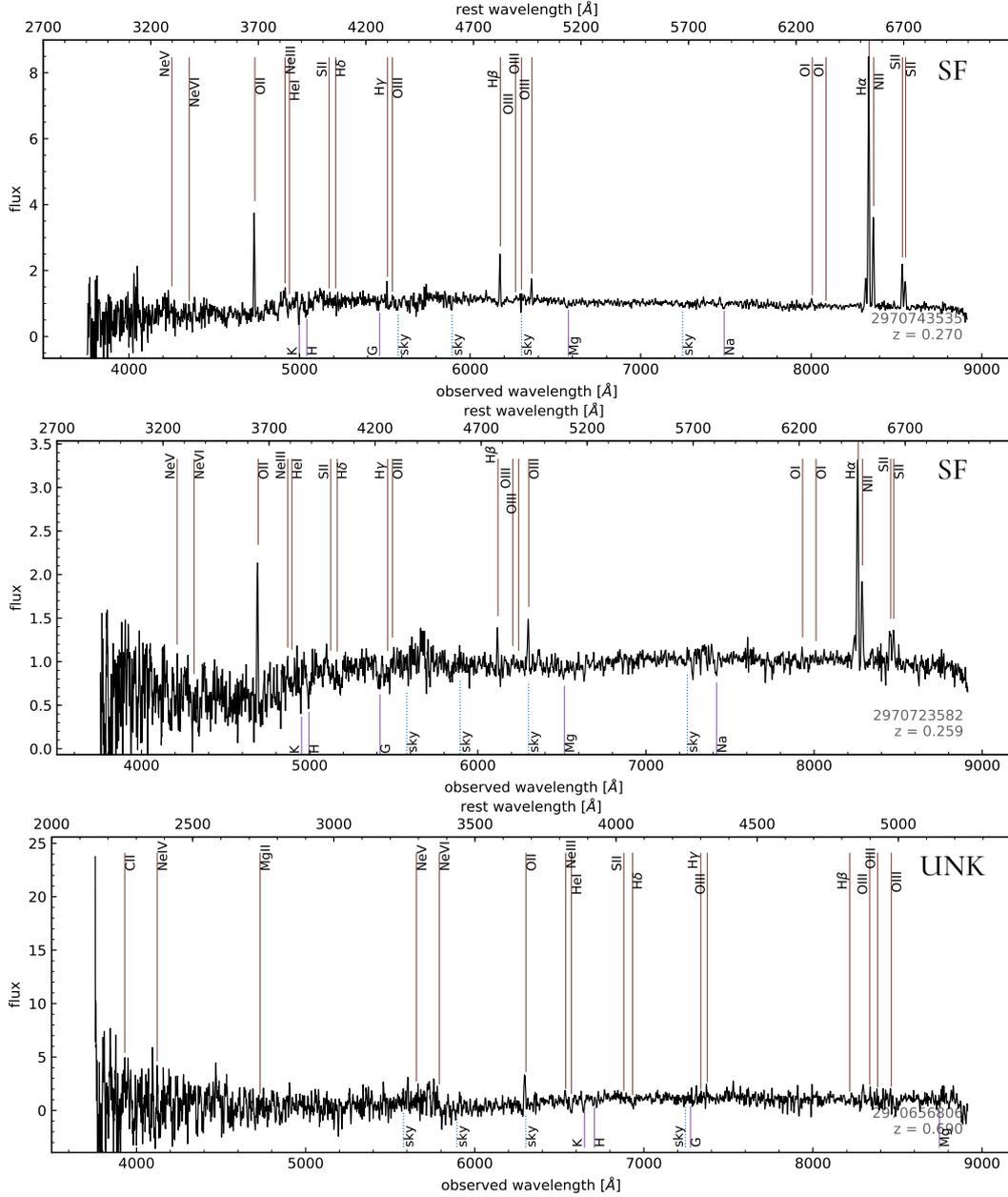


**Figure 2.1:** comparison of an OZDES spectrum as displayed for visual classification (Top), and the same spectrum fitted by pPXF (bottom). The top plot has been smoothed with a 3 pixel box-car convolution but is otherwise unmodified, with various spectral line locations marked. The bottom plot has been modified to match the resolution and wavelength span of the MILES stellar continuum library (note the truncation of [NII], H $\alpha$ , and [SII]). In this plot, the spectrum is displayed in black, with the fitted continuum in purple, fitted spectrum in orange, residuals of continuum fit in grey, and spectral line fits in blue.

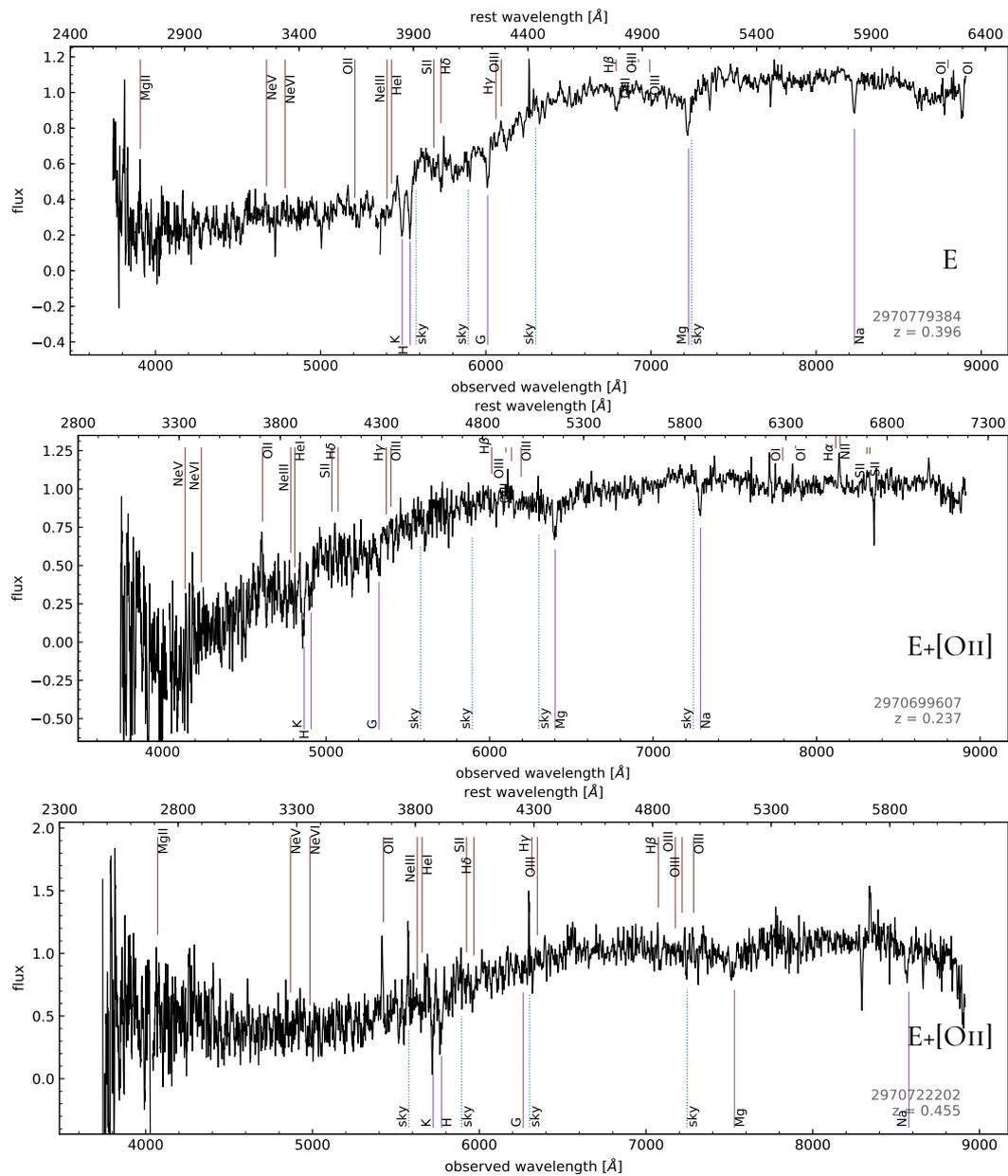
integer  $c \in \{1, 2, 3, 4\}$  for quality control. During visual classification, each spectrum was overlaid with the locations of relevant spectral features for visual reference.

Figures 2.2, 2.3, and 2.4 show examples of the various visual classifications. Figure 2.2 shows the spectra of two star-forming galaxies ( $z = 0.27$ , and  $z = 0.26$ ) and a spectrum classified as visually unknown due to the lack of distinguishing spectral features. Figure 2.3 shows the spectra of three elliptical galaxies ( $z = 0.396$ ,  $z = 0.237$ , and  $z = 0.455$ ). The first has no observable [OII] emission, the second spectrum has strong [OII] emission, and the third is an example of a spectrum with sufficient spectral features for a confident classification despite a low S/N and moderate redshift. Figure 2.4 shows the spectra of three sources classified as AGN ( $z = 0.377$ ,  $z = 0.878$ , and  $z = 1.70$ ); the first is a spectrum for which [NII] has been redshifted beyond the edge of the spectral window, but still is characteristic of an emission-line AGN. The second figure is a bright broad-line emission galaxy at moderate redshift, and the third was chosen as an example because, despite the noisy spectrum, it shows faint broad-line features allowing for a confident AGN classification. For the remainder of this thesis, I will consider

Elliptical, Elliptical with strong [OII], AGN emission line galaxies, and AGN broad line galaxies as AGN and will no longer distinguish between these AGN subclasses. Additionally, due to the low number of sources classified as AGN, I do not construct populations of HERGs and LERGs.

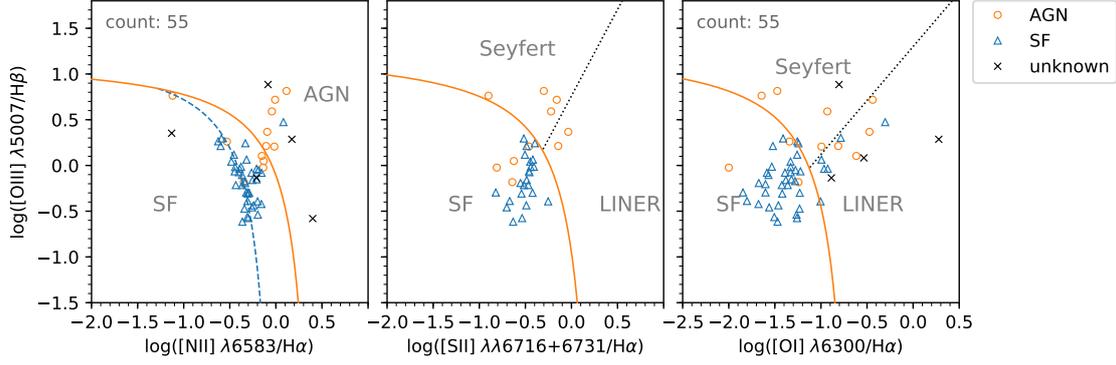


**Figure 2.2:** Three examples of OzDES spectra as they appear when classified visually. For each source, OzDES source designation and redshift are displayed in the lower-right, rest-frame wavelength is shown on the top x-axis, and labelled demarcations indicate the location of useful spectral features. Top: star-forming (SF) galaxy with high signal-to-noise (S/N). Middle: a similar SF galaxy with lower S/N. Bottom: a spectrum with no distinctive features classified as unknown.

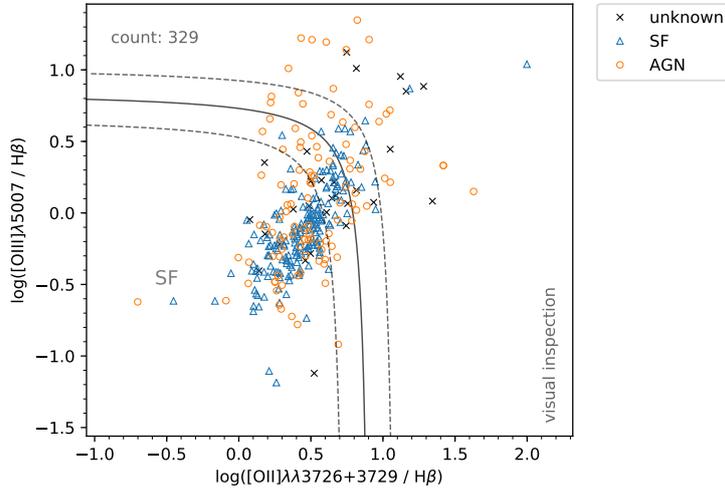


**Figure 2.3:** Three examples of OzDES spectra as they appear when classified visually. For each source, OzDES source designation and redshift are displayed in the lower-right, rest-frame wavelength is show on the top x-axis, and labelled demarcations indicate the location of useful spectral features. Top: Elliptical galaxy with high signal-to-noise (S/N). Middle: Elliptical galaxy with the presence of an [OII] emission line. Bottom: Low S/N example of an elliptical galaxy with [OII] emission.





**Figure 2.5:** [NII], [SII], and [OI] BPT classification diagrams. Sources are marked as star-forming, AGN, or Unknown based on the visual classification conducted in § 2.6. The lines are Kewley et al. (2001) theoretical classification lines (errors on these lines are  $\pm 0.1$  dex, these error lines are omitted for clarity).

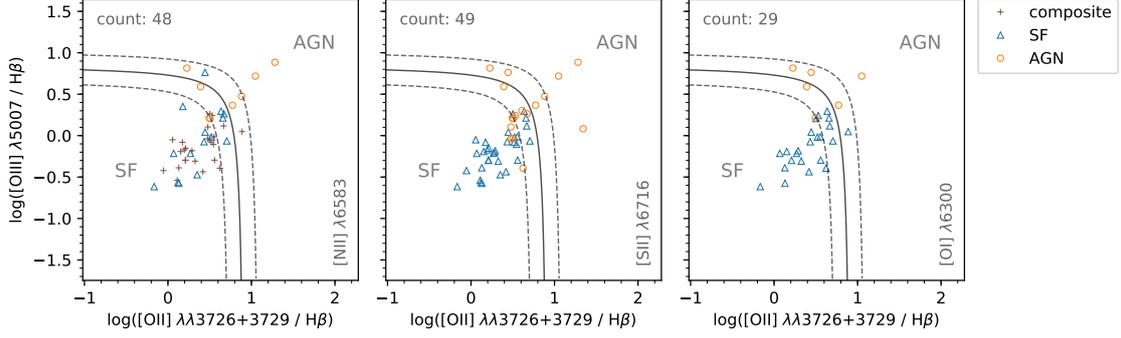


**Figure 2.6:** Lamareille (2010) blue-diagram of 329 OzDES spectra. Sources are marked as star-forming, AGN, or Unknown based on the visual classification conducted in § 2.6, as indicated in the lower-right of the plot. The population separation line (solid) is derived by Lamareille (2010), with  $\pm 0.2$  dex lines denoting the error on this line (dashed).

## 2.6.1 COMPARING VISUAL CLASSIFICATION WITH DIAGNOSTICS

## 2.7 COMBINING CLASSIFICATIONS

Each classification technique can benefit from the inclusion of additional diagnostics. For example, the blue-diagram fails at separating star-forming galaxies from Seyferts, BPT lacks clarity close to the composite region and is severely redshift restricted, visual inspection is subjective, qualitative, and lacks repeatability, and the far-infrared radio-correlation is subject to a ‘bright things are bright’ bias and cannot identify radio-quiet AGN (Wong et al. 2016). For



**Figure 2.7:** Three Lamareille (2010) blue-diagrams of OzDES spectra with the requisite spectral lines fitted by pPXF. Sources are marked as star-forming, AGN, or Composite based on the three the BPT diagnostic plots in Figure 2.5, as indicated by the unique species in the lower-right of each plot. The population separation line (solid) is derived by Lamareille (2010), with  $\pm 0.2$  dex lines denoting the error on this line (dashed).

these reasons, I have constructed a simple linear combination of the diagnostics used for the OzDES subset of ATLAS sources.

For each spectral classification method AGN, SF, composites and unknowns are scored a classification value,  $C$ , according to the following:

$$C = \begin{cases} -1 : & \text{star-forming,} \\ 0 : & \text{BPT composite or visual unknown,} \\ +1 : & \text{active galactic nuclei.} \end{cases} \quad (2.1)$$

For classifications with a segregation line (i.e., BPT [NII], [SII], and [OI], and blue-diagram) I calculate the absolute distance,  $|d|$ , to the (orange) AGN–SF segregation line (e.g., Figure 2.5) for each source. The final classification,  $C_{final}$  is then calculated via

$$C_{final} = (C_{visual}) + d_{[NII]}C_{[NII]} + d_{[SII]}C_{[SII]} + d_{[OI]}C_{[OI]} + d_{blue}C_{blue}. \quad (2.2)$$

Subscripts correspond to a classification method; where [NII], [SII], and [OI] uniquely identify each of the BPT diagnostics plots in Figure 2.5, *blue* refers to Lamareille (2010) blue-diagram (Figure 2.6), and *visual* refers to visual classification,  $C_{visual}$  is in parenthesis as it is only included for visual classifications with a confidence flag  $> 1$  (see § 2.6). In the event that a source has no classification then it is scored zero for that component\*. Naturally, sources with no classifica-

\*No contribution to Equation (2.2) is included from the radio–far-infrared correlation because of the lack of a well established population segregation –in particular the radio–far-infrared correlation can only identify bright AGN – instead this correlation is later used as an independent consistency check for  $C_{final}$ .

tions will retain a  $C_{final} = 0$ .

From here the galaxies are assigned to a category based on the sign of  $C_{final}$ :

$$C_{final} \begin{cases} < 0 : & \text{star-forming,} \\ = 0 : & \text{unknown}^{\dagger}, \\ > 0 : & \text{active galactic nuclei.} \end{cases} \quad (2.3)$$

This method of combining diagnostics guarantees several things: if a source is similarly classified in all diagnostics it will retain that classification; sources with only one classification will default to that classification; and a source with only one disagreement across multiple valid classifications can only be altered by that contrary result if the distance to the classification line ( $d$ , see Equation (2.2)) is sufficiently large to offset the other classifications.

### 2.7.1 SPECTRAL DIAGNOSTICS AND COMBINED CLASSIFICATION

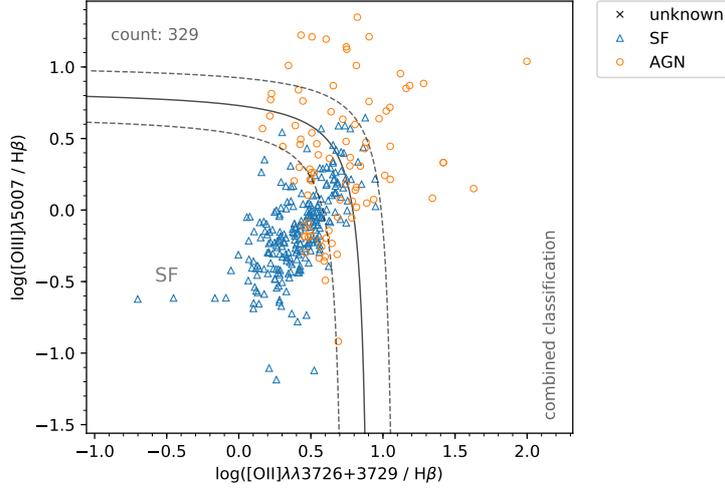
From Figure 2.5, visual classification produces reliable classification for most of the BPT spectra. There are several sources mixed across both regions that could be due to user error, or obfuscated spectra. Specifically, several sources fall into the AGN regime in each diagnostic from Figure 2.5 that I visually classified as unknown, or as a star-forming galaxy. Inspecting Figure 2.6 these sources along with a few others can also be seen in the AGN regime of the blue diagram. After calculating and combining the classification diagnostics, the final classifications are more broadly consistent with all diagnostics. Figures 2.8 and 2.9 show the BPT and blue-diagram with classifications marked according to the combined classification, there is now much better agreement with these diagnostics than when comparing to visual classifications only (see Figures 2.5 and 2.6).

### 2.8 RADIO–INFRARED CORRELATION

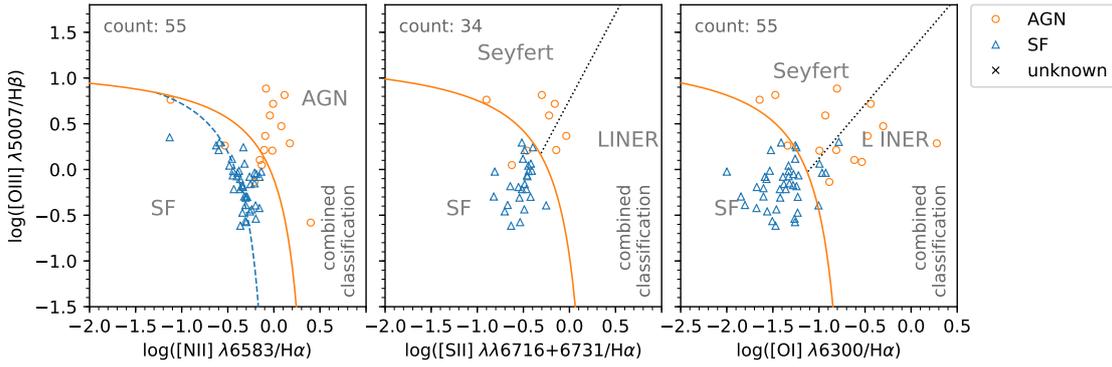
Figure 2.10 depicts the 1.4 GHz ATLAS radio flux density plotted against 24, 3.6, 4.5, 5.8, and 8  $\mu\text{m}$  infrared flux. Figure 2.11 shows the ratio of these plotted against  $\log(1 + z)$ . Both of these plots have sources marked in the same manner as the combined classifications. Looking at Figure 2.10, the majority of star-forming, AGN, and unknown galaxies lie along the locus, with several sources scattered towards higher 1.4 GHz flux density. These outliers are exaggerated in Figure 2.11, below  $S_{24\mu\text{m}}/S_{1.4} < 10^2$  there are 11 sources classified as unknown, 2 sources are AGN, and 2 are SF. It is expected that AGN will populate the high radio flux density region of these plots: the sources marked as SF are either host to an AGN buried inside a dusty SF galaxy (e.g., [Norris et al. 2012](#)), or they are SF galaxies with strong radio emission (e.g., ultra-luminous infrared galaxies [da Cunha et al. 2010](#)).

---

<sup>†</sup>For this thesis, I am only investigating star-forming and AGN galaxy populations; therefore, I do not distinguish between unknown galaxy classifications, BPT composites, or unclassified galaxies.

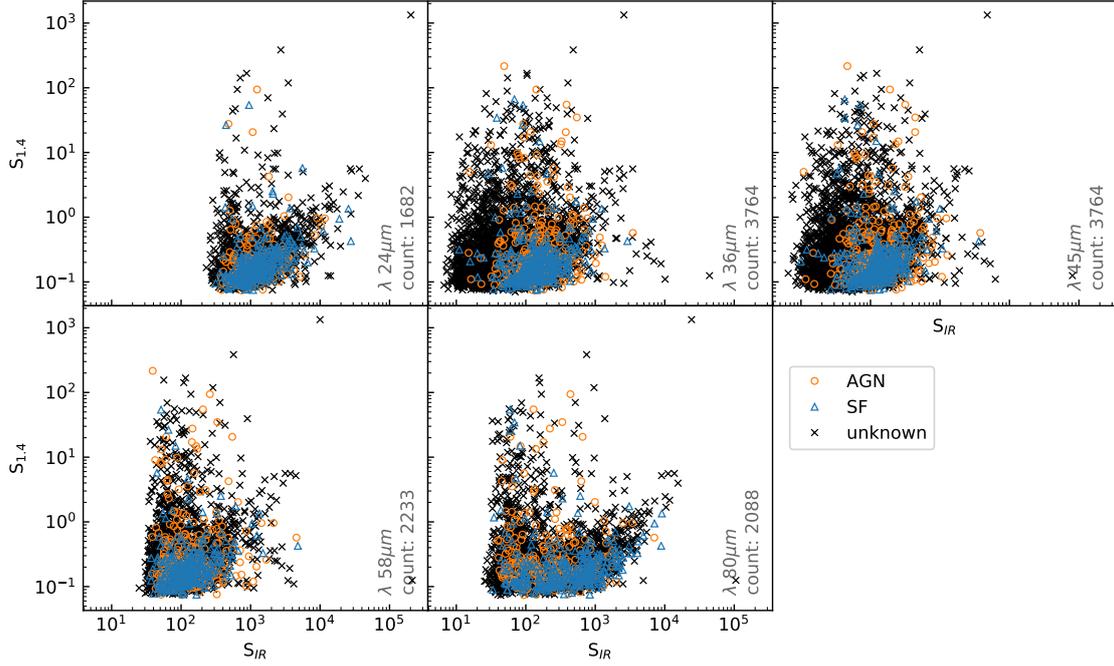


**Figure 2.8:** Lamareille (2010) blue-diagram of 329 OzDES spectra. Sources are marked as star-forming, AGN, or Unknown based on the combined classification, as indicated in the lower-right of the plot. The population separation line (solid) is derived by Lamareille (2010), with  $\pm 0.2$  dex lines denoting the error on this line (dashed).



**Figure 2.9:** [NII], [SII], and [OI] BPT classification diagrams. Sources are marked as star-forming, AGN, or Unknown based on the combined classification. The lines are Kewley et al. (2001) theoretical classification lines (errors on these lines are  $\pm 0.1$  dex, these error lines are omitted for clarity).

To determine if these sources are indeed AGN I reinspected the radio-infrared cutout and spectrum (where available) for all sources with  $z > 1.5$  and  $S_{24\mu\text{m}}/S_{1.4} < 10^3$  (see the dashed grey box in Figure 2.12). From these inspections 14 sources were confirmed to be AGN, 21 unknown galaxies were classified as AGN, and 6 SF classifications were changed to AGN. These classification corrections were decided on the basis of morphology (i.e., a multi-component extended source), radio intensity (the source must be compact and  $S_{1.4} > 128 \times \text{RMS}$ : i.e., seven radio contours), or exhibit AGN spectral characteristics: see the spectra and image cutouts for these 41 sources Figures C.2 to C.8 (Pages 85 to 91). Each source that was identified as an AGN



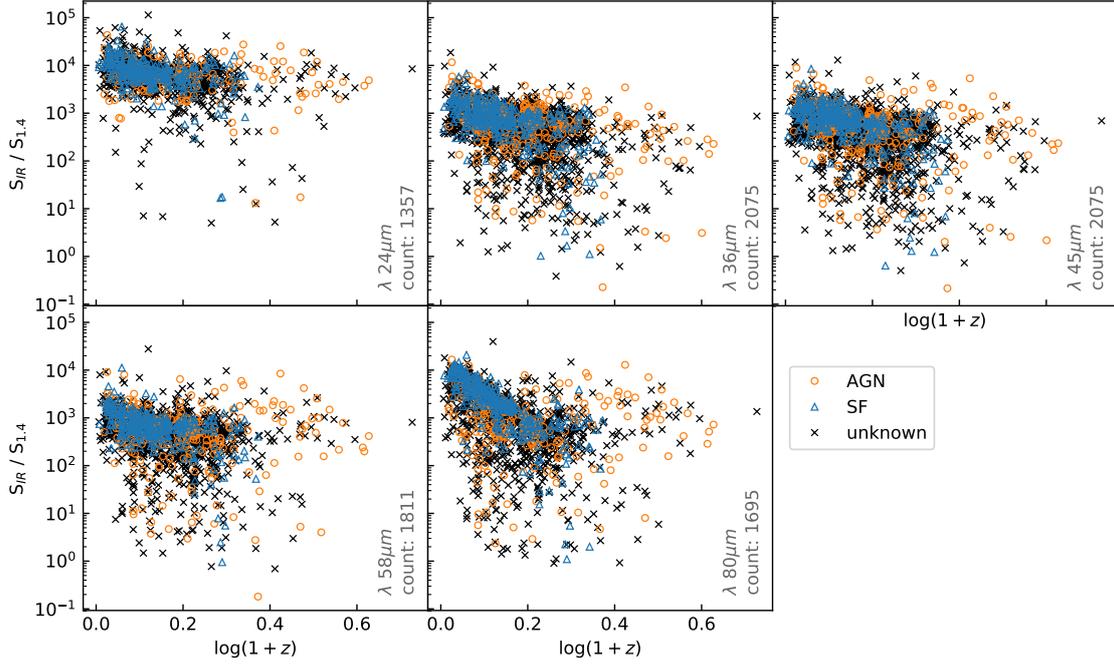
**Figure 2.10:** 1.4 GHz radio flux plotted against 24, 3.6, 4.5, 5.8, and 8  $\mu\text{m}$  infrared flux. The infrared wavelength and source total source count is indicated in the upper right of each plot, these counts change with the availability of the required constituents from each plot. Sources are marked as star-forming, AGN, or Unknown based on the combined classification.

via this inspection I manually changed the final combined classification to reflect this.

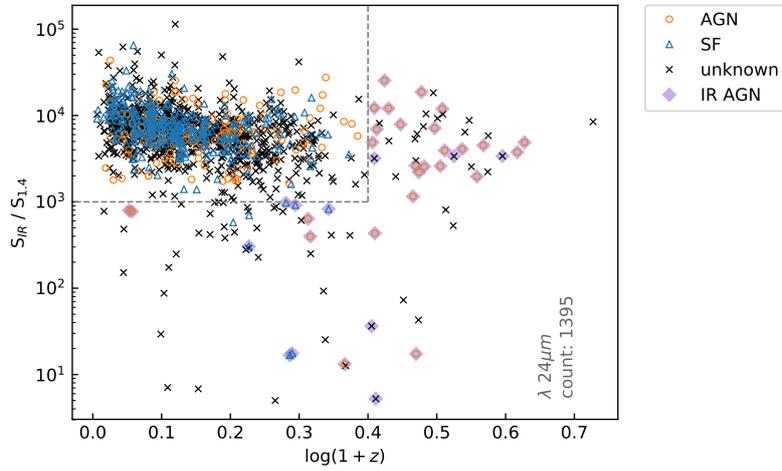
## 2.9 FINAL ATLAS RADIO POPULATION CLASSIFICATIONS

Figure 2.13 shows the distribution of star-forming galaxies, AGN, and unknown classifications for both visual spectral classification and the combined classification outlined in Equation (2.2), and include the manual corrections to sources that deviate from the  $S_{24\mu\text{m}}/S_{1.4}$  correlation. The bottom pane of Figure 2.13 illustrates the difference per bin of visual and the final combined classification. Combining classification techniques yields 39 new star-forming galaxies and 11 additional AGN compared to raw visual classifications. Of the 11 additional AGN, I had incorrectly classified 7 as SF via visual inspection; no incorrect AGN classifications were made. It is obvious then that visual-spectral classifications are sufficiently subjective that more than one pass should be made, more than one person should perform the visual classifications, or the visual classifications should be combined with additional diagnostics to correct similar errors.

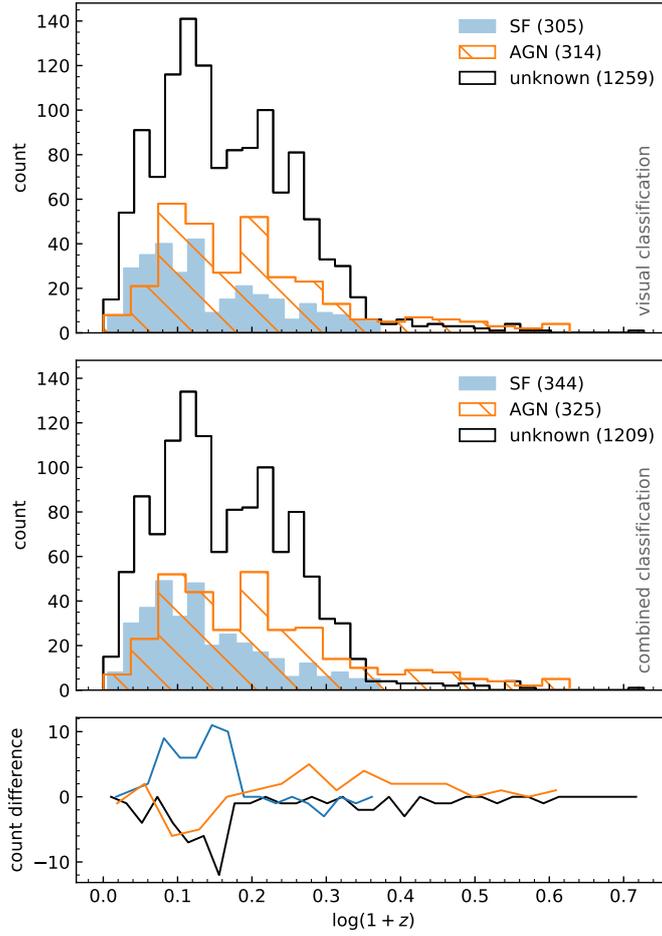
In the following chapter, I construct 1.4 GHz radio luminosity functions for the ATLAS sources and establish the evolution of this function across multiple redshift ranges. Of the 1851 ATLAS sources with redshifts (925 with spectra) 344 (19 per cent) are classified as star-forming,



**Figure 2.11:** Ratio of 1.4 GHz radio and 24, 3.6, 4.5, 5.8, and 8.0  $\mu\text{m}$  infrared flux plotted against  $\log(1+z)$ . The infrared wavelength and source total source count is indicated in the lower right of each plot, these counts change with the availability of the required constituents from each plot. Sources are marked as star-forming, AGN, or Unknown based on the combined classification.



**Figure 2.12:** Ratio of 1.4 GHz radio and 24  $\mu\text{m}$  infrared flux plotted against  $\log(1+z)$ . Sources are marked as star-forming, AGN, or Unknown based on the combined classification. All sources exterior to the grey bounding box were reinspected, and those sources determined to be AGN via this inspection are also marked as 'IR AGN'.



**Figure 2.13:** Histograms of ATLAS redshifts separated into star-forming, AGN, and unknown. For the top pane these are based on visual classification, and the middle pane uses the combined classification including corrections outlined in § 2.8. The bottom pane is the difference per bin (combined — visual) of these histograms.

and 325 (17 per cent) as AGN, using these classifications I also establish the evolution of ATLAS AGN and SF galaxies.

*Reality is that which, when you stop believing in it, doesn't go away*

– Philip K. Dick

# 3

## Evolution of ATLAS Radio-Galaxy Populations

The correlation of stellar velocity dispersion and stellar bulge mass to the mass of the supermassive black holes at the centre of galaxies is evidence for a coupling between this central engine and the broader galaxy properties. The most substantiated mechanism for this coupling is feedback from the active galactic nucleus via the injection of energy into the galaxy and surrounding environment. This energy originates from the accretion of matter onto the black hole from galaxy mergers, the hot halo, or cluster environment; and manifesting as relativistic jets tens to hundreds of kilo parsecs in scale (Cheung 2004). This mechanism has been shown to suppress star formation in the host galaxy and nearby cluster members (e.g., Binney & Tabor 1995; Best et al. 2006; Croton et al. 2006).

The accretion mechanisms that feed black holes are discussed in depth in § 0.4.2. To briefly re-iterate there are two modes of accretion: cold-mode and hot-mode. Cold-mode accretion is closely associated with powerful high-excitation radio galaxies (HERGs), in which cold gas from a gas-rich companion or environment is accreted onto a central supermassive black hole (SMBH) triggering powerful jets that drive intergalactic winds, these winds are capable of suppressing or enhancing star formation (e.g., Shabala et al. 2011; Davis et al. 2012). Hot-mode accretion is associated with low excitation radio galaxies (LERGs). LERGs produce less energetic radio jets, with little or no spectral emission. The hot-mode accretion occurs via cooling of

thermal halo gas or via regular ‘chaotic accretion’ from a large cold gas reservoir which can provide. The resulting jets not only suppress star-formation but also heat the in-falling gas feeding the AGN which can establish a self-regulating cycle (Best et al. 2006; Kaiser & Best 2007).

The coupling of AGN activity and star-formation quenching serves to couple their cosmic evolution. At high redshift AGN undergo rapid accretion resulting in black-hole growth and frequent SF quenching. This results in massive galaxies dominated by an older stellar population (Tadhunter 2016). In later epochs, there is a higher space-density of LERGS (Best & Heckman 2012) and evidence of more regular, less violent feedback periodically suppressing star formation (Fabian et al. 2006; Best et al. 2006; Tadhunter 2016; Ching et al. 2017).

### 3.1 RADIO LUMINOSITY FUNCTIONS

1.4 GHz radio-luminosity of a galaxy is a proxy for the energy output of AGN or star-formation rate, tracing the synchrotron emission excited by these processes. Best & Heckman (2012) established subclasses of AGN distinguished by core structure and accretion mechanism. They discovered that these subclasses follow different cosmic evolutionary tracks. Specifically, below  $L_{1.4} < 10^{26}$  W Hz<sup>-1</sup> the space density of LERGs is greater than HERGs and the location of this break in dominant population shifts changes with redshift, favouring an abundance of HERGs.

Brighter radio galaxies with strong optical emission lines (i.e., HERGs) are fast-evolving observing an increase in comoving spatial density between two to three orders of magnitude over the redshift range  $z \sim 0$  to  $z \sim 2$  (Dunlop & Peacock 1990; Willott et al. 2001; Clewley & Jarvis 2004; Donoso et al. 2009). Fainter radio galaxies ( $L_{1.4} < 10^{23}$  W Hz<sup>-1</sup>) with little optical spectral emission (i.e., LERGs) display very little evolution out to  $z < 0.7$  (Croom et al. 2004; Sadler et al. 2007; Smolčić et al. 2009; Donoso et al. 2009) and no more than a factor of 2 out to  $z < 1$  (Padovani et al. 2009; McAlpine et al. 2013).

The relatively faint luminosities (compared to AGN) produced by star-forming galaxies makes it difficult to establish a statistically reliable sample to gauge evolution at high redshifts. Recent radio surveys with modern instrumentation and sophisticated galaxy classification schemes have probed SF galaxy population densities to  $z \sim 0.75$  (Pracy et al. 2016), and  $z \sim 5$  (Novak et al. 2017). Ultra-violet selected SF galaxies have allowed for evaluation of population densities to redshifts as high as  $z \sim 10$  (Bouwens et al. 2015). These studies verify previous results indicating strong positive evolution of galaxies dominated by star formation (e.g., Hopkins et al. 1998; Best & Heckman 2012).

Continuing contributions from deep radio studies such as ATLAS contribute to the refinement of our understanding of these AGN and SF populations at intermediate to high redshifts. By observing the cosmic evolution of the volume density of these galaxies Longair (1966), we are able to understand the effect of AGN on the formation of modern galaxies, constrain cosmic assembly models, and guide further studies to ultimately build a deep understanding of the galactic and cluster environments that produce AGN.

## 3.2 CONSTRUCTING RADIO LUMINOSITY FUNCTIONS

### 3.2.1 VOLUME CORRECTIONS

Because the two ATLAS survey regions are small ( $\Omega < 4 \text{ deg}^2$ ) we can make small angle approximations for the calculation of the comoving volumes enclosed by the ATLAS survey; that is, we simply scale the spherical volume equation by  $\Omega/4\pi$ . This results in a simple conical volume for a comoving distance  $D_M^*$ :

$$V = \frac{1}{3} D_M^3 \Omega. \quad (3.1)$$

To correct for the under-counting of radio sources inherent to any flux-limited survey, we assume homogeneity of source distributions within the ATLAS volume. We consider  $V$  to be the volume calculated at the detection distance of a source, and  $V_{max}$  to be the corresponding volume for the distance at which that source would fall below the detection limit of the survey. Following the adaptation of the  $V/V_{max}$  correction (Schmidt 1968) established by Best & Heckman (2012) to allow for a sliding redshift window (i.e., the specification of a minimum redshift of interest corresponding to the minimum volume element,  $V_{min}$ ), the contribution to the comoving volume density of each source count is scaled by the volume weighting:

$$\omega_V = \frac{V - V_{min}}{V_{max} - V_{min}}. \quad (3.2)$$

Because the sources contributing to the radio-luminosity function require an optically obtained redshift as well as a 1.4 GHz radio flux-density,  $\omega_V$  needs to account for both radio and optical survey limitations; that is, a source will only be detectable within the most limiting volume. So we can take

$$V_{max} = \text{minimum} \left\{ V_{max,o}, V_{max,r}, V_{max,z} \right\}, \quad (3.3)$$

where  $V_{max,o}$ ,  $V_{max,r}$ , and  $V_{max,z}$  are the optical, radio and redshift maximum volumes respectively.  $V_{max,z}$  is included for cases where a manual redshift-cut is applied to the data that results in a smaller maximum comoving volume than either the optical or radio limitations (e.g., Figure 3.1). Going forward, all calculations adopt  $\Lambda$ -CDM cosmological parameters of  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (Spergel et al. 2003) to maintain consistency with literature.

---

\*In practice, volumes calculated in this chapter use `astropy.cosmology.comoving_volume` and `astropy.cosmology.comoving_distance` to correctly integrate over a non-spherical cosmology ( $\Omega_k \neq 0$ ).

### 3.2.2 MAXIMUM OPTICAL VOLUME

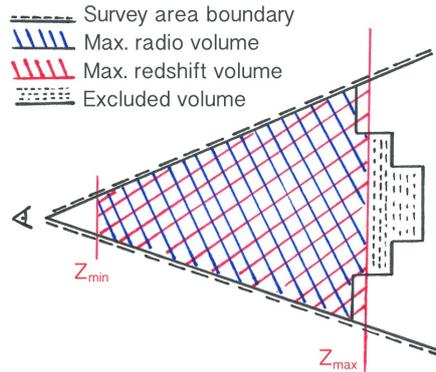
Calculation of  $V_{max,o}$  simply requires the calculation of an absolute magnitude for each source via the magnitude-distance relationship:

$$M = m - 5 [\log_{10}(D_L) - 1]. \quad (3.4)$$

followed by the inversion of this relationship to solve for a maximum luminosity distance  $D_{L,max}$  using the calculated absolute magnitudes and a specified limiting apparent magnitude,  $m_{lim}$ .

### 3.2.3 MAXIMUM RADIO VOLUME

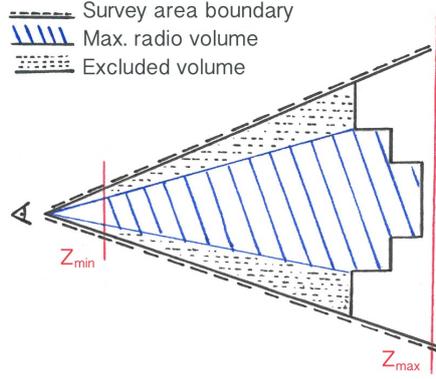
Calculations of the maximum radio volumes are more complicated than optical volumes due to the non-uniform sensitivity of the survey area (see Figure 3, Page 10). Instead of calculating a maximum volume for an averaged flux-density sensitivity limit, I break the survey area into discrete sensitivity bins (similar to Mao et al. 2012). The final maximum radio volume is the sum of the binned volumes for which the source would have been visible; that is, an individual bin only contributes to the summed  $V_{max}$  if the flux density limit of that bin is less than the flux-density of the source (see Figure 3.2).



**Figure 3.1:** Diagram demonstrating the effect of slice-wise volume calculations for a redshift volume limited source. Each tier represents the detectable volume of a source for a hypothetical RMS limit slice. The red lined region shows the maximum extent of an arbitrary redshift volume, and the stepped outline is the  $V_{max,r}$  over all bins for a source. The blue lined region is the final  $V_{max,r}$  once accounting for the excluded black dotted region and redshift regions.

The k-corrected 1.4 GHz luminosity of each source is calculated from the observed 1.4 GHz flux,  $S_{1.4}$ , via:

$$L_{1.4} = \frac{4\pi S_{1.4} D_M^2}{(1+z)^{\alpha-1}}, \quad (3.5)$$



**Figure 3.2:** Diagram demonstrating the effect of slice-wise volume calculations for a radio flux-density limited source. Each tier represents the detectable volume of a source for a hypothetical RMS limit slice. The blue lined region is the final  $V_{max,r}$  for this source. The black dotted regions are excluded as the source flux-density is less than five times the RMS of these bins.

where the spectral index  $\alpha = -0.7$  (Condon et al. 2002; Sadler et al. 2002) is chosen to maintain consistency with the literature used for comparison later in this chapter.

For a survey area separated into  $i$  regions, with corresponding sky-areas  $\Omega_i$  and flux-density limit,  $S_{lim,i}$ <sup>†</sup> the maximum co-moving distance for each source within this slice is calculated via a rearrangement of Equation (3.5), and the corresponding  $V_{max,r}$  for each source is then the sum of all contributing volume slices:

$$V_{max,r} = \frac{1}{3} \sum_{i=1}^n D_{M,max_i}^3 \Omega_i. \quad (3.6)$$

### 3.2.4 COMOVING SOURCE DENSITY

To assemble the radio-luminosity function sources are binned by 1.4 GHz luminosity in logarithmic space, and the comoving density of sources within these luminosity bins is calculated according to:

$$\Phi_{bin} = \frac{1}{\Delta_{bin}} \sum_{sources} \frac{\omega_{V,source}}{V_{source}} \times \frac{1}{C_{source}}, \quad (3.7)$$

<sup>†</sup>For all  $i$   $S_{lim,i}$  is taken to be five times the median root-mean-square (RMS) flux-density in that bin. Because the two ATLAS fields, CDFS and ELAIS-S1, have different RMS profiles these values are calculated independently. The binned median RMS values for ELAIS-S1 are [17.7, 25.1, 35.6, 51.5, 73.8, 106]  $\mu\text{Jy beam}^{-1}$  and for CDFS [15.9, 21.7, 33.4, 50.3, 74.7, 114]  $\mu\text{Jy beam}^{-1}$  covering [1.40, 0.76, 0.51, 0.36, 0.31, 0.27]  $\text{deg}^2$  and [1.46, 1.37, 0.72, 0.54, 0.41, 0.37]  $\text{deg}^2$  respectively.

in this equation  $\omega_{V,source}$  is calculated from Equation (3.2) according to the most limiting volume (see Equation (3.3)). Regardless of the limiting volume,  $C_{source}$  is the spectroscopic completeness at the corresponding magnitude of each source (see § 3.4.1). The density in each bin is normalised by the (logarithmic) width of the bin,  $\Delta_{bin}$ .

### 3.3 RADIO-LUMINOSITY FUNCTION PIPELINE

To limit the possibility of unintentional biases (including under- and over-binning), and to maximise the reproducibility and consistency of RLF construction across various redshift, radio flux-density, and optical magnitude ranges I endeavoured to partially automate the RLF assembly. I first attempt to reproduce the [Best & Heckman \(2012\)](#) RLF (for their full sample), before constructing the ATLAS RLFs later in this chapter. In this way, any techniques I develop can be compared to an established result before being used to calculate an unknown result. The SDSS [Best & Heckman \(2012\)](#) sample consists of  $> 18\,000$  sources in the redshift range  $0.01 < z < 0.3$ . Their online data catalogue provides the radio data and redshift of each source, and unique identifiers for the SDSS spectra allows for manual retrieval of  $G$ ,  $R$ , and  $I$  magnitudes. Assuming a uniform radio flux density sensitivity limit of 5.0 mJy, and discarding sources with flux densities less than this  $\sim 12\,000$  sources remain.

The luminosity of each source is recalculated directly from the 1.4 GHz flux-densities and redshifts, and the radio luminosity function constructed from these data in the manner outlined in § 3.2. Before experimenting with alternative binning techniques, I successfully reproduced BH12 RLF (with no apparent disagreement) using their original binning to confirm my volume calculations and corrections.

#### 3.3.1 AUTOMATED RLF BINNING

Because an RLF is inherently a histogram of counts weighted by volume calculations, it is important that statistical consideration is given to the binning of the data. Having too many or too few bins can either introduce noise or smooth genuine structure. The shape of an RLF is particularly sensitive to these effects, and poor binning can produce spurious breaks leading to false population conclusions ([Cara & Lister 2008](#)). There are several well established automatic statistical binning techniques, for example [Scott \(1979\)](#); [Freedman & Diaconis \(1981\)](#) and [Knuth \(2006\)](#) <sup>#</sup>. However for the incomplete volume-corrected RLF data, the Scott and Freedman-Diaconis algorithms produce too many bins resulting in false ‘breaks’ in the RLF, and the Knuth algorithm produces too few bins. Both of these are an issue that would be exacerbated if applied to the much smaller ATLAS sample.

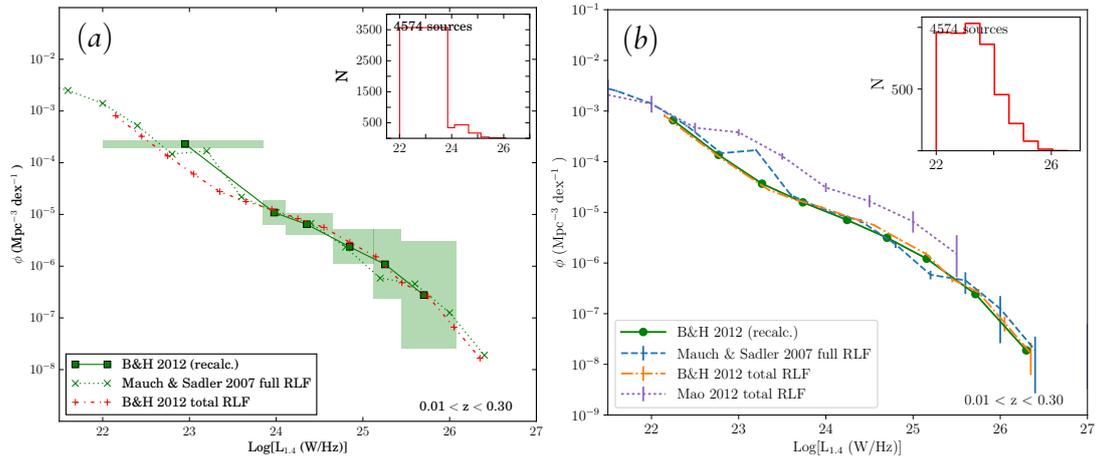
Passing the luminosity data into the [Scargle et al. \(2013\)](#) Bayesian binning algorithm produces intelligently allocated ‘Bayesian blocks’ of variable width (see Figure 3.3a). Because this

---

<sup>#</sup>All histograms in this thesis are constructed using the Knuth algorithm ([Knuth 2006](#)) implemented via `astropy.stats.knuth_bin_width`.

algorithm only considers actual source counts – that is, it does not recognise volume corrections for ‘missing’ detections – the lowest luminosity bin is very broad. As  $\Phi_{bin}$  is scaled by each bin width the result is a reasonably accurate reproduction of the BH12 RLF using only six bins. However, I feel that this binning does not accurately represent the function of a volume-corrected RLF and opted to use a fixed-width binning that piggy-backs on the Bayesian algorithm.

Figure 3.3*b* is the BH12 RLF constructed using linearly spaced bins calculated as the median width of the Bayesian variable bin widths in Figure 3.3*a*. Additionally, because the distribution of source luminosities within each bin will not be normally distributed about the bin centres, the demarcation for each bin is instead the median luminosity of all sources in that bin. The resulting luminosity function shows practically zero deviation from the established BH12 RLF. This automatic binning technique benefits from the flexibility of the Bayesian block allocation but avoids the under-representation of low-luminosity sources, and avoids the arbitrary 0.5 dex bin widths commonly used in RLF construction. The RLF constructed in the remainder of this thesis uses this ‘Bayesian-median’ binning technique.

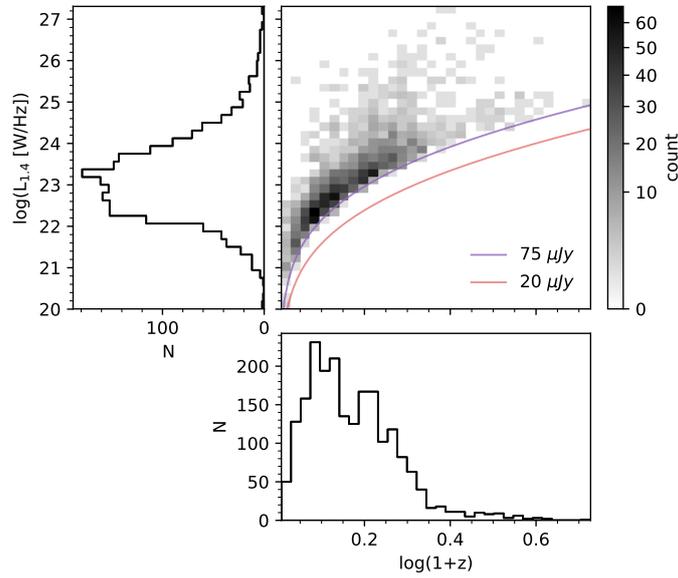


**Figure 3.3:** Reconstruction of the full sample Best & Heckman (2012) radio luminosity function. The calculations for BH12 RLF is constructed according to Equation (3.7), as the BH12 SDSS data is spectroscopically complete (i.e.  $C_{source} = 1$ ) this allows for the independent verification of the per-source volume corrections and automatic binning established in this chapter. Pane *a* binning is based on Scargle et al. (2013) Bayesian block representation, error bars are Poisson  $\sqrt{n}$ . Pane *b* uses a fixed median width binning discussed in § 3.3.1, error bars are omitted for easier comparison to the BH12 result. Embedded in the top-right of each pane is a histogram of per-bin source counts.

### 3.4 ATLAS SOURCE STATISTICS AND RESTRICTIONS

Figure 3.4 shows the final distribution of 1851 ATLAS sources (of the 4903 for which a redshift is available) plotted in 1.4 GHz luminosity–redshift space. Inspecting the redshift histogram

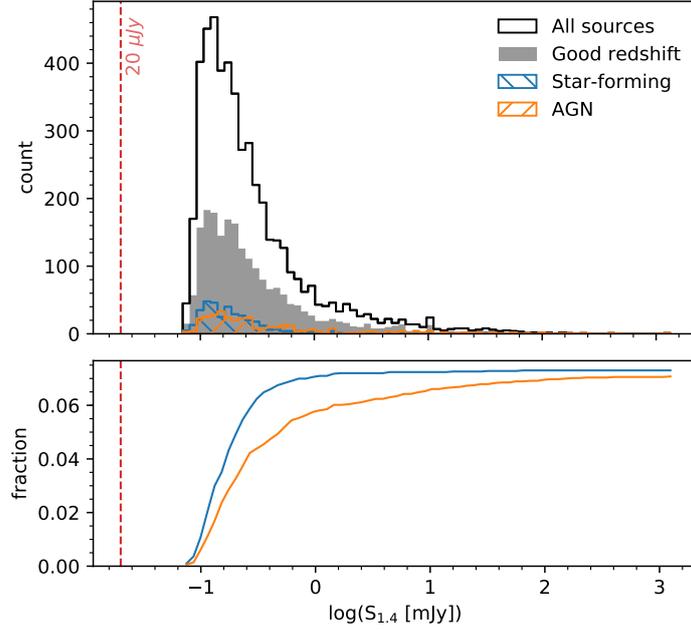
source counts begin to drop off rapidly at  $\log(1+z) \sim 0.3$  indicating that spectroscopic completeness of the ATLAS sample beyond  $z > 1.0$  is too low to be used for population inferences. The two dimensional histogram of luminosity versus redshift shows the majority of the ATLAS sources occupying  $L_{1.4} < 10^{24}$ , and over-plotted on this histogram are two curves representing the expected luminosity-redshift cut-off for a survey with flux-density sensitivity limits of  $20 \mu\text{Jy}$  and  $75 \mu\text{Jy}$ . This illustrates that the ATLAS  $\sim 20 \mu\text{Jy}/\text{beam}$  sensitivity corresponds to a source-detection sensitivity floor of  $S_{1.4} \sim 75 \mu\text{Jy}$ . Figure 3.5 presents the  $S_{1.4}$  distribution of the ATLAS star-forming and AGN galaxies (see § 2.9 for details on these classifications). Star-forming galaxies have a higher representation of faint sources than AGN: 6.5 per cent of SF galaxies appear below  $S_{1.4} < 5 \text{ mJy}$  compared to a similar fraction of AGN only being attained below  $S_{1.4} < 80 \text{ mJy}$ .



**Figure 3.4:** 1.4 GHz luminosity versus redshift of the 1851 ATLAS sources with redshift measurements. The central pane is a two dimensional histogram showing the distribution of source luminosities with redshift; the redshifts are plotted as  $\log(1+z)$ . In the central pane, hypothetical survey flux-density sensitivity limits of  $20 \mu\text{Jy}$  and  $75 \mu\text{Jy}$  are shown. The bottom and left pane are individual histograms of redshift and luminosity which serve to more clearly illustrate their individual distributions.

### 3.4.1 OPTICAL COMPLETENESS CORRECTIONS

From the 4903 ATLAS sources, only 1851 (38 per cent) have reliable redshifts, and 925 (19 per cent) have available spectra. Figure 1.10 shows the source counts in each of the four DES photometric bands. In addition to the total source distribution, each panel distinguishes sources as having a successful redshift measurement, or an attempted redshift measurement (the difference between these would, therefore, be unsuccessful redshift measurements).

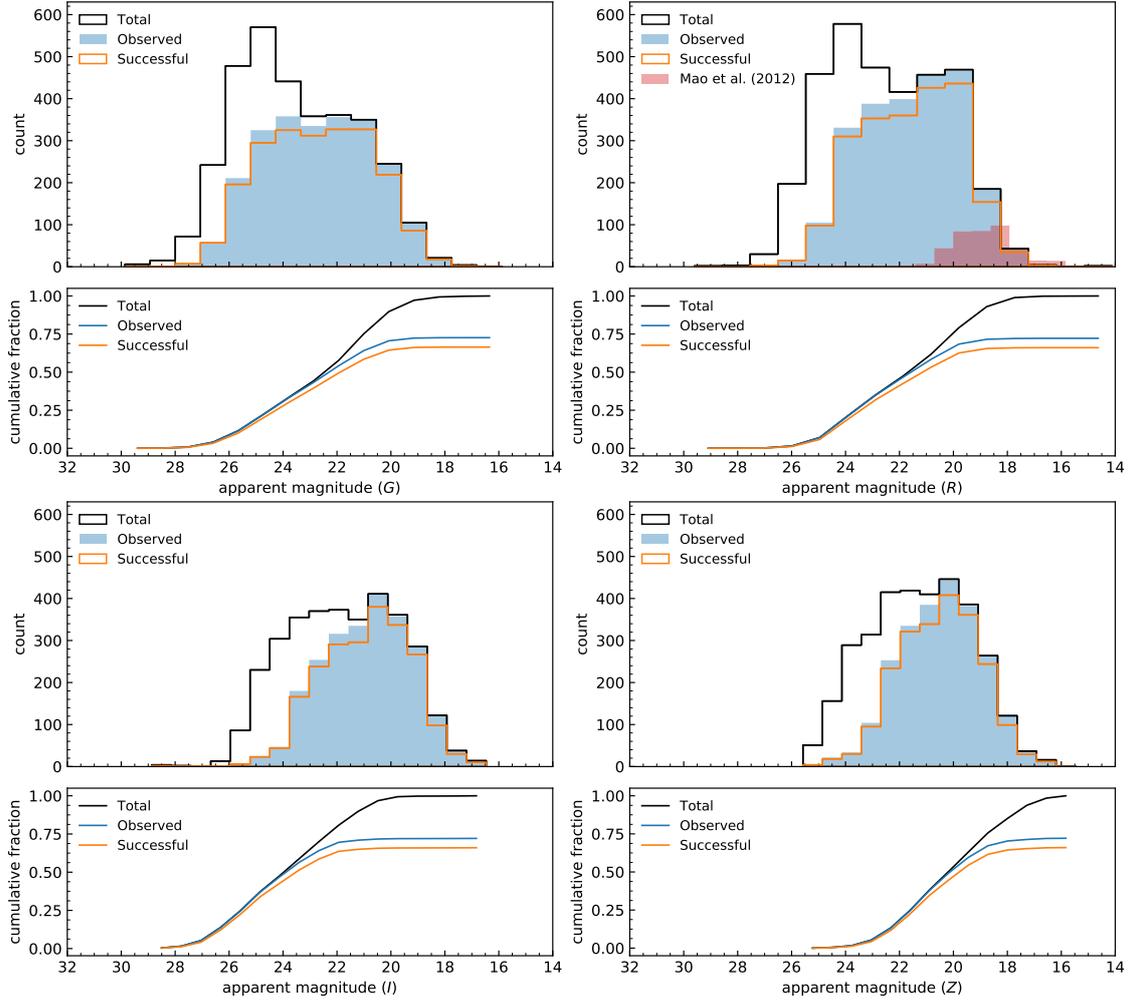


**Figure 3.5:** 1.4 GHz flux-density histogram (top), cumulative fraction (bottom). Sources are distinguished as possessing a ‘Good redshift’ (i.e.,  $cflag \geq 2$ ), or classified as an AGN or star-forming galaxy. The bottom pane only shows the cumulative fraction (of ‘All sources’) for star-forming and AGN subsets. A  $20 \mu\text{Jy}$  line is plotted for reference.

A reliable redshift is defined as  $cflag \geq 2$  (see Table C.1), and the success rate is the number of sources with a successful redshift divided by the number of sources observed for a redshift measurement regardless of success. Since the optical data for the ATLAS sources are drawn from the OzDES survey, I use the photometric and spectroscopic completeness of all OzDES sources instead of for the ATLAS subset. This avoids inadvertently including radio-optical selection effects in the optical completeness corrections when constructing the RLF. Figure 3.7 shows the redshift success rate per  $G, R, I, Z$  apparent magnitude bin. In order, the survey achieves 80 per cent completeness at 24.8, 24.0, 23.2, and 23.0. These spectroscopic completeness fractions are a significant improvement over ATLAS DR1 (Mao et al. 2012), in which the spectroscopic completeness restricted sources to  $m_R < 20$  (see Figure 3.6).

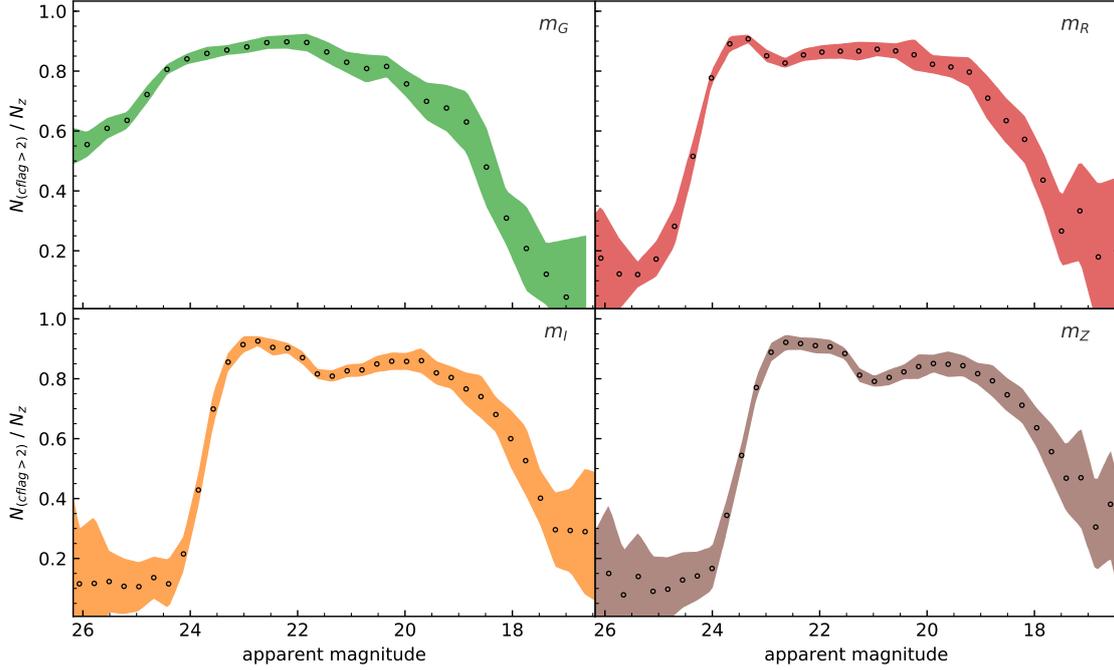
### 3.4.2 FINAL ATLAS SAMPLE RESTRICTIONS

Any restrictions applied to the ATLAS sample are a balance between preserving reliability (by eliminating sources of high incompleteness) and preserving sample size (by not zealously applying restrictions). If made correctly, these restrictions improve the conclusive power of the ATLAS RLF without introducing strong selection effects. The RLF constructed in this chapter will only consider sources with  $S_{1.4} > 80 \mu\text{Jy}$  to prevent the assembly of flux-density incomplete luminosity bins (see Cara & Lister 2008). Sources are further restricted to  $m_R < 23.5$  based on



**Figure 3.6:** Spectroscopic completeness of the final cross-matched ATLAS catalogue for each of the photometric bands observed by DES:  $m_G$ ,  $m_R$ ,  $m_I$ , and  $m_Z$ . The top panel of each pane shows the histogram of all ATLAS sources for the given apparent magnitude (black), as well as a histogram of all ATLAS sources targeted for a redshift measurement (blue), and finally a histogram of all ATLAS sources *with* a redshift for that optical band (orange). The bottom panel of each pane is the cumulative fraction of these histograms (normalised to the black ‘Total’ histogram). In the histogram of ATLAS DR1  $m_R$  magnitudes (Mao et al. 2012) are included for comparison.

the  $R$ –band spectroscopic completeness (Figure 3.7). To remove galactic objects from the sample the constraint  $z > 0.005$  will also be enforced. Due to the rapid decline in source counts beyond  $z > 1.0$  (Figure 3.4), and to avoid making population inferences from spectroscopically incomplete data, a maximum redshift  $z < 1.0$  is applied. Finally, when evaluating the (redshift-binned) cosmic evolution of the ATLAS 1.4 GHz source density, I construct the RLF over bins spanning  $z = 0.26 - 0.58 - 1.0$ : these redshift spans are chosen to maintain linear



**Figure 3.7:** Spectroscopic completeness of OzDES  $G$ ,  $R$ ,  $I$ ,  $Z$  sources. This completeness is defined as the ratio of successful redshifts (i.e.,  $cflag \geq 2$ ) to the number of attempted (or total) redshifts. Errors are plotted as a filled area, these are the *first* and *fourth* quartiles of Monte-Carlo resampled catalogue smoothed with a first-order interpolation.

spacing in  $\log(z + 1)$  (i.e., 0.1, 0.2, 0.3), partially offsetting the redshift dependence of  $L_{1.4}$  (see Equation (3.5)) allowing for a cleaner analysis of the cosmic evolution (see e.g., [Seymour et al. 2009](#)).

After these restrictions are applied 1633 (33 per cent of 4903) ATLAS sources remain, of which 756 (15 per cent) have spectra for a possible classification (established in § 2).

### 3.5 THE COMPLETE ATLAS RADIO LUMINOSITY FUNCTION

Before investigating the RLF of ATLAS sub-populations, the full ATLAS RLF is constructed across the three chosen redshift bins:  $0.005 < z < 0.26$ ,  $0.26 < z < 0.58$ ,  $0.58 < z < 1.0$  in Figure 3.8. In the lowest redshift bin ( $0.005 < z < 0.26$ ) the ATLAS RLF broadly agrees with literature. The break in the shape of the RLF at  $L_{1.4} \sim 10^{23.1} \text{ W Hz}^{-1}$  suggests that there are two distinct populations. The middle of the ATLAS RLF bins presents slightly higher source densities than the corresponding literature ( $\Delta\phi < 0.25$  dex). Due to selection effects (such as different instrument sensitivities) no two literary radio luminosity functions are expected to agree perfectly, and the uncertainty in the ATLAS RLF is large enough that this difference can be attributed to statistical variance.

There are two notable deviations of the  $z < 0.26$  ATLAS RLF from the compared literature,

whilst the low luminosity ( $L_{1.4} < 10^{21}$  W Hz $^{-1}$ ) difference should be considered insignificant due to the poor  $S_{1.4}$  completeness of these bins (see the embedded histogram in Figure 3.8), the  $\Delta\phi \sim 1.0$  dex difference in source density in the ATLAS  $L_{1.4} = 10^{25.38}$  W Hz $^{-1}$  bin (see Table 3.1) is worth investigating. This bin contains only two ATLAS sources (CI0723, CI2844), which appear to be genuine high-luminosity ( $L_{1.4} = 3.9, 2.2 \times 10^{25}$  W Hz $^{-1}$ ) AGN at low redshift ( $z = 0.108, 0.255$ ). The presence of these sources in the low-redshift CDFS field is likely a fluke of cosmic variance.

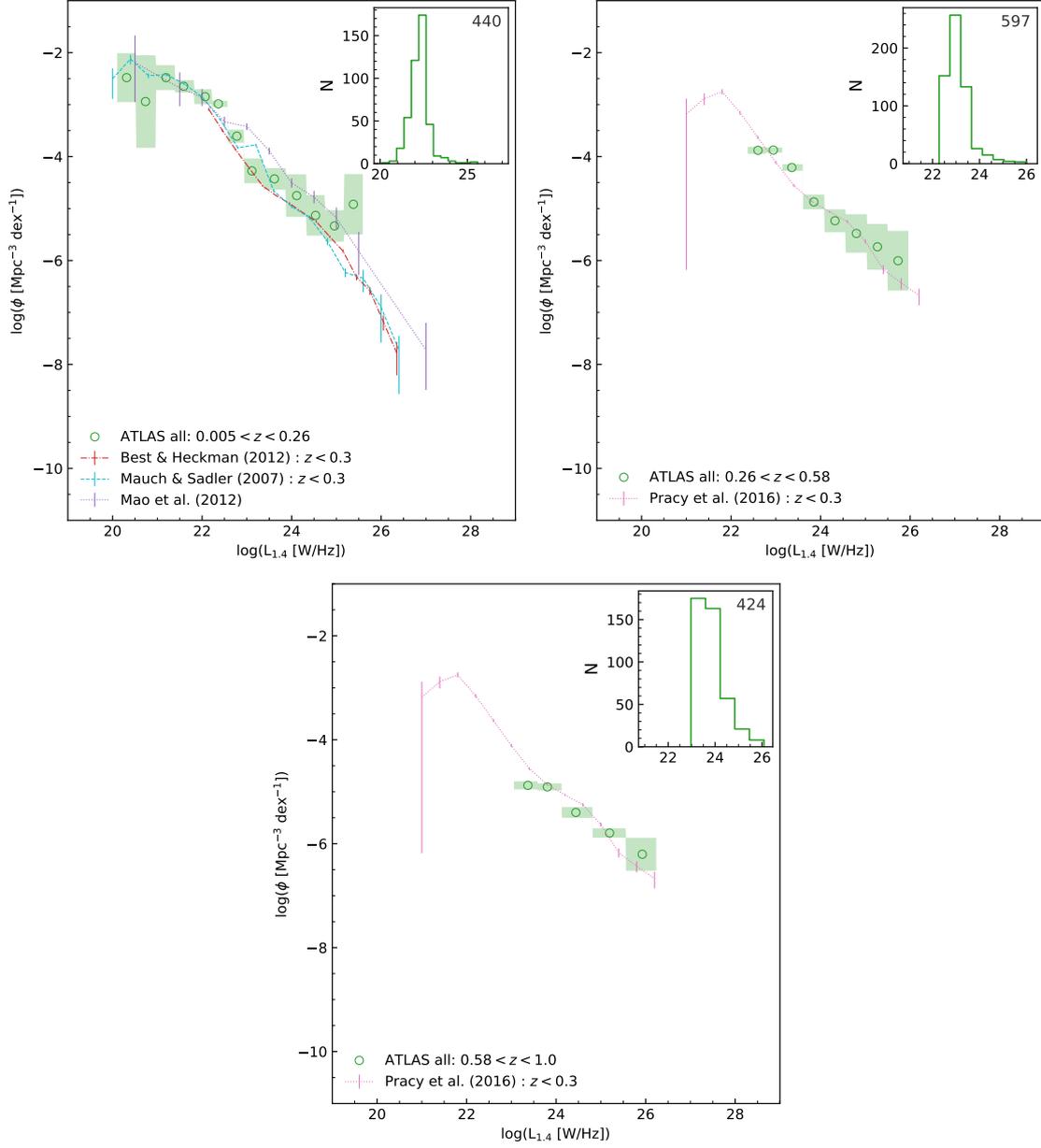
The higher redshift ranges of the ATLAS RLF maintain a fairly good agreement with the low-redshift ( $z < 0.3$ ) literature RLFs. The only changes of note are the truncation of faint sources (associated with different  $S_{1.4}$  completeness turn-overs), and a small deviation from the [Pracy et al. \(2016\)](#) RLF at  $L_{1.4} < 10^{25}$  W Hz $^{-1}$ . The lack of change to the ATLAS RLF across the three redshift bins indicates that there is no change in the comoving source density out to  $z < 1.0$ , or that any change is beyond the resolution of these data.

**Table 3.1:** The 1.4 GHz radio luminosity function for all ATLAS sources separated into three redshift bins as indicated by the top-most row. Due to the dynamic binning (see § 3.3.1) the RLFs are centered around a different set of  $\log(L_{1.4})$  values for each redshift bin, rows are aligned for similar  $\log(L_{1.4})$  centers.

0.005 < z < 0.26			0.26 < z < 0.58			0.58 < z < 1.0		
$\log(L_{1.4})$	$N$	$\log(\phi)$	$\log(L_{1.4})$	$N$	$\log(\phi)$	$\log(L_{1.4})$	$N$	$\log(\phi)$
(W Hz $^{-1}$ )		(Mpc $^{-3}$ dex $^{-1}$ )	(W Hz $^{-1}$ )		(Mpc $^{-3}$ dex $^{-1}$ )	(W Hz $^{-1}$ )		(Mpc $^{-3}$ dex $^{-1}$ )
20.31	1	$-2.48^{+0.43}_{-0.04}$	-	-	-	-	-	-
20.74	3	$-2.94^{+0.25}_{-0.64}$	-	-	-	-	-	-
21.19	18	$-2.48^{+0.04}_{-0.20}$	-	-	-	-	-	-
21.59	54	$-2.65^{+0.01}_{-0.11}$	-	-	-	-	-	-
22.06	121	$-2.85^{+0.05}_{-0.09}$	-	-	-	-	-	-
22.36	174	$-2.99^{+0.01}_{-0.07}$	-	-	-	-	-	-
22.78	46	$-3.61^{+0.00}_{-0.13}$	22.60	152	$-3.88^{+0.01}_{-0.05}$	-	-	-
23.11	9	$-4.27^{+0.02}_{-0.25}$	22.94	257	$-3.87^{+0.01}_{-0.04}$	-	-	-
23.61	7	$-4.43^{+0.01}_{-0.20}$	23.36	133	$-4.21^{+0.01}_{-0.06}$	23.37	175	$-4.88^{+0.02}_{-0.06}$
24.11	3	$-4.75^{+0.08}_{-0.33}$	23.85	26	$-4.87^{+0.03}_{-0.12}$	23.80	163	$-4.91^{+0.01}_{-0.06}$
24.53	1	$-5.13^{+0.27}_{-0.11}$	24.32	15	$-5.23^{+0.03}_{-0.19}$	24.44	57	$-5.40^{+0.03}_{-0.08}$
24.96	1	$-5.33^{+0.26}_{-0.04}$	24.78	7	$-5.48^{+0.09}_{-0.28}$	-	-	-
25.38	2	$-4.92^{+0.12}_{-0.46}$	25.27	4	$-5.74^{+0.11}_{-0.33}$	25.19	21	$-5.79^{+0.03}_{-0.06}$
-	-	-	25.73	3	$-6.00^{+0.20}_{-0.38}$	25.92	8	$-6.20^{+0.12}_{-0.20}$

### 3.6 SPLIT RADIO LUMINOSITY FUNCTIONS

Due to the fundamentally different emission mechanisms of star-forming galaxies and AGN, these classes are expected to follow distinct cosmic evolution. To assess this, 1.4 GHz radio lumi-



**Figure 3.8:** The 1.4 GHz radio luminosity functions for the ATLAS sample over three consecutive redshift bins (indicated in each legend). Errors for each bin are plotted as the the *first* and *fourth* quartile of Monte-Carlo resampled data over the width of the bins. A histogram of per-bin source counts is embedded in the top-right of each panel. In the first panel the ATLAS RLF is compared to Mao et al. (2012); Best & Heckman (2012); Pracy et al. (2016) RLFs, in higher redshift spans  $z < 0.30$  Pracy et al. (2016) is retained for a low-redshift reference.

nosity functions are constructed for these classes independently across the same redshift bins as the complete RLF. This segregation follows the classifications established in § 2.

### 3.6.1 STAR-FORMING GALAXIES

Figure 3.9 shows the star-forming RLF across the same three redshift bins used for the full ATLAS RLF. Unfortunately, there is little sample overlap between ATLAS  $z < 0.26$  and compared literature past  $L_{1.4} > 10^{23} \text{ W Hz}^{-1}$ . This is likely to be a manifestation of differences in selection effects, and classification techniques between surveys as well as the differences in sample sizes (only 358 ATLAS sources were classified as star-forming galaxies, this corresponds to 7 per cent of the 4903 ATLAS sources, and 39 per cent of the 925 available spectra).

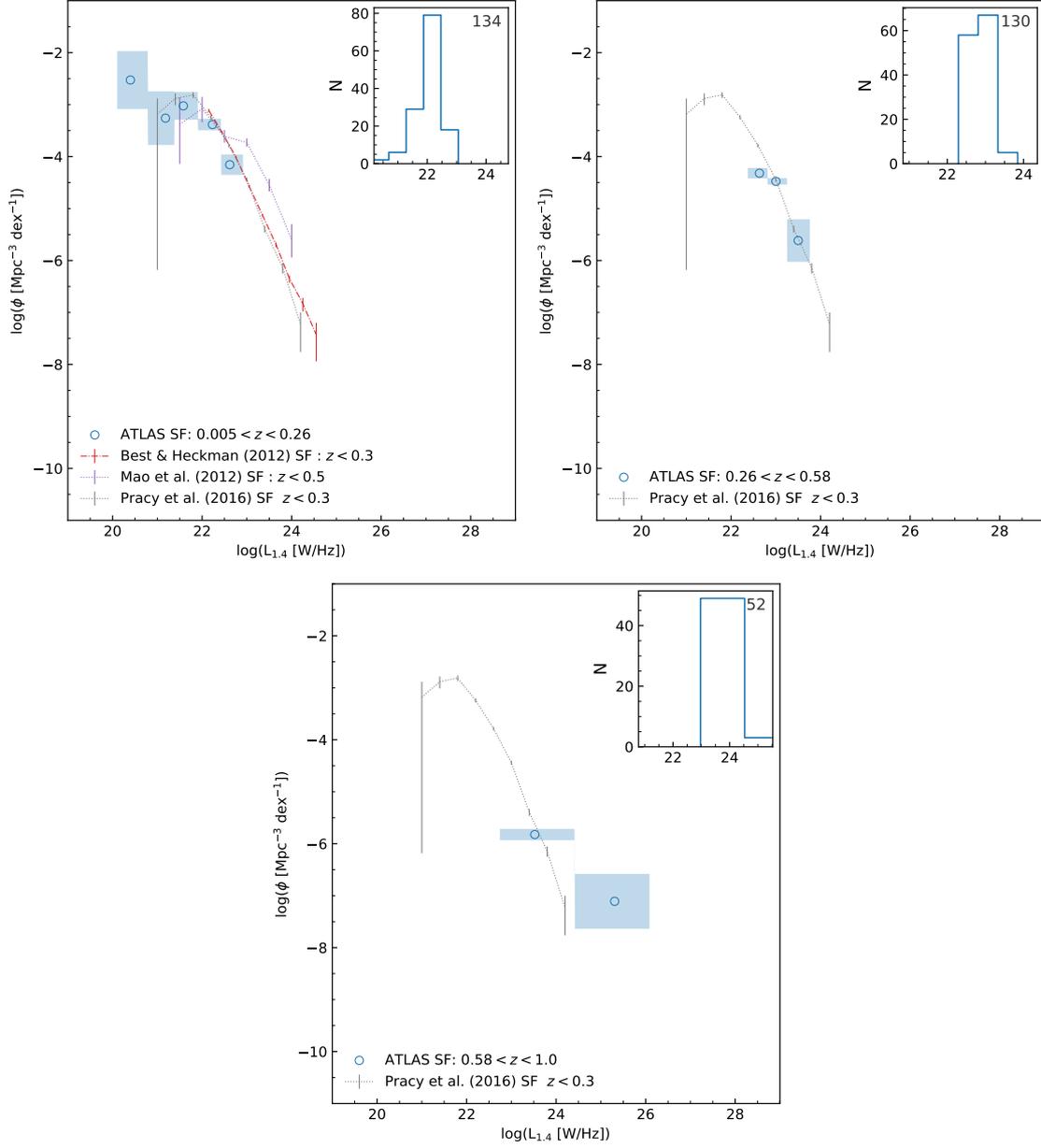
The ATLAS SF RLF does serve to challenge the established low-luminosity RLF. At lower luminosities ( $L_{1.4} < 10^{22} \text{ W Hz}^{-1}$ ) the ATLAS SF RLF begins to turn-over indicating an  $S_{1.4}$  incompleteness at approximately the same location as [Pracy et al. \(2016\)](#). For ATLAS, however, this turn-over is interrupted by the lowest luminosity bin ( $L_{1.4} = 10^{20.4} \text{ W Hz}^{-1}$ ; see Table 3.2) which suggests a continuation of the general shape of the RLF. Comparison with comprehensive infrared surveys that establish star-forming rates and infrared luminosity functions out to higher redshifts ( $z \sim 2$  [Magnelli et al. 2013](#);  $z \sim 4$ : [Grupponi et al. 2013](#)) reveals this turn-over to be a spurious result (that is, these studies see no break at similar luminosities over a wide redshift range). Direct comparison with these infrared studies would require considerations of the different selection effects between 1.4 GHz and far-infrared populations, and more comprehensive SF galaxy classification of the ATLAS sample and is therefore left for future work.

In the next redshift range  $0.26 < z < 0.58$ , the ATLAS SF RLF maintains an agreement with [Pracy et al. \(2016\)](#)  $z < 0.3$  SF volume density indicating no cosmic evolution of the ATLAS SF population. As discussed in § 3.1 it is expected that the source density of SF galaxies shows negligible evolution, and it is not surprising that the limited ATLAS sample is not sensitive enough to observe a small change in source density.

The ATLAS star-forming RLF constructed for the final redshift range ( $0.58 < z < 1.0$ ) consist of only two bins assembled from low source counts. The lower luminosity bin maintains consistency with  $z < 0.3$  [Pracy et al. \(2016\)](#) source densities, whereas the higher luminosity bin indicates (at least superficially) that there is a population of bright ( $L_{1.4} = 10^{24} \text{ W Hz}^{-1}$ ) star-forming galaxies unseen in [Pracy et al. \(2016\)](#). As there are only three sources in this bin, it is unlikely that this is the case, a more likely explanation would be that these sources are AGN misclassified as SF galaxies. Upon visual investigation of these sources, it appears that they are exceptionally bright point-like sources with (noisy) AGN spectral characteristics. These sources meet the criteria for the classification-override for the reinspected sources that deviated from the radio-infrared correlation (see § 2.8 Page 43). This misclassification is further testament to the necessity of conducting visual classifications multiple times, and the adoption of consistency checks (such as those used in § 2.7) to prevent population contamination.

**Table 3.2:** The 1.4 GHz radio luminosity functions for ATLAS sources over three consecutive redshift bins as indicated by the top-most row. The table includes the RLF for galaxies classified as star-forming, active galactic nuclei, and unknown (i.e., no classification was resolved in § 2). Due to the dynamic binning (see § 3.3.1) the RLFs are centered around a different set of  $\log(L_{1.4})$  values for each redshift bin of each classification, rows are aligned for similar  $\log(L_{1.4})$  centers.

0.005 < z < 0.26			0.26 < z < 0.58			0.58 < z < 1.0		
$\log(L_{1.4})$	$N$	$\log(\phi)$	$\log(L_{1.4})$	$N$	$\log(\phi)$	$\log(L_{1.4})$	$N$	$\log(\phi)$
(W Hz <sup>-1</sup> )		(Mpc <sup>-3</sup> dex <sup>-1</sup> )	(W Hz <sup>-1</sup> )		(Mpc <sup>-3</sup> dex <sup>-1</sup> )	(W Hz <sup>-1</sup> )		(Mpc <sup>-3</sup> dex <sup>-1</sup> )
<i>Star-Forming</i>								
20.40	2	-2.53 <sup>+0.16</sup> <sub>-0.40</sub>	-	-	-	-	-	-
21.18	6	-3.26 <sup>+0.12</sup> <sub>-0.39</sub>	-	-	-	-	-	-
21.58	29	-3.02 <sup>+0.05</sup> <sub>-0.22</sub>	-	-	-	-	-	-
22.23	79	-3.38 <sup>+0.04</sup> <sub>-0.07</sub>	-	-	-	-	-	-
22.62	18	-4.16 <sup>+0.07</sup> <sub>-0.13</sub>	22.63	58	-4.32 <sup>+0.02</sup> <sub>-0.08</sub>	-	-	-
-	-	-	23.00	67	-4.48 <sup>+0.01</sup> <sub>-0.05</sub>	-	-	-
-	-	-	23.50	5	-5.62 <sup>+0.17</sup> <sub>-0.24</sub>	23.52	49	-5.82 <sup>+0.02</sup> <sub>-0.09</sub>
-	-	-	-	-	-	25.31	3	-7.11 <sup>+0.24</sup> <sub>-0.29</sub>
<i>Active Galactic Nuclei</i>								
21.68	7	-3.47 <sup>+0.23</sup> <sub>-0.65</sub>	-	-	-	-	-	-
22.36	41	-3.92 <sup>+0.01</sup> <sub>-0.10</sub>	-	-	-	-	-	-
22.98	6	-4.81 <sup>+0.03</sup> <sub>-0.30</sub>	23.04	92	-4.83 <sup>+0.03</sup> <sub>-0.05</sub>	23.53	52	-5.45 <sup>+0.04</sup> <sub>-0.05</sub>
24.22	2	-5.20 <sup>+0.10</sup> <sub>-0.38</sub>	24.47	12	-5.94 <sup>+0.10</sup> <sub>-0.17</sub>	24.11	22	-5.86 <sup>+0.08</sup> <sub>-0.13</sub>
25.13	2	-5.25 <sup>+0.02</sup> <sub>-0.46</sub>	-	-	-	24.75	5	-6.50 <sup>+0.15</sup> <sub>-0.50</sub>
-	-	-	-	-	-	25.36	6	-6.39 <sup>+0.07</sup> <sub>-0.15</sub>
<i>Unknown</i>								
21.22	11	-2.84 <sup>+0.00</sup> <sub>-0.35</sub>	-	-	-	-	-	-
21.81	53	-2.92 <sup>+0.00</sup> <sub>-0.13</sub>	-	-	-	-	-	-
22.28	127	-3.14 <sup>+0.03</sup> <sub>-0.09</sub>	-	-	-	-	-	-
22.76	42	-3.76 <sup>+0.01</sup> <sub>-0.11</sub>	22.68	125	-4.14 <sup>+0.01</sup> <sub>-0.07</sub>	-	-	-
-	-	-	23.13	171	-4.22 <sup>+0.02</sup> <sub>-0.04</sub>	-	-	-
23.59	9	-4.42 <sup>+0.03</sup> <sub>-0.17</sub>	23.59	41	-4.83 <sup>+0.04</sup> <sub>-0.08</sub>	23.48	160	-5.00 <sup>+0.05</sup> <sub>-0.04</sub>
23.93	2	-5.12 <sup>+0.22</sup> <sub>-0.38</sub>	24.18	16	-5.21 <sup>+0.04</sup> <sub>-0.20</sub>	24.02	89	-5.25 <sup>+0.03</sup> <sub>-0.06</sub>
24.49	2	-5.09 <sup>+0.09</sup> <sub>-0.41</sub>	24.84	5	-5.71 <sup>+0.12</sup> <sub>-0.20</sub>	24.82	28	-5.82 <sup>+0.07</sup> <sub>-0.08</sub>
-	-	-	25.47	3	-5.95 <sup>+0.10</sup> <sub>-0.50</sub>	25.59	8	-6.27 <sup>+0.11</sup> <sub>-0.14</sub>

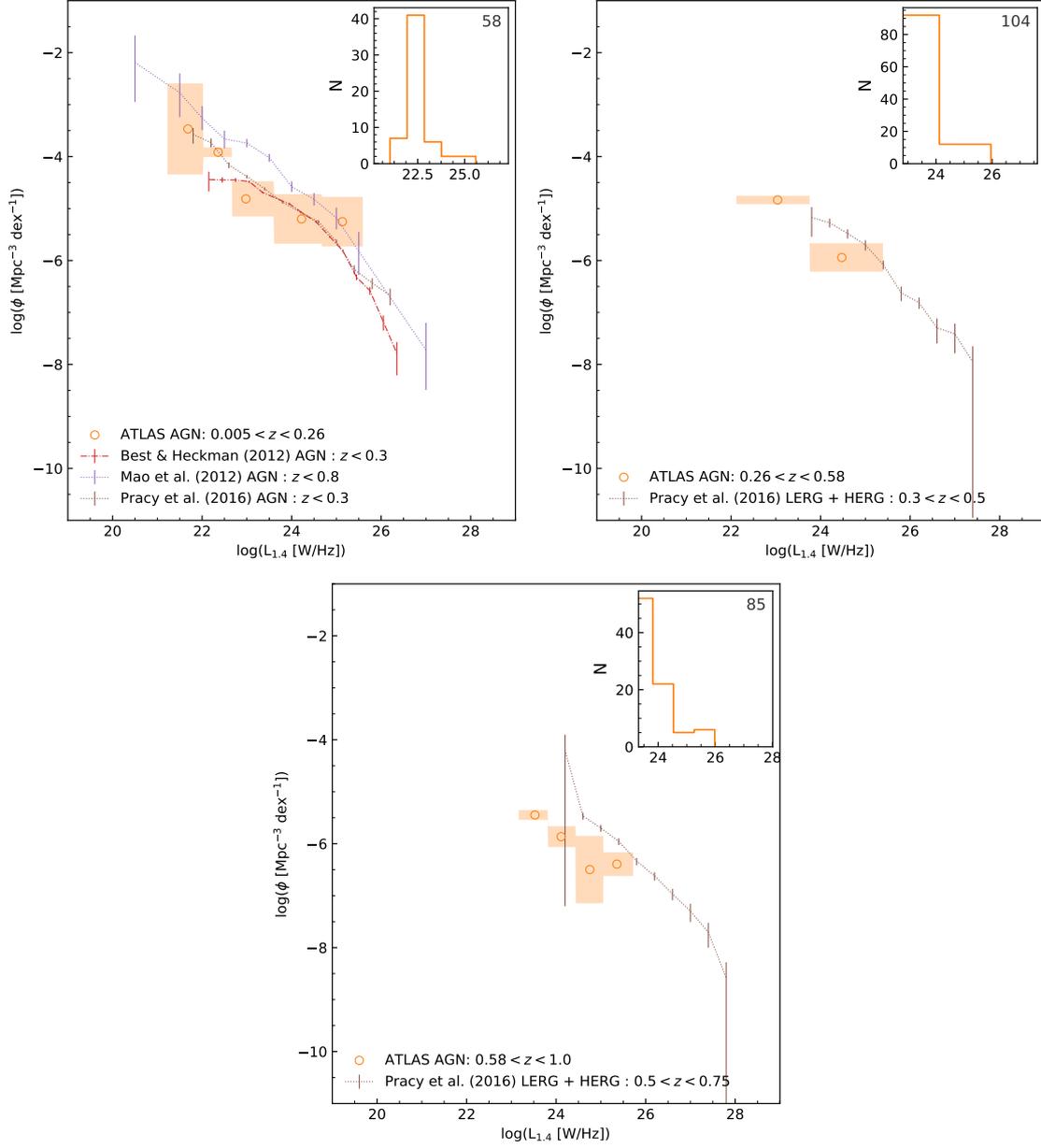


**Figure 3.9:** The 1.4 GHz radio luminosity functions for ATLAS star-forming galaxies over three consecutive redshift bins (indicated in each legend). Errors for each bin are plotted as the the *first* and *fourth* quartile of Monte-Carlo resampled data over the width of the bins. A histogram of per-bin source counts is embedded in the top-right of each panel. In the first panel the ATLAS RLF is compared to Mao et al. (2012); Best & Heckman (2012); Pracy et al. (2016) SF RLFs, in higher redshift spans  $z < 0.30$  Pracy et al. (2016) is retained for a low-redshift reference.

### 3.6.2 ACTIVE GALACTIC NUCLEI

Despite the ATLAS sample containing only 58 AGN for  $z < 0.26$  the AGN RLF agrees well with Pracy et al. (2016) and Best & Heckman (2012). Additionally, the ATLAS RLF probes

<sup>2</sup>Because Pracy et al. (2016) does not present an ‘AGN’ RLF across multiple redshifts, the summed co-moving source density of their HERG and LERG populations are used.



**Figure 3.10:** The 1.4 GHz radio luminosity functions for ATLAS active galactic nuclei over three consecutive redshift bins (indicated in each legend). Errors for each bin are plotted as the the *first* and *fourth* quartile of Monte-Carlo resampled data over the width of the bins. A histogram of per-bin source counts is embedded in the top-right of each panel. In the first panel the ATLAS RLF is compared to Mao et al. (2012); Best & Heckman (2012); Pracy et al. (2016) AGN RLFs, in higher redshift spans Pracy et al. (2016) AGN RLF<sup>2</sup> for complimentary redshift spans are used for comparison.

lower luminosities than Best & Heckman (2012), confirming the measured source-density of faint Pracy et al. (2016) sources. Therefore, it is likely then that the turnover shown in the Best

& Heckman (2012) sample is due to  $S_{1.4}$  incompleteness. The ATLAS  $L_{1.4} = 10^{23} \text{ W Hz}^{-1}$  bin does deviate slightly, but this is within uncertainty. However the  $\Delta\phi \sim 0.5$  dex disagreement at  $L_{1.4} = 10^{25.13} \text{ W Hz}^{-1}$  is significant, but is likely to be a spurious difference rather than a genuine result.

Shifting the AGN redshift window to  $0.26 < z < 0.58$  finds a general agreement of population source densities with Pracy et al. (2016)  $0.30 < z < 0.50$  AGN. The ATLAS sources are only separated into two bins, but it does sample fainter luminosities than Pracy et al. (2016) suggesting a continuation of the RLF shape established in their larger sample.

The highest ATLAS redshift window  $0.58 < z < 1.0$  shows poor agreement with Pracy et al. (2016)  $0.50 < z < 0.75$  AGN RLF, with  $\Delta\phi > 1.0$  dex for the the largest disparity. A difference of this magnitude is likely to be an issue with data completeness, and I will demonstrate that the inclusion of the sources previously classified as unknown (see § 2.7 Page 41) yields a much better agreement.

### 3.6.3 UNKNOWN CLASSIFICATIONS

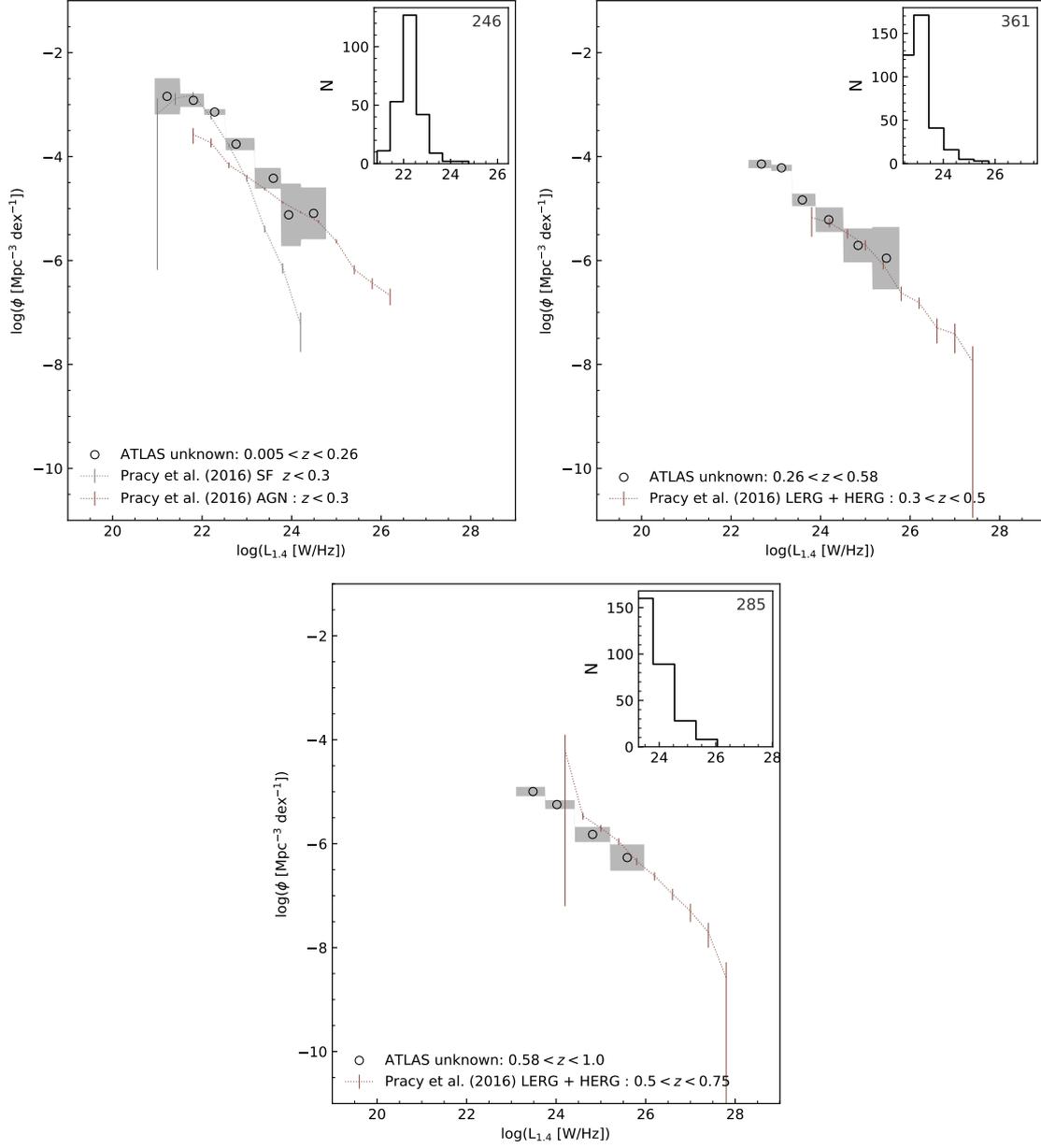
Comparing the RLF of the unknown ATLAS sources to SF and AGN RLF (Figure 3.11) reveals the presence of mixed populations within the unknown class. In particular, the lowest redshift RLF ( $z < 0.26$ ) agrees (generally) with the SF RLF at low luminosities and the AGN RLF at higher luminosities. It is likely that this  $z > 0.26$  sample of unknown sources is a genuine mix of SF galaxies and AGN. Comparatively, at higher redshifts, the SF component of the RLF disappears and the ATLAS unknown RLF agrees excellently with Pracy et al. (2016) AGN RLF (from similar redshifts) and it is therefore likely that the unknown population at high redshifts is dominated by AGN.

### 3.6.4 BLENDED CLASSIFICATIONS

Figure 3.12 considers all unknown classifications beyond  $z > 0.26$  to be AGN and combines these with the classified AGN for the two redshift windows  $0.26 < z < 0.58$  and  $0.58 < z < 1.0$ . Both of these RLF agree better with the corresponding Pracy et al. (2016) AGN RLF than the ATLAS AGN sources alone (Figure 3.10). The additional sources also serve to establish the faint end of AGN RLF at these redshifts. It is very likely that the unclassified sources detected at these distances are indeed AGN that lacked sufficient spectral information for classification (i.e., LERGs).

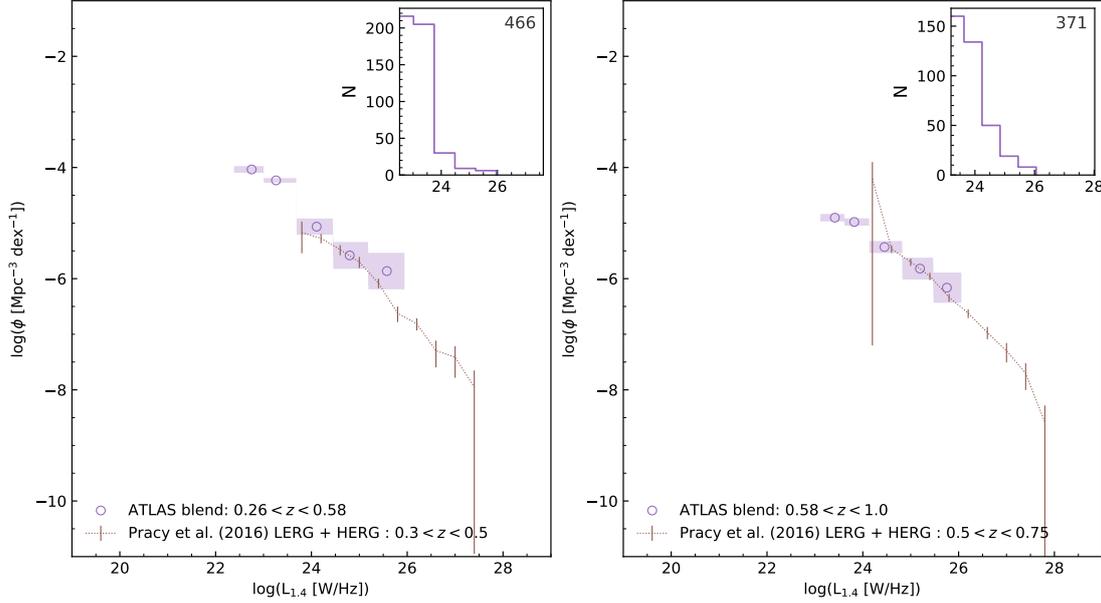
### 3.6.5 HIGH-REDSHIFT AGN

There are an additional 169 sources beyond  $z > 1.0$ : 62 AGN, 18 SF, and 89 unknown. Figure 3.13 is a construction of an RLF using the combined AGN and unknown sources. Compared to the Pracy et al. (2016) lower redshift AGN sample, the co-moving source density of ATLAS AGN appears to have decreased with a significant increase in redshift. This is a result contrary



**Figure 3.11:** The 1.4 GHz radio luminosity functions for ATLAS ‘unknown’ classified sources over three consecutive redshift bins (indicated in each legend). Errors for each bin are plotted as the *first* and *fourth* quartile of Monte-Carlo resampled data over the width of the bins. A histogram of per-bin source counts is embedded in the top-right of each panel. In the first panel the ATLAS RLF is compared to Pracy et al. (2016) AGN and SF RLFs, in higher redshift spans Pracy et al. (2016) AGN RLF for complimentary redshift spans are used for comparison.

to other findings which suggest an increase in AGN source density of up to three orders of magnitude (e.g., Clewley & Jarvis 2004; Donoso et al. 2009). Due to the poor spectroscopic complete-



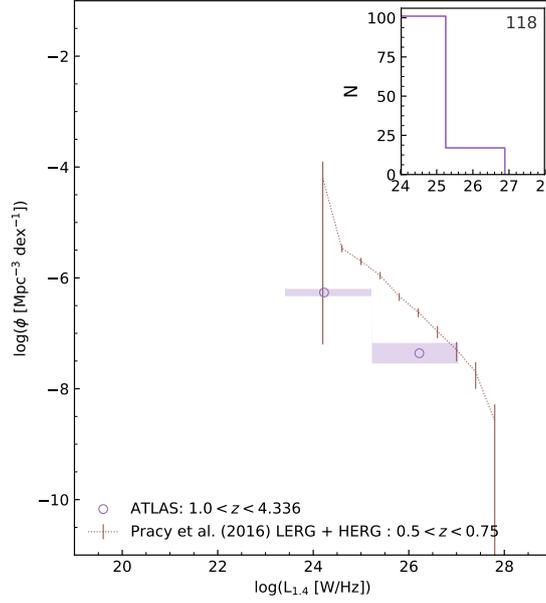
**Figure 3.12:** The 1.4 GHz radio luminosity functions for the combined populations of ATLAS ‘unknown’ classified sources and ATLAS AGN over two consecutive redshift bins (indicated in each legend). Errors for each bin are plotted as the the *first* and *fourth* quartile of Monte-Carlo resampled data over the width of the bins. A histogram of per-bin source counts is embedded in the top-right of each panel. In each panel the ATLAS RLF is compared to Pracy et al. (2016) AGN RLF for complimentary redshift spans.

ness past  $z > 1.0$  (see Figure 3.4) of the ATLAS sample, it is not surprising that positive source-density evolution is not observed. Therefore, the ATLAS results should only be considered as a lower limit to the AGN RLF for  $z > 1.0$ .

### 3.7 FINAL THOUGHTS

The  $S_{1.4}$  completeness of each 1.4 GHz luminosity bin is essential for constructing an accurate radio luminosity function. The ATLAS data pushes the sensitivity limit of these luminosity functions with a flux density limit 10 – 100 times more sensitive than similar surveys (e.g., Machalski & Godlowski (2000); Mauch & Sadler (2007); Best & Heckman (2012)). The deeper ATLAS survey confirms literature results and helps to establish information on the comoving source-density at lower luminosities. These ATLAS results also provide a significant improvement over ATLAS DR1 findings (Mao et al. 2012).

Fewer than half of the ATLAS sources have a redshift measurement necessary to build RLF, while this incompleteness is accounted for (see § 3.4.1) the missing sources cannot be perfectly accounted for with a correction and this affects the conclusive power of the results. When extending analysis further by defining sub-populations of galaxies this effect is magnified: I was only able to reliably classify several hundred galaxies (compared to several thousand in similar



**Figure 3.13:** The high-redshift 1.4 GHz radio luminosity functions for the combined populations of ATLAS ‘unknown’ classified sources and ATLAS AGN. This RLF samples galaxies beyond  $z > 1$  to the most distant ATLAS galaxy:  $z < 4.34$ . Errors for each bin are plotted as the the *first* and *fourth* quartile of Monte-Carlo resampled data over the width of the bins, this error only reflects the sample uncertainty and does not represent the uncertainty due to the severe spectral incompleteness of ATLAS data beyond  $z > 1.0$ . A histogram of per-bin source counts is embedded in the top-right of each panel. The ATLAS RLF is compared to the highest redshift [Pracy et al. \(2016\)](#) AGN RLF.

studies).

Despite the imperfect data, this ATLAS sample is sufficient to confirm the established RLF of full samples, star-forming galaxies and active galactic nuclei as well as constrain lower-luminosity source densities. The evolution of [Pracy et al. \(2016\)](#) AGN was also confirmed, and the lack of evidence of an evolving star-forming population in the ATLAS sample suggests that star-forming galaxies must evolve less strongly with redshift than AGN. This result is contrary to established literature SF galaxy population evolution (e.g., [Magnelli et al. 2013](#); [Gruppioni et al. 2013](#)) which demonstrate strong evolution of SF galaxies to  $z \sim 3$ .

A deeper analysis of these ATLAS RLFs could be achieved by fitting power-law luminosity functions such as those in [Padovani et al. \(2015\)](#); [Pracy et al. \(2016\)](#); [Novak et al. \(2017\)](#). This would also allow the quantification and comparison of the redshift evolution of ATLAS RLFs. To achieve these results with sufficient precision more thorough SF and AGN classification would need to be undertaken; including higher resolution and higher bandwidth spectral fitting with pPXF, and further development and verification of the *classification value* developed in § 2.7. Such development would increase the classified fraction of ATLAS sources and potentially the redshift ranges over which comparisons can be made.

*The greater our knowledge increases, the greater our ignorance unfolds.*

– John F. Kennedy

# 4

## Discussion and Summary

The data presented in this thesis and the analyses conducted, demonstrate and establish robust techniques for cross-matching and classifying small ( $N < 10\,000$ ) galaxy surveys. In chapter one I designed an open-source Python program (MCVCM) that easily allows the visual cross-matching of multi-frequency catalogues: MCVCM is currently being used in GLASS (GAMA Legacy ATCA Southern Survey) and GLEAM (GaLactic and Extragalactic All-sky MWA Survey). Using MCVCM I improved the reliability of the ATLAS 1.4 GHz radio catalogue via a more complete accounting of extended radio sources with isolated components, and blended radio sources located in dense fields or from errors in automated catalogue assembly. To more thoroughly resolve these errors, MCVCM should be used on a catalogue of radio-flux peaks that have not been combined into radio sources.

I also developed a simple automatic technique for choosing an optimum radius for blind nearest neighbour cross-matching of catalogues, and demonstrated the improvement of cross-matching a radio catalogue to an optical spectroscopic or photometric catalogue via an infrared intermediate catalogue. Future work would see the development of a hybrid cross-matching method, establishing thresholds at which MCVCM is used in place of BNN cross-matching. The aim of such a hybrid system would be to maximise cross-matching reliability by using MCVCM on complex sources or sources in dense fields, but allowing BNN to reliably cross-match compact sources. This system would require further refinement of the BNN techniques used in this thesis, such as: a quantification of a ‘false match’ probability (e.g., [Novak et al. 2017](#)), and exploration of how the ‘blocking effect’ described in [Smolčić et al. \(2017\)](#) affects the reliability of

BNN cross-matching of faint optical or infrared counterparts.

This thesis has revealed that with simple BNN cross-matching there is a high probability of spurious matches, and that manual visual cross-matching with a tool such as MCVCM can alleviate these issues. Using the results of the MCVCM cross-matching of ATLAS via SWIRE infrared, I assembled a complete catalogue containing ATLAS 1.4 GHz data, DES *G, R, I, Z* photometry, SWIRE 3.6, 4.5, 5.8, 8.0, 24  $\mu\text{m}$  infrared, and OzDES redshifts which can be used by the ATLAS team in ongoing research, and the development of automated intelligent cross-matching techniques.

In chapter two, I classified AGN and SF sub-populations of the ATLAS radio galaxies using a combination of spectral emission diagnostics, visual spectral classification, and the radio-infrared correlation. I developed a simple method of combining these diagnostics that employs distance-to-line weighting of spectral line diagnostics to minimise the effect of spectral line diagnostics on ‘transition’ sources (i.e., sources that a diagnostic can not strongly classify). This combined classifier folds information from multiple sources, allowing the shortcomings of any isolated classification to be strengthened.

In chapter three I constructed Radio Luminosity Functions for the improved ATLAS catalogue developed in chapter one for the full ATLAS sample, as well as for the sub-populations classified in chapter two. I demonstrated a simple automatic RLF binning technique that removes potential user bias during the selection of binning and automatically adjusts for changes in data (for example, if redshift restrictions are applied or altered). The ATLAS RLF constructed in this manner agree well with established literature and improve the low-luminosity end of these RLFs due to the better sensitivity of the ATLAS survey. These new ATLAS RLFs improve dramatically upon the ATLAS DR1 RLF which were constructed from a much smaller sample (Mao et al. 2012), demonstrating the improvements of increased sample sizes and observational depth on population analyses.

The ATLAS SF population show no evolution of source density over a wide redshift, suggesting that any evolution is too small for the resolution of the classifications established in chapter two. I discovered that sources classified as ‘unknown’ in chapter two, when included in the AGN RLF, improved agreement with literature AGN RLF at high-redshift. These ‘blended’ AGN RLF show cosmic evolution that agree with the results of Pracy et al. (2016) as well as improving the accounting of low-luminosity sources at higher redshifts. This agreement suggests that the ‘combined classification’ established in chapter two is particularly sensitive to emission-line AGN (HERGs) and that the addition of other classification schemes are needed to reliably classify high-redshift LERGs. Chapter three demonstrates that a small sample of radio galaxies is adequate for building reliable radio luminosity functions if the catalogue assembly and source classifications are handled thoroughly.

The limited fraction of ATLAS sources classified in this thesis restricts these data to providing lower limits of population densities, and verification of established results below  $z < 0.75$ . Future work on these data would involve robust spectral classification through the use of better matched stellar templates during the pPXF spectral line fitting, and further development of the

classification value established in § 2.7. The inclusion of more ATLAS sources in SF and AGN classifications would allow for the fitting of power-law RLFs to these data, providing a means of directly comparing population turn-over luminosities and quantifying the evolution of these RLF with redshift.

It is fundamentally important that the techniques conducted in this thesis are refined before the next generations of data are processed. The manual techniques employed in this thesis are impossible to scale to the quantity of data produced in next-generation surveys. The scientific focus is consequently shifting towards automation of these data-processing steps via intelligent automated techniques (such as machine learning algorithms, Bayesian statistics, citizen science, or combinations thereof). The training or quality verification of all of these techniques rely upon relatively small sets (compared to a full survey) of data manually processed by one (or a few) experts. Invariably, there will be mistakes in these ‘golden samples’ and edge cases that the experts were unable to resolve. Additionally, intelligent techniques that are trained on samples of low-redshift galaxies, or on samples with an over-representation of a given galaxy type or environment will suffer from a class-imbalance.

If any of these issues persist into the training of intelligent automated techniques, or if the source of these issues are omitted from the training, then there is a risk that the errors will be systematically reproduced or that the edge cases will be numerous enough that the intelligent techniques are no more useful than blind nearest-neighbour cross-matching (in the case of multi-catalogue cross-matching) or diagnostic diagrams (in the case of galaxy classification), but with the additional cost of computing resources and software engineering.

Intelligent automated techniques are essential for this science to scale with growing technology, but these automated techniques must be constructed and trained with far greater statistical rigour than the processing of any individual survey as these intelligent techniques will invariably become black-boxes used naively by all but the few who constructed them. The refined ATLAS catalogue, MCVCM program, and techniques developed in this thesis endeavour to provide a basis to avoid these potentially devastating errors.



# MCVCM: A Python Program for Multi-Catalogue Visual Cross-Matching

## A.1 FOREWORD

This software was initially designed and coded by me, *the author*, to facilitate the cross-matching of the ATLAS 1.4 GHz radio source catalogue with DES photometric source catalogue. Initially, this cross-matching was performed directly between the radio data and optical  $I$ -band (e.g., Figure 1.7) and was later adapted to work with SWIRE 3.6  $\mu\text{m}$  infrared (achieving a more reliable cross-match, as established in § 1). During the necessary modifications to facilitate this, I endeavoured to make this tool generalised and adaptable enough to be usable with any catalogue pairings. Following this, I have made the code available in a public repository ([github.com/FriskyGrub/MCVCM](https://github.com/FriskyGrub/MCVCM)). MCVCM is being used in other projects – GLASS (GAMA Legacy ATCA Southern Survey) and GLEAM (GaLactic and Extragalactic All-sky MWA Survey) – which has encouraged the ongoing development of the software.

## A.2 DESCRIPTION

MCVCM (Multi-Catalogue Visual Cross-Matching) is written in `python 3.6` and is designed to facilitate the visual-manual cross-matching of two catalogues provided by the user.

### A.3 CUTOUT ASSEMBLY

When MCVCM is first run, the cutout generator creates a memory mapping of the mosaicked radio continuum, radio RMS and infrared images provided by the user. As the user iterates through catalogue sources in MCVCM sources, a slice of these images centred about the target source is retrieved for the generation of each cutout. Using this memory-mapping technique, the requisite data can be retrieved fast enough that the images need not be stored entirely in computer random access memory. Using a combination of `astropy world coordinate system` utilities and the `astropy` affiliated `reproject` tool, the radio data is re-sampled (via interpolation) to match the resolution and orientation of the infrared image. These data are then overlaid and displayed to the user using the `matplotlib imshow` and `contour` plotting utilities. User interactivity is established with a combination of `matplotlib canvas` listeners and Python's standard graphical user interface package `TkInter`.

#### A.3.1 RADIO CONTOUR LEVELS

MCVCM accepts a continuum map, as well as an RMS map for each of the fields being cross-matched. Using the sliced region of these data (about the source to be cross-matched) radio contours are calculated as  $2^n \times \langle \text{RMS} \rangle$  for this slice and overlaid onto the infrared image, this method allows for consistent contours throughout the radio field. I have not established a reliable means for drawing consistent contours when a complimentary RMS map is unavailable, but I will incorporate this in future iterations of MCVCM.

### A.4 USER INTERACTION

The script is designed to be called from within the folder containing `MCVCM.py` in the following manner:

```
>> MCVCM.py {CDF3,ELAIS,Field3} [-h] [-v] [-t] [-x] [-d] [--savefigs {png,eps,pdf}]
```

MCVCM requires the user to specify the field that they will be cross-matching as a positional argument (in this example either 'CDF3', 'ELAIS', or 'Field3'). These field options are automatically generated from the `./path_config.json` file once edited by the user.

positional arguments:

{CDFS,ELAIS,Field3} Specify if we are working on CDFS, ELAIS, or Field3

optional arguments:

-h, --help show this help message and exit  
-v toggles verbose output  
-t toggles output of function timings  
(requires verbose mode)  
-x if specified, MCVCM processes only sources  
previously skipped by the user  
--savefigs {png,eps,pdf}  
if provided with an extension, saves final  
cross-matched source cutout to that format  
inside ./MCVCM-figs/ (e.g. --savefigs png)

#### A.4.1 CONFIGURATION

Three configurations are needed once MCVCM has been installed to begin cross-matching of data. The first step is editing the `./path_config.json` file within the main directory to specify the paths to the required data files. The specified data files can exist anywhere on the user's machine (so long as the full file-path is provided) and are opened as read-only. The following is an example of the configuration file used for the cross-matching of the ATLAS radio sources with SWIRE infrared, CDFS and ELAIS are the two ATLAS fields, and Field3 is an example. This configuration file can contain unlimited field-data structures.

```

<./path_config.json>
{
  "ELAIS": {
    "radio_continuum": "data/ELAIS/ELAISmosaic_allch_8March2015.fits",
    "radio_rms": "data/ELAIS/ELAISmosaic_allch_noise_8March2015.fits",
    "image_mosaic": "data/ELAIS/elais_mosaic.fits",
    "radio_catalog": "data/ELAIS/ATLASDR3_ELAIscmpcat_23July2015.dat",
    "infrared_catalog": "data/ELAIS/SWIRE_ELAISS1-FALL05.fits"
  },
  "CDFS": {
    "radio_continuum": "data/CDFS/CDFSmosaic_allch_8March2015.fits",
    "radio_rms": "data/CDFS/CDFSmosaic_allch_noise_8March2015.fits",
    "image_mosaic": "data/CDFS/cdfs_mosaic.fits",
    "radio_catalog": "data/CDFS/ATLASDR3_CDFSscmpcat_23July2015.dat",
    "infrared_catalog": "data/CDFS/SWIRE-CDFS-FALL05.fits"
  }
  "Field3": {
    "radio_continuum": "path/to/continuum_map.fits",
    "radio_rms": "path/to/rms_map.fits",
    "image_mosaic": "path/to/image_mosaic.fits",
    "radio_catalog": "path/to/primary(radio)_source_catalogue.fits",
    "infrared_catalog": "path/to/secondary(infrared)_catalogue.fits"
  }
}

```

The second configuration provides information to MCVCM that allows it to read the correct columns from the source catalogues for source positions and identifications, the size (in image pixels) to be rendered from the image files, and two user-configuration options: the start-index (counting from 0, which row to start cross-matching from) and the figure position (this will be the position of the MCVCM cross-matching window on the users screen - (0,0) is top-left). The third configuration is adjusting the scaling of the infrared image for standardised source visibility.

```

<./parameter_config.json>
{
  "column_names": {
    "radio_ra": "RA_deg",
    "radio_dec": "Dec_deg",
    "radio_id": "ID",
    "infrared_ra": "ra",
    "infrared_dec": "dec",
    "infrared_id": "object"
  },
  "cutout_pixels": {
    "radio": 95,
    "infrared": 170
  },
  "start_index": 0,
  "figure_position": {
    "horizontal": 0,
    "vertical": 0
  }
  "image_scaling": {
    "max_saturation": 1.5,
    "power_normalise": 0.7
  }
}

```

#### A.4.2 CALIBRATING IMAGE

Calibration of the infrared image is done via a utility packaged with MCVCM. Once `./path_config.json` and `./parameter_config.json` are established by the user, the calibration script produces an MCVCM cutout with two variable sliders that adjust the normalisation factor, and maximum saturation threshold for the image scaling. When satisfied the user can enter the displayed numbers into the `"max_saturation"`, and `"power_normalise"` fields of `./parameter_config.json`.

#### A.4.3 CROSS-MATCHING

Once the MCVCM instance is created, the catalogue cross-matching is split into three phases:

1. **Infrared Host Selection:** The user is presented with a cutout image of the infrared postage stamp with radio contours overlaid (e.g., Figure A.1a). Infrared catalogue source positions are marked as small crosses that the user can select by clicking with the mouse. In

this phase the user may only make one selection, this will be marked with a large hollow black cross-hair that is retained through the remaining phases (Figure A.1d). If a mistake is made, the user can reset by pressing 'r'.

2. **Radio Core Selection:** The image remains the same, with the cross-hair marking the location of the selected infrared host still visible, and the infrared catalogue source positions are replaced with the radio component peak positions (Figure A.1d). The user can now select one of these sources as the radio core if there is no viable core this phase can be skipped.
3. **Radio Component Selection:** The image will remain the same as Phase 2 with a slight change to the radio component marker style to notify the user of the new phase (Figure A.1e). The user can now select any number of radio components that are associated with this cross-identification (Figure A.1f).

At any time during the cross-matching, the user may press one of the following keys or key combinations for the associated effect.

- {h} Print this help list
- {r} Reset current progress of source cross-matching and re-draw starting at Phase One
- {b} Increase the displayed area of the cutout by 33% (this will reset current source progress)
- {t} toggle on and off the marked catalogue source positions
- {s} Save an automatically named cutout of the current figure state to ./MCVCM-figs/
- {i} Print information on current source from the radio source catalogue
- {j} Print the last 25 rows of the cross-match catalogue being assembled in ./MCVCM-tables/
- {c} Open a dialogue box for the user to enter a comment for the current source. This comment is saved in the final table under the column 'MCVCM\_comment'
- {1, 2, 3, 4} provide an integer confidence flag for the cross-matching of this source
- {shift + x} Skip cross-matching of the current source, skipped sources can be cross-matched at any time by starting the script with the '-x' flag
- {shift + q} Safely close the MCVCM instance

Each phase is completed when the user presses ‘spacebar’, and the cross-matching of a particular source is finished when the user presses ‘d’ or ‘enter’, provided that at least one radio component has been selected. If no infrared core was selected in Phase one, then the cross-identification is given a ‘no-ir-source’ in place of the infrared source ID. If no core is selected, then the component ‘C0’ is used as the core for catalogue management. After cross-matching a unique association string is generated for each of the selected radio sources according to the following:

```
<infrared ID>*<radio core ID>*m<number of components>*C<component number>
```

For example:

```
Infrared host: SWIRE3_J003940.76-432549*EI1896*m3
Radio host:    SWIRE3_J003940.76-432549*EI1896*m3*C0
Component #1: SWIRE3_J003940.76-432549*EI1896*m3*C1
Component #2: SWIRE3_J003940.76-432549*EI1896*m3*C2
```

Note: Once a source is cross-matched it will be skipped by the MCVCM source incrementation, but it is not write-protected. During cross-matching, the MCVCM cutout will centre on the lowest-indexed radio component without a cross-match (skipping sources that were force-skipped by the user via ‘shift + x’). If during the cross-matching of this new source, the user deliberately or accidentally cross-matches a source other than the one that the catalogue is centred on, that cross-match is correctly stored in the cross-match catalogue (even if it was previously cross-matched). Once this cross-matching is complete MCVCM will again load the lowest indexed radio-source without a cross-match. This does allow the user to correct any radio-component mis-associations that they might have made earlier but it also allows for the possibility of erroneously altering an earlier cross-match.

#### A.4.4 SAVING A FIGURE

The saving of a figure can either be done manually or automatically. If MCVCM is started with the `--savefigs {png,eps,pdf}` flag then the final cross-matched image will be saved upon completion. This can take up to one second per source and is not recommended. The final version of MCVCM will include an option to iteratively generate these image cutouts once the entire catalogue has been cross-matched. Individual figures can be saved manually by the user either via the `matplotlib` save icon in the MCVCM window (the user will need to specify file name and location) or by pressing the ‘f’ key at any point during the cross-matching of the source.

#### A.5 VERSION CONTROL

The master catalogue which MCVCM is appending to as cross-matches are made is located at `./output/tables/<field>_mcvcm_table.dat`, there will be separate files for each `<field>`

being cross-matched. Each time MCVCM is resumed from the terminal, a copy of this file is made with an incremented numerical suffix. These act as back-ups of the cross-matching at the state when the MCVCM instance was launched. At any point, the user can revert to one of these states by following the ‘Rolling Back’ procedure.

#### A.5.1 ROLLING BACK

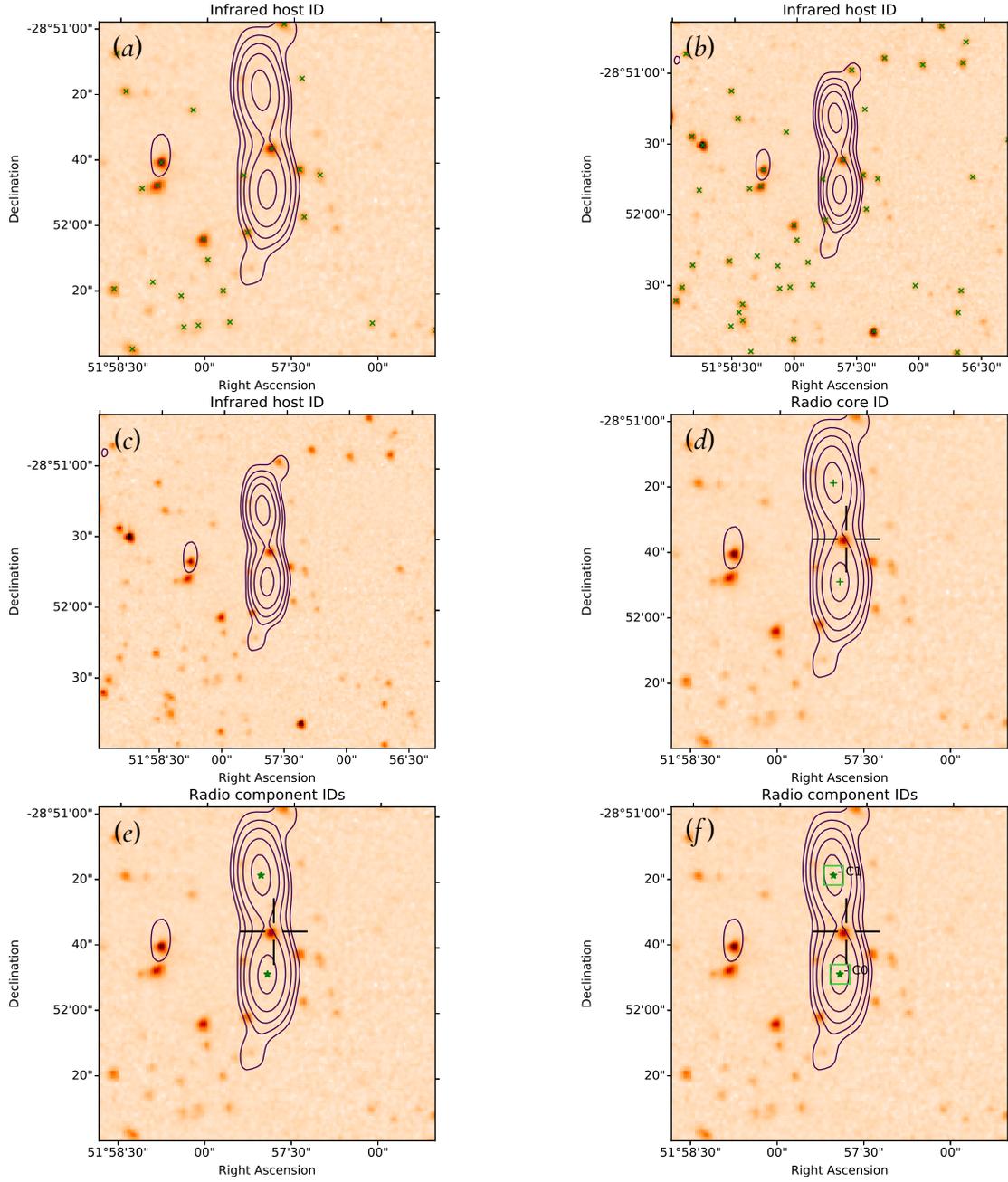
Rolling-back to a previous state of the cross-matched catalogue currently requires manual file management by the user, in the future this will be handled by MCVCM. To roll-back to a previous state locate the desired version (e.g., `./output/tables/<field>_mcvcm_table-015.dat`) and replace the master file (`./output/tables/<field>_mcvcm_table.dat`) with this version. It is a good idea to make a backup of both files in case of a mistake.

#### A.6 CATALOGUE MANAGEMENT

Once the cross-matching has been performed, a separate script needs to be run if the user wishes to create an amalgamated catalogue. Due to the vast differences in catalogue formats and conventions (for example, the filling of missing values), MCVCM does not include a generalised means of performing this merging. Upon request, I am happy to provide assistance in constructing this tool or provide the example used for the handling of the ATLAS data for reference.

#### A.7 CONTINUING DEVELOPMENT

Future work will see MCVCM handling as many complimentary data sets as the user desires, including the ability to over-lay multi-frequency radio contours and switch between (for example) infrared, optical and x-ray images on the fly. Supporting software will be developed to allow more robust version control and automated restoration, and final cross-matched catalogue assembly. MCVCM will continue to be developed along-side the users so that the user interface, user experience, and scientific utility are improved.



**Figure A.1:** Examples of various states during the three phases of source cross-identification in MCVCM. In all images, radio contours are overlaid onto an infrared image. (a) - Phase One: no action taken by user infrared catalogue sources are marked with green crosses. (b) - Phase One: user has incremented the field of view, infrared sources are marked (c) - Phase One: user has incremented the field of view and toggled off infrared sources (d) - Phase Two: user makes an infrared host selection in Phase One (black cross-hair) and has moved to Phase Two, radio source positions are now marked for core selection. (e) - Phase Three: identical to (d); no radio core was selected, and radio position markers have changed to indicate the new phase. (f) - Phase Three: the user has selected two radio components and is ready to finish this source cross-identification.

# B

## Measuring Cross-Matching Contamination via Random Perturbation versus Catalogue Offset

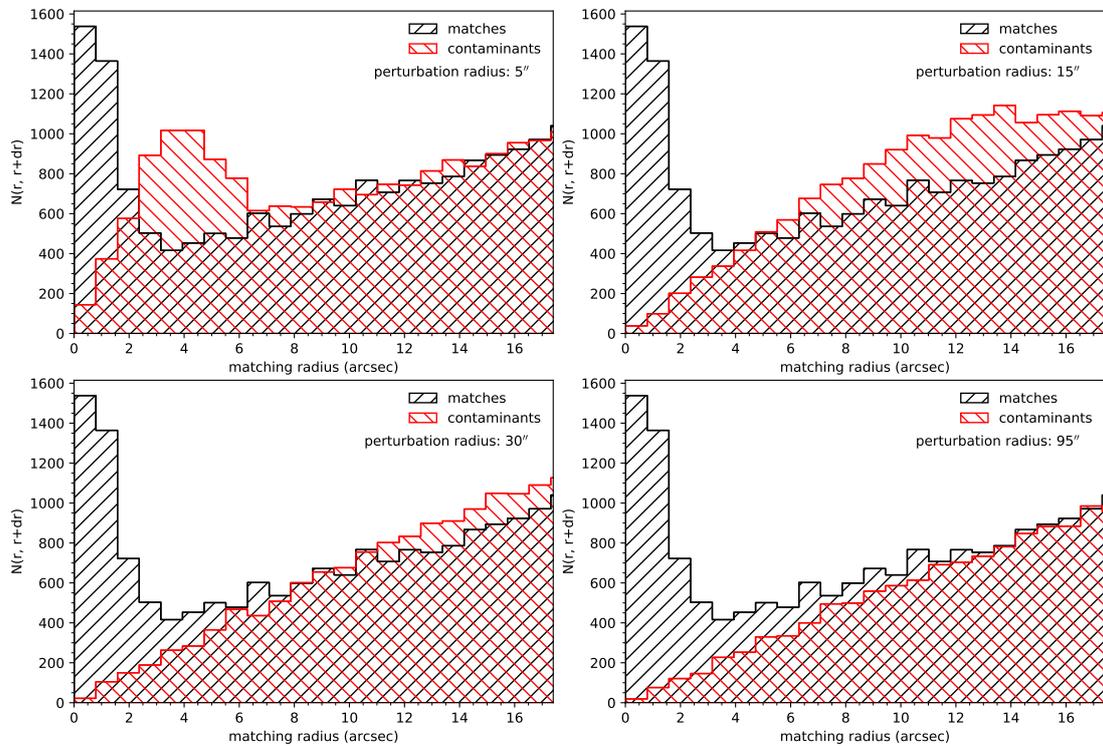
Contamination is not a measure of how many of the matches are wrong, that is unquantifiable, contamination is a measure of how many of the matches are purely due to chance alignment. That is, if we have two catalogues that cover the same sky area that are otherwise unassociated, how many matches would we expect for a given matching radius? Ideally, we would maintain source count and individually randomly perturb the positions of all sources in both catalogues. This preserves the critical representative density of sources whilst completely removing the physical association. Sufficiently, randomising one catalogue achieves the same result as any alignment 'correct' or otherwise is purely by chance.

A 'shifted-catalogue' technique only mimics randomness by taking advantage of the isotropic nature of our Universe. The observed source density of any particular region is expected to persist across the entire sky; the sheer quantity of sources makes any two regions *effectively* randomly unassociated. To invert this logic, a first-order model of the night sky (ignoring stars in our galaxy) would involve merely placing extra-galactic sources randomly across a black canvas. A second-order model would preferentially place these sources in clusters, and so on.

In the limiting case of two catalogues each containing a single source, a 'shifted catalogue'

technique would no longer be viable, as a decision on how far to shift the catalogue has to be made. At that point, it devolves to the question “do I align the sources, or not?”. Contrarily, in this scenario, the technique of randomly perturbing the source positions would still be valid.

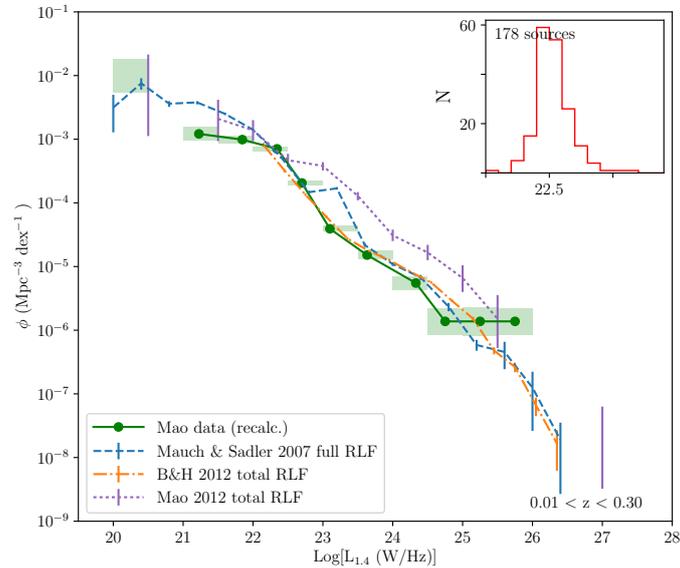
When generating the random catalogue, I tested the cross-matching to a randomised catalogue with perturbation radii of 5, 15, 30, 90'' on the cross-matching of ATLAS–SWIRE, this is shown in Figure B.1. For perturbation radii of 5'' and 15'' there is structure in the pairing of ATLAS sources to the randomised catalogue, but this almost disappears for a perturbation radius of 30'' and is entirely gone with a perturbation radius of 90''. For the generation of these random catalogues in § 1 I use a perturbation radius of 45'' for the generation of the random catalogue for all catalogue pairings. These tests were performed with the ATLAS–SWIRE catalogue pairing as the matching radii for these catalogues is much higher SWIRE–DES, for example. For catalogues that can be paired with a smaller cross-matching radius, a proportionally smaller perturbation radius can be used if desired.



**Figure B.1:** Visualisation of the effect of increasing cross-matching radius for blind nearest neighbour (BNN) cross-matching techniques. Catalogues being cross-matched are the ATLAS 1.4 GHz radio and the SWIRE 3.6  $\mu\text{m}$  catalogue. These four panels represent the difference in cross-match results to the random catalogue generated by source-position perturbation as discussed in § 1.6. The perturbation radius is noted in the top-right of each panel.

# C

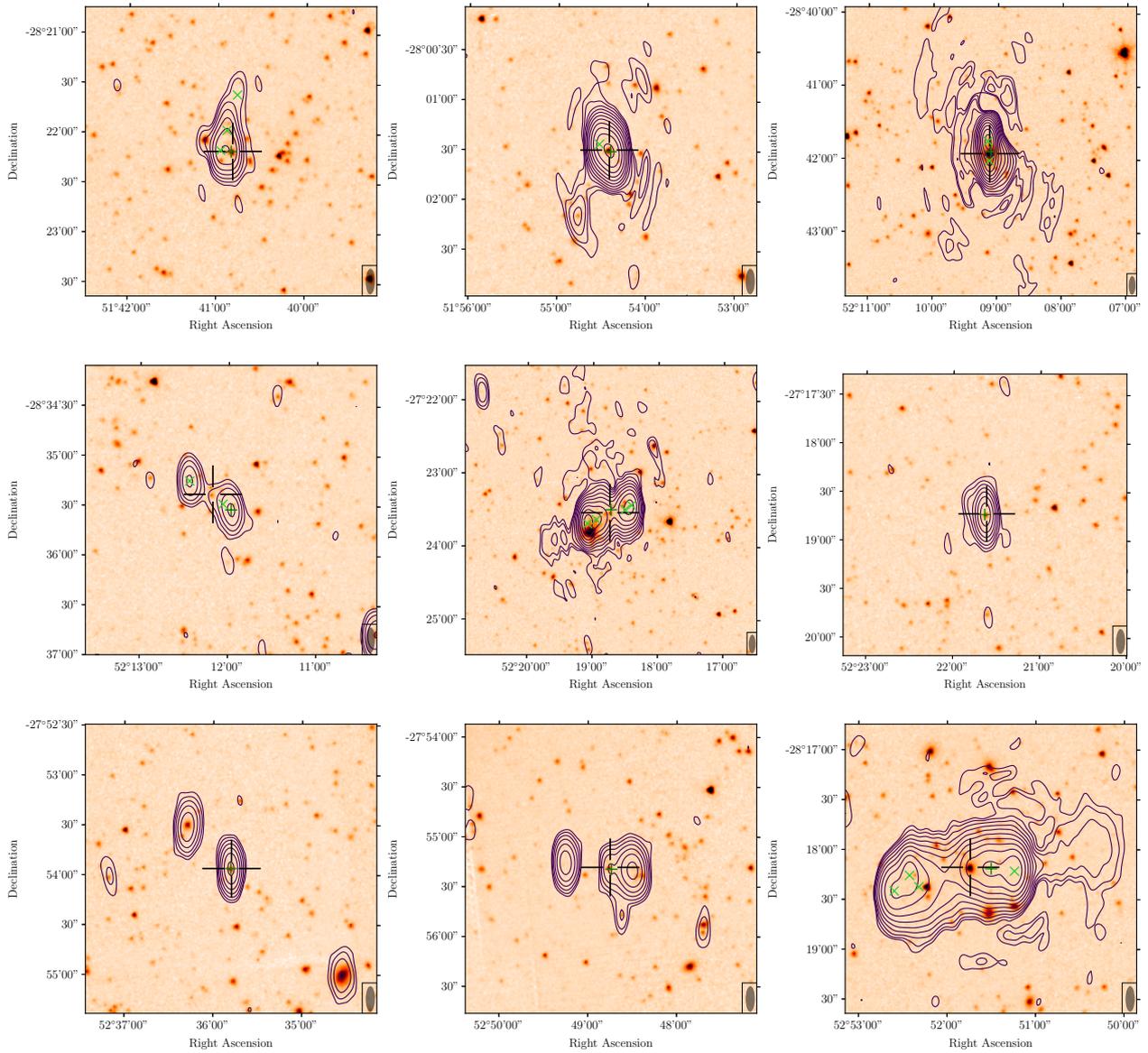
## Data and Figures



**Figure C.1:** Recalculation of [Mao et al. \(2012\)](#) 1.4 GHz radio luminosity function. This recalculated RLF (green) is generated using the original ATLAS-DR1 data with calculations performed by the RLF pipeline developed in this thesis (see § 3.2), and is compared to the published RLF (purple). Per-bin source count is displayed in the upper right.

**Table C.1:** Redshift quality flags from the various spectral sources included in the final OzDES catalogue. These are mapped to a standardised '*Final Flag*' based on qualitative descriptions of these source flags. The '*Final Flag*' are:  
 '−1' - Certain star; '4' - Certain redshift; '3' - Probable redshift; '2' - Possible redshift; '1' - Unknown redshift

<i>Source Flag</i>	<i>Final Flag</i>	<i>Source Flag</i>	<i>Final Flag</i>
NED		ACES	
"	2	'3'	3
'FoF'	2	'4'	4
'SN'	2	SpARCS	
'1LIN'	2	'1'	4
'TOMO'	2	'2'	3
'PAH'	2	'2.0'	4
'.'	2	'1.0'	3
'?'	1	6dF	
DES_AAOmega		'4'	4
'1'	1	NOAO_0334	
'3'	2	'3'	3
'2'	3	'4'	4
'4'	4	'6':	−1
'6'	−1	NOAO_0522	
VVDS_CDFS		'3'	3
'24'	4	'4'	4
'13'	4	'6':	−1
'14'	4	2dFGRS	
'23'	4	'3'	3
'3'	4	'5'	4
'4'	4	'4'	4
PRIMUS		ATLAS	
'3'	3	'.'	2
'4'	4		



**Figure C.2:** 1–9 of 23 sources classified as or confirmed to be AGN based on these cutouts, reinspection was prompted by their location in  $S_{24\mu m}/S_{1.4}$  vs.  $\log(1+z)$  (Figure 2.12). Cutouts were generated from MCVCVM, in each panel 1.4 GHz, ATLAS radio continuum contours are overlaid onto SWIRE  $3.6 \mu m$  infrared emission, contour levels are  $2^n \times \langle \text{RMS} \rangle$  calculated for each cutout. The large black cross-hair shows the location of the selected infrared core, and the green crosses show the location of the radio component peaks. In the bottom right of each image is the ATCA beam for that field.

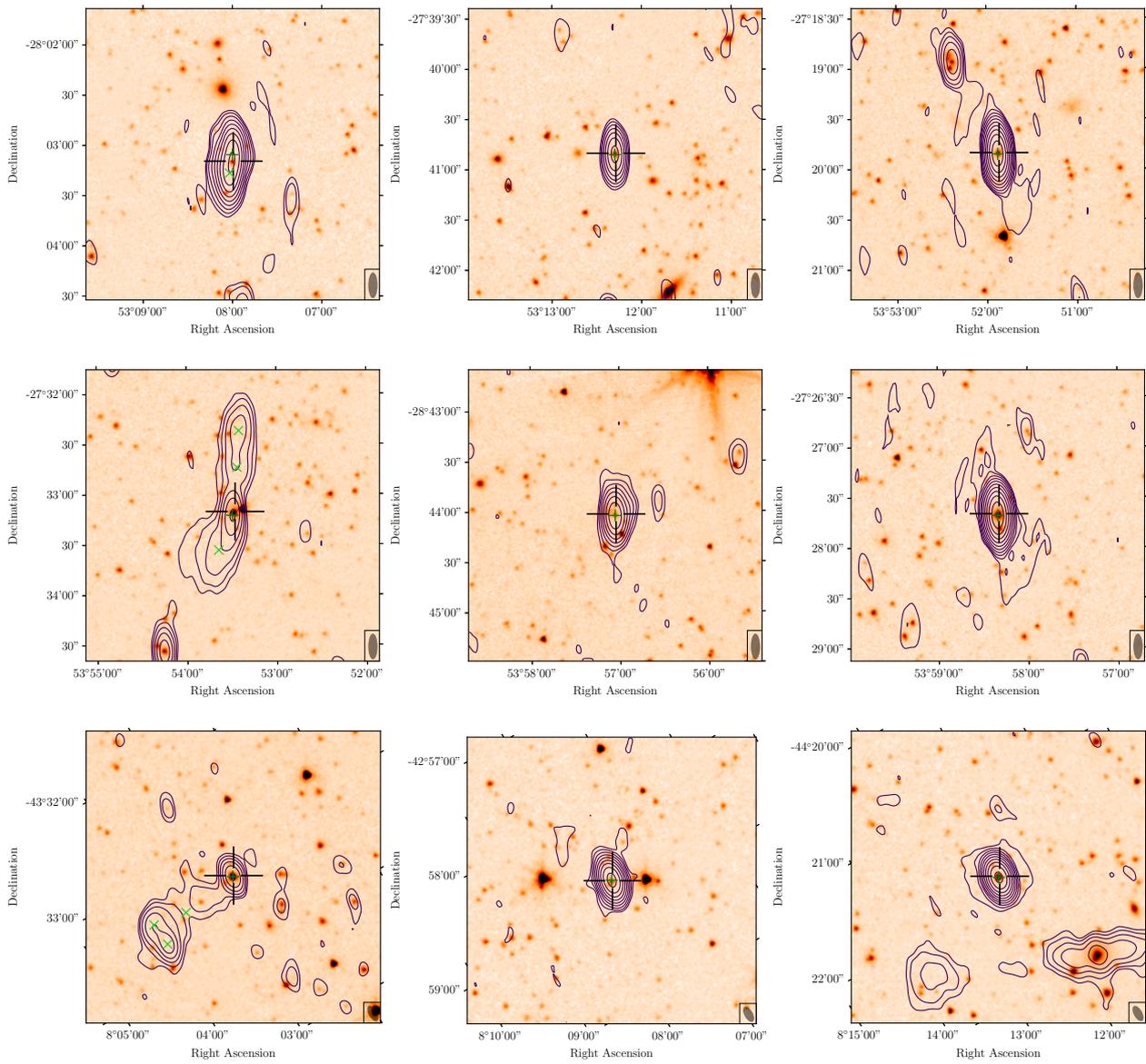


Figure C.3: Figure C.2 continued: 10–18 of 23 sources.

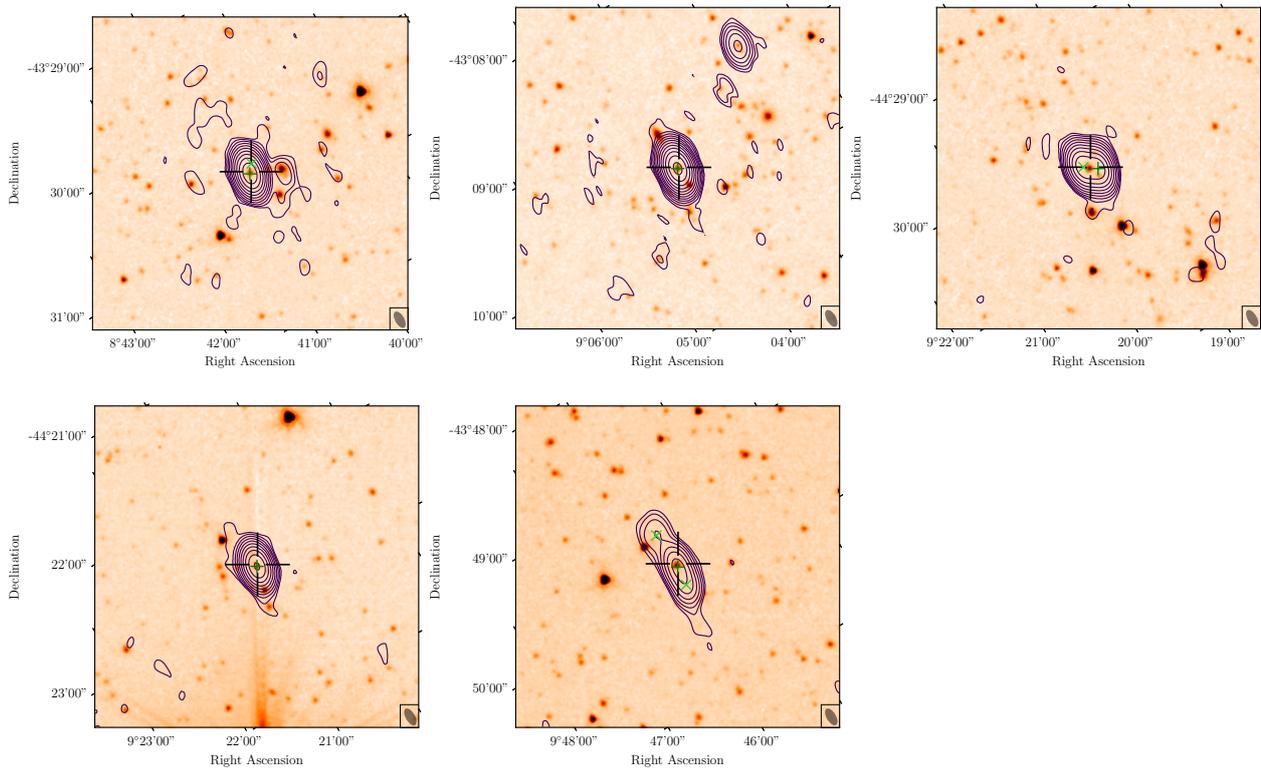


Figure C.4: Figure C.2 continued: 19–23 of 23 sources.



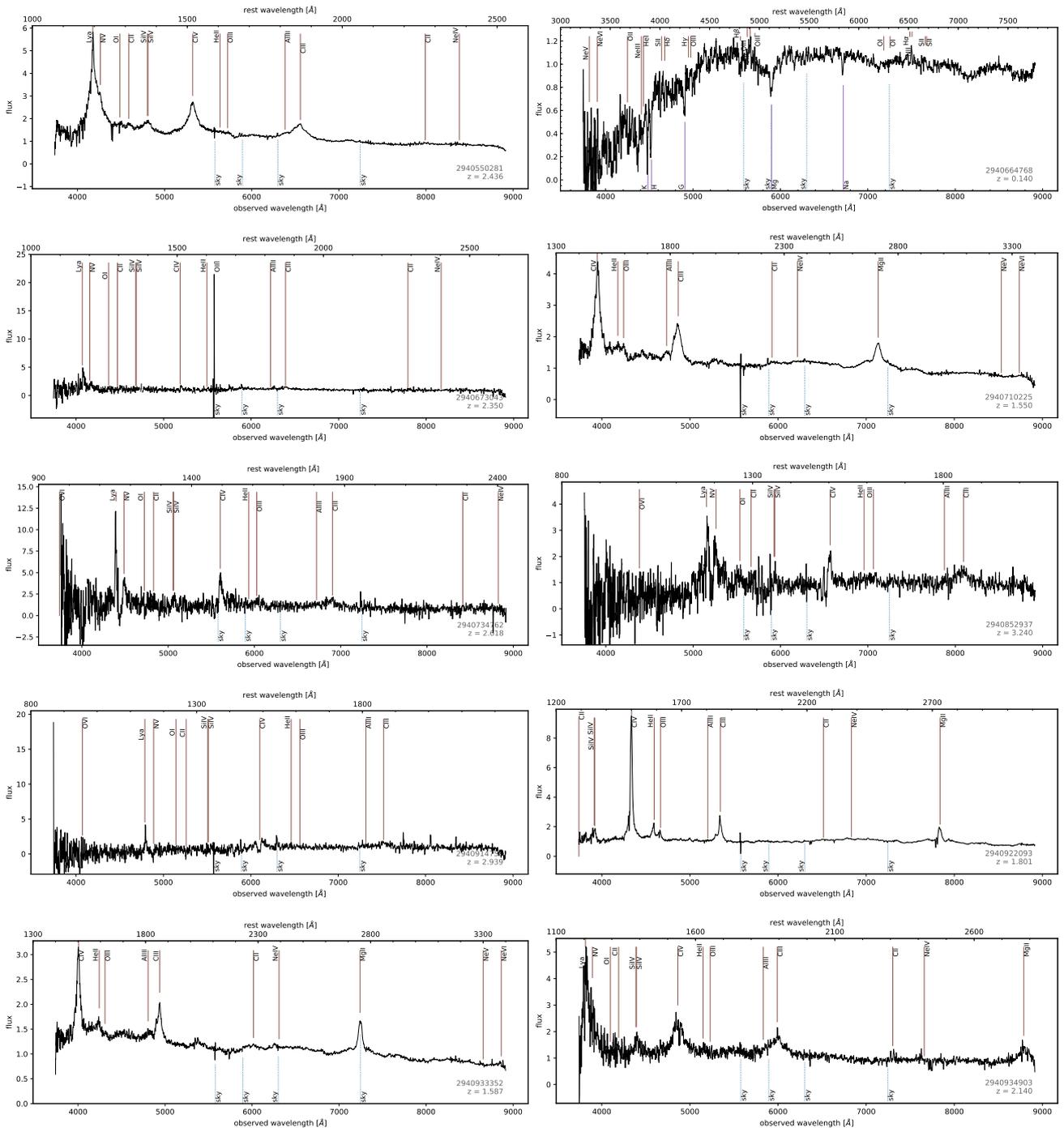


Figure C.6: Figure C.5 continued: 11–20 of 38 sources.

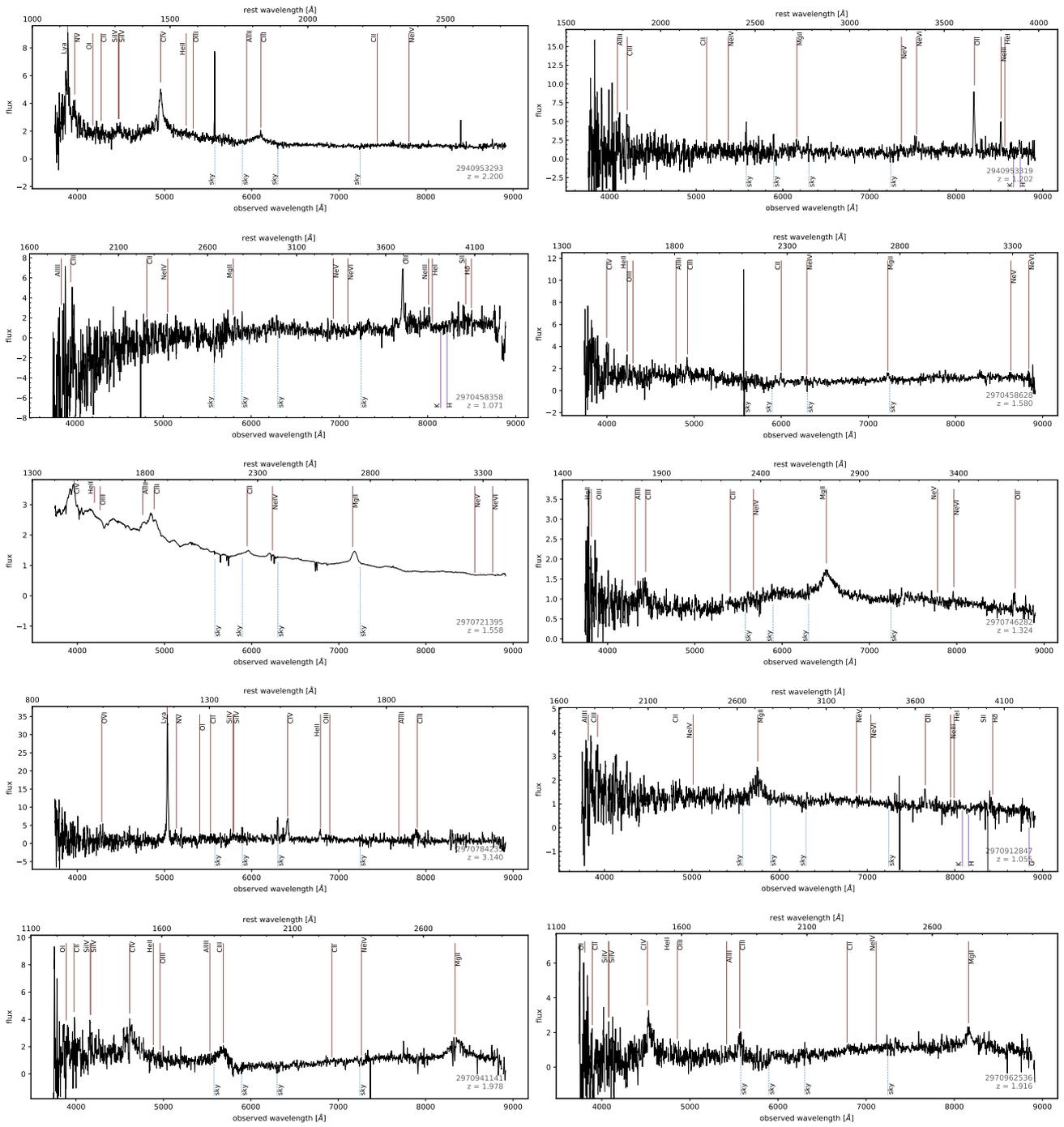


Figure C.7: Figure C.5 continued: 21–30 of 38 sources.

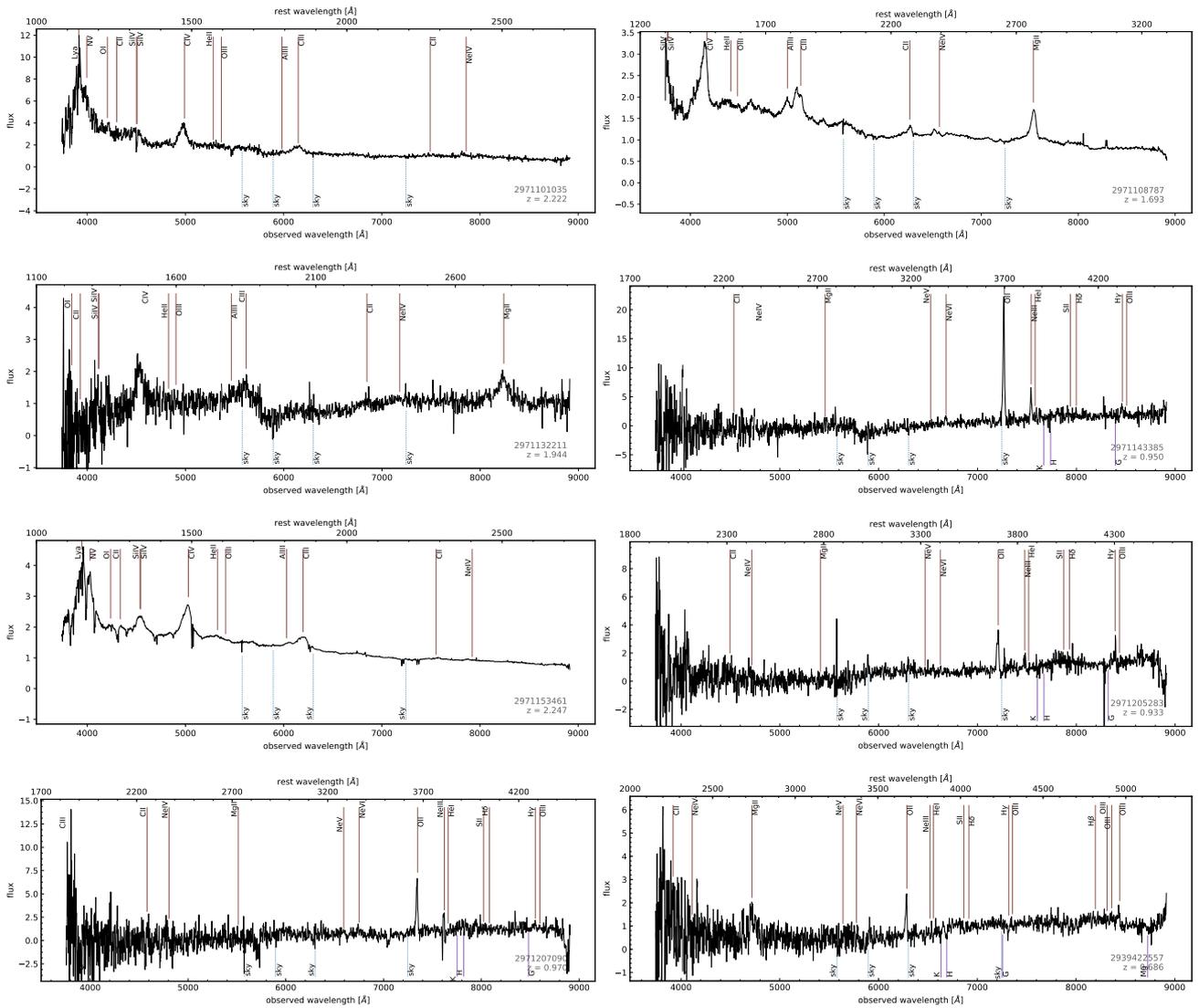


Figure C.8: Figure C.5 continued: 21–30 of 38 sources.

# References

- Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ArXiv e-prints, arXiv:1801.03181
- Annuar, A., Alexander, D. M., Gandhi, P., et al. 2017, *Astrophysical Journal*, 836, 165
- Antonucci, R. 1993, *Annual Review of Astronomy and Astrophysics*, 31, 473
- Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, *Astrophysical Journal*, 772, 26
- Assef, R. J., Eisenhardt, P. R. M., Stern, D., et al. 2015, *Astrophysical Journal*, 804, 27
- Babbedge, T. S. R., Rowan-Robinson, M., Vaccari, M., et al. 2006, *Monthly Notices of the RAS*, 370, 1159
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *Publications of the ASP*, 93, 5
- Banfield, J. K., Wong, O. I., Willett, K. W., et al. 2015, *Monthly Notices of the RAS*, 453, 2326
- Beckmann, V., & Shrader, C. R. 2012, *Active Galactic Nuclei*
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1984, *Reviews of Modern Physics*, 56, 255
- Benson, A. J., Bower, R. G., Frenk, C. S., et al. 2003, *Astrophysical Journal*, 599, 38
- Berriman, G. B., Good, J. C., Rusholme, B., & Robitaille, T. 2016, ArXiv e-prints, arXiv:1608.02649
- Best, P. N., & Heckman, T. M. 2012, *Monthly Notices of the RAS*, 421, 1569
- Best, P. N., Kaiser, C. R., Heckman, T. M., & Kauffmann, G. 2006, *Monthly Notices of the RAS*, 368, L67
- Binney, J., & Tabor, G. 1995, *Monthly Notices of the RAS*, 276, 663
- Bolatto, A. D., Warren, S. R., Leroy, A. K., et al. 2013, *Nature*, 499, 450
- Bordoloi, R., Fox, A. J., Lockman, F. J., et al. 2017, *Astrophysical Journal*, 834, 191

- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *Astrophysical Journal*, 803, 34
- Bower, G. C., Wright, M. C. H., Backer, D. C., & Falcke, H. 1999, *Astrophysical Journal*, 527, 851
- Brandl, B. R., Bernard-Salas, J., Spoon, H. W. W., et al. 2006, *Astrophysical Journal*, 653, 1129
- Calabretta, M. R., & Greisen, E. W. 2002, *Astronomy and Astrophysics*, 395, 1077
- Cappellari, M. 2017, *Monthly Notices of the RAS*, 466, 798
- Cappellari, M., & Emsellem, E. 2004, *Publications of the ASP*, 116, 138
- Cara, M., & Lister, M. L. 2008, *Astrophysical Journal*, 686, 148
- Cheung, C. C. 2004, PhD thesis, BRANDEIS UNIVERSITY
- Childress, M. J., Lidman, C., Davis, T. M., et al. 2017, *Monthly Notices of the RAS*, 472, 273
- Ching, J. H. Y., Sadler, E. M., Croom, S. M., et al. 2017, *Monthly Notices of the RAS*, 464, 1306
- Clewley, L., & Jarvis, M. J. 2004, *Monthly Notices of the RAS*, 352, 909
- Condon, J. J., Cotton, W. D., & Broderick, J. J. 2002, *Astronomical Journal*, 124, 675
- Croom, S. M., Smith, R. J., Boyle, B. J., et al. 2004, *Monthly Notices of the RAS*, 349, 1397
- Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *Monthly Notices of the RAS*, 365, 11
- da Cunha, E., Charmandaris, V., Díaz-Santos, T., et al. 2010, *Astronomy and Astrophysics*, 523, A78
- Daddi, E., Cimatti, A., Renzini, A., et al. 2004, *Astrophysical Journal*, 617, 746
- Davis, T. A., Krajnović, D., McDermid, R. M., et al. 2012, *Monthly Notices of the RAS*, 426, 1574
- Di Matteo, P., Combes, F., Melchior, A.-L., & Semelin, B. 2007, *Astronomy and Astrophysics*, 468, 61
- Di Matteo, T., Springel, V., & Hernquist, L. 2005, *Nature*, 433, 604
- Diehl, H. T., Neilsen, E., Gruendl, R., et al. 2016, in *Proceedings of the SPIE*, Vol. 9910, *Observatory Operations: Strategies, Processes, and Systems VI*, 99101D
- Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al. 2008, *Nature*, 455, 78
- Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, *Astrophysical Journal*, 748, 142
- Donoso, E., Best, P. N., & Kauffmann, G. 2009, *Monthly Notices of the RAS*, 392, 617

- Drlica-Wagner, A., Sevilla-Noarbe, I., Rykoff, E. S., et al. 2018, *Astrophysical Journal*, Supplement Series, 235, 33
- Dunlop, J. S., & Peacock, J. A. 1990, *Monthly Notices of the RAS*, 247, 19
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *Astrophysical Journal*, 633, 560
- Evans, I. N., Primini, F. A., Glotfelty, C. S., et al. 2014, *VizieR Online Data Catalog*, 9045
- Fabian, A. C. 1999, *Monthly Notices of the RAS*, 308, L39
- Fabian, A. C., Celotti, A., & Erlund, M. C. 2006, *Monthly Notices of the RAS*, 373, L16
- Fan, D., Budavári, T., Norris, R. P., & Hopkins, A. M. 2015, *Monthly Notices of the RAS*, 451, 1299
- Ferrarese, L., & Merritt, D. 2000, *Astrophysical Journal*, Letters to the Editor, 539, L9
- Fiore, F., Grazian, A., Santini, P., et al. 2008, *Astrophysical Journal*, 672, 94
- Fiore, F., Puccetti, S., Brusa, M., et al. 2009, *Astrophysical Journal*, 693, 447
- Flaugher, B. 2005, *International Journal of Modern Physics A*, 20, 3121
- Franx, M., Labbé, I., Rudnick, G., et al. 2003, *Astrophysical Journal*, Letters to the Editor, 587, L79
- Franzen, T. M. O., Banfield, J. K., Hales, C. A., et al. 2015, *Monthly Notices of the RAS*, 453, 4020
- Freedman, D., & Diaconis, P. 1981, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57, 453
- Gaspari, M., Ruszkowski, M., & Oh, S. P. 2013, *Monthly Notices of the RAS*, 432, 3401
- Gillessen, S., Eisenhauer, F., Trippe, S., et al. 2009, *Astrophysical Journal*, 692, 1075
- Goulding, A. D., & Alexander, D. M. 2009, *Monthly Notices of the RAS*, 398, 1165
- Greenstein, J. L., & Schmidt, M. 1964, *Astrophysical Journal*, 140, 1
- Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, *Monthly Notices of the RAS*, 436, 2875
- Gürkan, G., Hardcastle, M. J., & Jarvis, M. J. 2014, *Monthly Notices of the RAS*, 438, 1149
- Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2007, *Monthly Notices of the RAS*, 376, 1849

- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, *Astrophysics and Space Science*, Supplement Series, 47, 1
- Herrera Ruiz, N., Middelberg, E., Deller, A., et al. 2017, *Astronomy and Astrophysics*, 607, A132
- Hopkins, A. M., Mobasher, B., Cram, L., & Rowan-Robinson, M. 1998, *Monthly Notices of the RAS*, 296, 839
- Hopkins, P. F., Somerville, R. S., Hernquist, L., et al. 2006, *Astrophysical Journal*, 652, 864
- Hubble, E. P. 1925, *Astrophysical Journal*, 62, doi:10.1086/142943
- . 1936, *Realm of the Nebulae*
- Hunstead, R. W. 1972, *Astrophysics Letters*, 12
- Jacob, J. C., Katz, D. S., Berriman, G. B., et al. 2010, *ArXiv e-prints*, arXiv:1005.4454
- Jaffe, W. J., & Perola, G. C. 1973, *Astronomy and Astrophysics*, 26, 423
- Juneau, S., Dickinson, M., Alexander, D. M., & Salim, S. 2011, *Astrophysical Journal*, 736, 104
- Kaiser, C. R., & Best, P. N. 2007, *Monthly Notices of the RAS*, 381, 1548
- Kauffmann, G., Heckman, T. M., & Best, P. N. 2008, *Monthly Notices of the RAS*, 384, 953
- Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, *Monthly Notices of the RAS*, 346, 1055
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *Astrophysical Journal*, 556, 121
- Kewley, L. J., Maier, C., Yabe, K., et al. 2013, *Astrophysical Journal*, Letters to the Editor, 774, L10
- Knuth, K. H. 2006, *ArXiv Physics e-prints*, physics/0605197
- Kormendy, J., & Ho, L. C. 2013, *Annual Review of Astronomy and Astrophysics*, 51, 511
- Kormendy, J., & Richstone, D. 1995, *Annual Review of Astronomy and Astrophysics*, 33, 581
- Lacy, M., Petric, A. O., Sajina, A., et al. 2007, *Astronomical Journal*, 133, 186
- Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, *Astrophysical Journal*, Supplement Series, 154, 166

- Laing, R. A., Jenkins, C. R., Wall, J. V., & Unger, S. W. 1994, in *Astronomical Society of the Pacific Conference Series*, Vol. 54, *The Physics of Active Galaxies*, ed. G. V. Bicknell, M. A. Dopita, & P. J. Quinn, 201
- Lamareille, F. 2010, *Astronomy and Astrophysics*, 509, A53
- Levin, Y., & Beloborodov, A. M. 2003, *Astrophysical Journal, Letters to the Editor*, 590, L33
- Lewis, I. J., Cannon, R. D., Taylor, K., et al. 2002, *Monthly Notices of the RAS*, 333, 279
- Lípari, S., Tsvetanov, Z., & Macchetto, F. 1997, *Astrophysical Journal, Supplement Series*, 111, 369
- Longair, M. S. 1966, *Monthly Notices of the RAS*, 133, 421
- Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, *Publications of the ASP*, 115, 897
- Lukic, V., Brüggén, M., Banfield, J. K., et al. 2018, *Monthly Notices of the RAS*, 476, 246
- Machalski, J., & Godłowski, W. 2000, *Astronomy and Astrophysics*, 360, 463
- Machalski, J., Jamrozny, M., & Konar, C. 2010, *Astronomy and Astrophysics*, 510, A84
- Magnelli, B., Popesso, P., Berta, S., et al. 2013, *Astronomy and Astrophysics*, 553, A132
- Mao, M. Y., Sharp, R., Norris, R. P., et al. 2012, *Monthly Notices of the RAS*, 426, 3334
- Marocco, J., Hache, E., & Lamareille, F. 2011, *Astronomy and Astrophysics*, 531, A71
- Mauch, T., & Sadler, E. M. 2007, *Monthly Notices of the RAS*, 375, 931
- Mauduit, J.-C., Lacy, M., Farrah, D., et al. 2012, *Publications of the ASP*, 124, 714
- McAlpine, K., Jarvis, M. J., & Bonfield, D. G. 2013, *Monthly Notices of the RAS*, 436, 1084
- Menéndez-Delmestre, K., Blain, A. W., Smail, I., et al. 2009, *Astrophysical Journal*, 699, 667
- Meyer, L., Schödel, R., Eckart, A., et al. 2006, *Astronomy and Astrophysics*, 458, L25
- Middelberg, E., Norris, R. P., Cornwell, T. J., et al. 2008, *Astronomical Journal*, 135, 1276
- Middelberg, E., Deller, A. T., Norris, R. P., et al. 2013, *Astronomy and Astrophysics*, 551, A97
- Miller, N. A., Bonzini, M., Fomalont, E. B., et al. 2013, *Astrophysical Journal, Supplement Series*, 205, 13
- Miyoshi, M., Moran, J., Herrnstein, J., et al. 1995, *Nature*, 373, 127

- Morganti, R., Tadhunter, C. N., & Oosterloo, T. A. 2005, *Astronomy and Astrophysics*, 444, L9
- Müller, S. A. H., Haas, M., Siebenmorgen, R., et al. 2004, *Astronomy and Astrophysics*, 426, L29
- Mushotzky, R. F. 1984, *Advances in Space Research*, 3, 157
- Nagashima, M., Lacey, C. G., Okamoto, T., et al. 2005, *Monthly Notices of the RAS*, 363, L31
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *Astrophysical Journal*, 462, 563
- . 1997, *Astrophysical Journal*, 490, 493
- Norris, R. P. 2017, *Nature Astronomy*, 1, 671
- Norris, R. P., Lenc, E., Roy, A. L., & Spoon, H. 2012, *Monthly Notices of the RAS*, 422, 1453
- Norris, R. P., Middelberg, E., & Boyle, B. J. 2007, in *Astronomical Society of the Pacific Conference Series*, Vol. 380, *Deepest Astronomical Surveys*, ed. J. Afonso, H. C. Ferguson, B. Mobasher, & R. Norris, 229
- Norris, R. P., Afonso, J., Appleton, P. N., et al. 2006, *Astronomical Journal*, 132, 2409
- Novak, M., Smolčić, V., Delhaize, J., et al. 2017, *Astronomy and Astrophysics*, 602, A5
- Nyland, K., Lacy, M., Sajina, A., et al. 2017, *VizieR Online Data Catalog*, 223
- O'Dea, C. P. 1998, *Publications of the ASP*, 110, 493
- Padovani, P., Bonzini, M., Kellermann, K. I., et al. 2015, *Monthly Notices of the RAS*, 452, 1263
- Padovani, P., Mainieri, V., Tozzi, P., et al. 2009, *Astrophysical Journal*, 694, 235
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *Astrophysical Journal*, 517, 565
- Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2014a, *Astronomy and Astrophysics*, 571, A11
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014b, *Astronomy and Astrophysics*, 571, A12
- . 2014c, *Astronomy and Astrophysics*, 571, A16
- Pracy, M. B., Ching, J. H. Y., Sadler, E. M., et al. 2016, *Monthly Notices of the RAS*, 460, 2
- Puccetti, S., Comastri, A., Fiore, F., et al. 2014, *Astrophysical Journal*, 793, 26

- Rees, M. J. 1971, *Nature*, 229, 312
- Reid, M. J., & Brunthaler, A. 2004, *Astrophysical Journal*, 616, 872
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *Astronomical Journal*, 116, 1009
- Robertson, B., Yoshida, N., Springel, V., & Hernquist, L. 2004, *Astrophysical Journal*, 606, 32
- Rowan-Robinson, M., Babbedge, T., Oliver, S., et al. 2008, *Monthly Notices of the RAS*, 386, 697
- Sadler, E. M., Jackson, C. A., Cannon, R. D., et al. 2002, *Monthly Notices of the RAS*, 329, 227
- Sadler, E. M., Cannon, R. D., Mauch, T., et al. 2007, *Monthly Notices of the RAS*, 381, 211
- Sajina, A., Lacy, M., & Scott, D. 2005, *Astrophysical Journal*, 621, 256
- Salpeter, E. E. 1964, *Astrophysical Journal*, 140, 796
- Sanders, D. B., Soifer, B. T., Elias, J. H., Neugebauer, G., & Matthews, K. 1988, *Astrophysical Journal*, Letters to the Editor, 328, L35
- Scargle, J. D., Norris, J. P., Jackson, B., & Chiang, J. 2013, *Astrophysical Journal*, 764, 167
- Schawinski, K., Thomas, D., Sarzi, M., et al. 2007, *Monthly Notices of the RAS*, 382, 1415
- Schmidt, M. 1968, *Astrophysical Journal*, 151, 393
- Schödel, R., Ott, T., Genzel, R., et al. 2002, *Nature*, 419, 694
- Scott, D. W. 1979, *Biometrika*, 66, 605
- Seymour, N., Huynh, M., Dwelly, T., et al. 2009, *Monthly Notices of the RAS*, 398, 1573
- Shabala, S., & Alexander, P. 2009, *Astrophysical Journal*, 699, 525
- Shabala, S. S., Kaviraj, S., & Silk, J. 2011, *Monthly Notices of the RAS*, 413, 2815
- Sijacki, D., Springel, V., Di Matteo, T., & Hernquist, L. 2007, *Monthly Notices of the RAS*, 380, 877
- Silk, J., & Rees, M. J. 1998, *Astronomy and Astrophysics*, 331, L1
- Smith, E. P., & Heckman, T. M. 1989, *Astrophysical Journal*, 341, 658
- Smith, G. A., Saunders, W., Bridges, T., et al. 2004, in *Proceedings of the SPIE*, Vol. 5492, Ground-based Instrumentation for Astronomy, ed. A. F. M. Moorwood & M. Iye, 410–420

- Smith, H. J., & Hoeffleit, D. 1963, *Astronomical Journal*, 68, 292
- Smolčić, V., Zamorani, G., Schinnerer, E., et al. 2009, *Astrophysical Journal*, 696, 24
- Smolčić, V., Delvecchio, I., Zamorani, G., et al. 2017, *Astronomy and Astrophysics*, 602, A2
- Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, *Astrophysical Journal*, Supplement Series, 148, 175
- Spoon, H. W. W., Keane, J. V., Tielens, A. G. G. M., Lutz, D., & Moorwood, A. F. M. 2001, *Astronomy and Astrophysics*, 365, L353
- Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, *Monthly Notices of the RAS*, 371, 972
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, *Astrophysical Journal*, 592, 728
- Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, *Astrophysical Journal*, 631, 163
- STScI Development Team. 2013, pynphot: Synthetic photometry software package, *Astrophysics Source Code Library*, ascl:1303.023
- Supercosmos Sky, S. 2007, *VizieR Online Data Catalog*, 6124
- Surace, & et al. 2012, *VizieR Online Data Catalog*, 2302
- Surace, J. A., Shupe, D. L., Fang, F., et al. 2005, in *Bulletin of the American Astronomical Society*, Vol. 37, American Astronomical Society Meeting Abstracts, 1246
- Tadhunter, C. 2016, *Astronomy and Astrophysics Reviews*, 24, 10
- Taylor, M. 2013, *Starlink User Note*, 253
- Trouille, L., & Barger, A. 2010, in *IAU Symposium*, Vol. 267, *Co-Evolution of Central Black Holes and Galaxies*, ed. B. M. Peterson, R. S. Somerville, & T. Storchi-Bergmann, 142–142
- Trouille, L., Barger, A. J., Cowie, L. L., Yang, Y., & Mushotzky, R. F. 2008, *Astrophysical Journal*, Supplement Series, 179, 1
- . 2009, *Astrophysical Journal*, 703, 2160
- Trouille, L., Barger, A. J., & Tremonti, C. 2011, *Astrophysical Journal*, 742, 46
- Urry, C. M., & Padovani, P. 1995, *Publications of the ASP*, 107, 803
- Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, *Monthly Notices of the RAS*, 404, 1639

- Walter, F., Weiss, A., & Scoville, N. 2002, *Astrophysical Journal*, Letters to the Editor, 580, L21
- Watson, D., Christensen, L., Knudsen, K. K., et al. 2015, *Nature*, 519, 327
- Whysong, D., & Antonucci, R. 2004, *Astrophysical Journal*, 602, 116
- Williams, W. L., Calistro Rivera, G., Best, P. N., et al. 2018, *Monthly Notices of the RAS*, 475, 3429
- Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M., & Eales, S. A. 2001, *Monthly Notices of the RAS*, 322, 536
- Wolf, C., Aragón-Salamanca, A., Balogh, M., et al. 2009, *Monthly Notices of the RAS*, 393, 1302
- Wong, O. I., Koss, M. J., Schawinski, K., et al. 2016, *Monthly Notices of the RAS*, 460, 1588
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *Astronomical Journal*, 140, 1868
- Wyithe, J. S. B. 2006, *Monthly Notices of the RAS*, 365, 1082
- Yan, H., Yan, L., Zamojski, M. A., et al. 2011a, *Astrophysical Journal*, Letters to the Editor, 728, L22
- Yan, R., Ho, L. C., Newman, J. A., et al. 2011b, *Astrophysical Journal*, 728, 38
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *Astronomical Journal*, 120, 1579
- Yuan, F., Lidman, C., Davis, T. M., et al. 2015, *Monthly Notices of the RAS*, 452, 3047
- Zel'dovich, Y. B., & Novikov, I. D. 1964, *Soviet Physics Doklady*, 9, 246
- Zhang, K., & Hao, L. 2018, *Astrophysical Journal*, 856, 171
- Znajek, R. L. 1977, *Monthly Notices of the RAS*, 179, 457

it is not the critic who counts; not the man who points out how the strong man stumbles, or where the doer of deeds could have done them better. the credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs, who comes short again and again, because there is no effort without error and shortcoming; but who does actually strive to do the deeds; who knows great enthusiasms, the great devotions; who spends himself in a worthy cause; who at the best knows in the end the triumph of high achievement, and who at the worst, if he fails, at least fails while daring greatly, so that his place shall never be with those cold and timid souls who neither know victory nor defeat.

TR

THIS PAGE INTENTIONALLY LEFT BLANK

**T**HIS THESIS WAS TYPESET using L<sup>A</sup>T<sub>E</sub>X: originally developed by Leslie Lamport and based on Donald Knuth's T<sub>E</sub>X. The template for this thesis borrowed heavily from a template released under the permissive AGPL license by its lead author, Jordan Suchow; that template can be found online at [github.com/suchow/Dissertate](https://github.com/suchow/Dissertate). This modified template was altered to suit the expected structure of a thesis for The University of Tasmania, as well as to suit my taste. If you would like a template for this thesis or any of the Python code used within, I am contactable at [jesseaswan@gmail.com](mailto:jesseaswan@gmail.com).

