TRACING LARGE-SCALE STRUCTURE WITH RADIO SOURCES

by

Samuel Nathan Lindsay

A THESIS SUBMITTED TO THE UNIVERSITY OF HERTFORDSHIRE
IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

February 2015

Typeset in LaTeX 2_ε.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Samuel Nathan Lindsay

Acknowledgements

Firstly, this thesis is a credit to the hard work of my parents, Roger and June, for far longer than the 4 years it took to complete. They have gone without, so I could have the excellent education I have been given for the last two decades, and I may never be able to repay that kindness.

At Haberdashers' Aske's Boys' School, no teacher of mine has better served the school's motto of *"Nurturing Excellence"* than Richard Kingdon, who taught me in 5 of his first 6 years at the school. His bad luck was my good fortune, and undoubtedly a factor in deciding to study Physics, and earning a place at my first-choice university. I am grateful to Imperial College London for providing the world-class learning environment that ultimately compelled me to pursue a PhD, and particularly to Professors Carlo Contaldi, João Magueijo and Andrew Jaffe – the project supervisor, examiner and lecturer, respectively, who helped set my sights on the structure of the Universe, rather than of atoms. This is also where I met my best friend Hannah, who has supported and vicariously lived through this PhD with me, reminding me why it has been worth the hard work.

Thank you to the friends I have shared an office with at UH: Jose, James, Eleni, Hanifa, Sam, Chris, Mahesh, Phil, and honorary office mate, Niall. Sometimes you were all that got me out of bed in the morning and helped me get through the days. Special appreciation is reserved for Phil, who became my comedy partner, sounding board and, latterly, proofreader over the final 2 years of my PhD, and beyond.

Rowing has kept me busy while not in the office; my first 9 months were spent rowing for Imperial College and University of Hertfordshire. After sensibly retiring, I later returned to a warm welcome offered by Vesta Rowing Club in September 2012. I owe all involved a great debt of gratitude, for accepting me, allowing me some escape from the PhD to be a part of a team, and reminding me what it feels like to be good at something (and to enjoy the hard work along the way).

Finally, this thesis wouldn't be here without the opportunity, faith, encouragement and support of my supervisor. Professor Matt Jarvis has been, and continues to be, my ideal supervisor. His attention, while hard to come by, has always been directed at helping me to become a well-rounded scientist for the future. After countless meetings where I entered the room with my head down about my work, I have always left feeling rejuvenated by the consideration, support and encouragement Matt has provided. Professionally, and personally, I couldn't have wished for better.

Abstract

In this thesis, I investigate the spatial distribution of radio sources, and quantify their clustering strength over a range of redshifts, up to $z \sim 2.2$, using various forms of the correlation function measured with data from several multi-wavelength surveys.

I present the optical spectra of 30 radio AGN ($S_{1.4} > 100$ mJy) in the GAMA/H-ATLAS fields, for which emission line redshifts could be deduced, from observations of 79 target sources with the EFOSC2 spectrograph on the NTT. The mean redshift of these sources is $z = 1.2$; 12 were identified as quasars (40 per cent), and 6 redshifts (out of 24 targets) were found for AGN hosts to multiple radio components. While obtaining spectra for hosts of these multi-component sources is possible, their lower success rate highlights the difficulty in acheiving a redshift-complete radio sample.

Taking an existing spectroscopic redshift survey (GAMA) and radio sources from the FIRST survey ($S_{1.4} > 1$ mJy), I then present a cross-matched radio sample with 1,635 spectroscopic redshifts with a median value of $z = 0.34$. The spatial correlation function of this sample is used to find the redshift-space (s_0) and real-space correlation lengths ($r_0 \sim 8.2 \ h^{-1}$Mpc), and a mass bias of ~ 1.9. Insight into the redshift-dependence of these quantities is gained by using the angular correlation function and Limber inversion to measure the same spatial clustering parameters. Photometric redshifts from SDSS/UKIDSS are incorporated to produce a larger matched radio sample at $z \simeq 0.48$ (and low- and high-redshift subsamples at $z \simeq 0.30$ and $z \simeq 0.65$), while their redshift distribution is subtracted from that taken from the SKADS radio simulations to estimate the redshift distribution of the remaining unmatched sources ($z \simeq 1.55$). The observed bias evolution over this redshift range is compared with model predictions based on the SKADS simulations, with good agreement at low redshift. The bias found at high redshift significantly exceeds these predictions, however, suggesting a more massive population of galaxies than expected, either due to the relative proportions of different radio sources, or a greater typical halo mass for the high-redshift sources.

Finally, the reliance on a model redshift distribution to reach to higher redshifts is removed, as the angular cross-correlation function is used with deep VLA data ($S_{1.4} > 90 \ \mu$Jy) and optical/IR data from VIDEO/CFHTLS ($K_\mathrm{s} < 23.5$) over 1 square degree. With high-quality photometric redshifts up to $z \sim 4$, and a high signal-to-noise clustering measurement (due to the $\sim 100{,}000 \ K_\mathrm{s}$-selected galaxies), I am able to find the bias of a matched sample of only 766 radio sources (as well as of

the VIDEO sources), divided into 4 redshift bins reaching a median bias at $z \simeq 2.15$. Again, at high redshift, the measured bias appears to exceed the prediction made from the SKADS simulations. Applying luminosity cuts to the radio sample at $L > 10^{23}$ W Hz^{-1} and higher (removing any non-AGN sources), I find a bias of 8–10 at $z \sim 1.5$, considerably higher than for the full sample, and consistent with the more numerous FRI AGN having similar mass to the FRIIs ($M \sim 10^{14} \, M_{\odot}$), contrary to the assumptions made in the SKADS simulations. Applying this adjustment to the model bias produces a better fit to the observations for the FIRST radio sources cross-matched with GAMA/SDSS/UKIDSS, as well as for the high-redshift radio sources in VIDEO.

Therefore, I have shown that we require a more robust model of the evolution of AGN, and their relation to the underlying dark matter distribution. In particular, understanding these quantities for the abundant FRI population is crucial if we are to use such sources to probe the cosmological model as has been suggested by a number of authors (e.g. Raccanelli et al., 2012; Camera et al., 2012; Ferramacho et al., 2014)

Contents

List of Figures

List of Tables

1

Introduction

"There will come a time when our descendants will be amazed that we did not know things that are so plain to them . . . Many discoveries are reserved for ages still to come, when memory of us will have been effaced. Our universe is a sorry little affair unless it has in it something for every age to investigate . . . Nature does not reveal her mysteries once and for all."

<div align="right">Lucius Annaeus Seneca</div>

The topics discussed in this thesis and the intended applications of this work are concerned with the structure of the Universe on the very largest scales. For the majority of human history, this could be said for astronomy in general; the stars and other heavenly bodies represented the greatest distances conceivable, beyond the scale of our own Earth. However, the revolutionary progress made in astronomy (particularly in the last century) has opened our eyes to the true scale of the observable Universe, and the importance of studying its structure and dynamics as a whole, as well as the individual objects within it. Of particular relevance to my work is the discovery of extragalactic radio sources and the new science that has been opened up by observing in this part of the electromagnetic spectrum. This literal and figurative broadening of horizons has occurred in tandem with a similarly rapid advance in fundamental particle physics reaching far deeper at the subatomic level, and just as 21st century particle physicists have the Large Hadron Collider, the cosmology and astrophysics community looks forward

to its own epic new experiment at radio wavelengths with the forthcoming Square Kilometre Array. The unprecedented combination of depth and sheer scale of the coming radio surveys stands to benefit observational cosmology in a variety of ways, the early stages of which have been the subject of the research presented here.

1.1   Cosmology

1.1.1   The Extragalactic Universe

Astronomy has long struggled against the human urge to be considered special, and at the centre of everything. The geocentric models of Ptolemy survived until the 16th century, when Copernicus put forward a heliocentric model where the Earth orbited the Sun, rather than vice versa. The work of Kepler and Newton allowed this model to be refined and eventually brought about the consensus that the 'fixed' stars were like our Sun, but evenly distributed through infinite space. Not until the 20th century would it be known that the Milky Way was just one of many galaxies like it, and that we do not occupy a special location in our galaxy, or in our model of the Universe – this is known as the Copernican principle.

In 1771, Charles Messier, while looking for comets, was frustrated by several 'fuzzy' sources (rather than point-like stars) which were dubbed "nebulae". He catalogued 103 of these sources, numbered M1 to M103 (although several more were added to the Messier catalogue later). The combined work of the Herschels (William, sister Caroline, and son John) and John Dreyer brought the New General Catalogue (NGC) of 7,840 such objects in 1888, which Dreyer extended with the Index Catalogues (IC), in 1895 and 1908, containing a total of 5,386 nebulae. A clue as to the nature of these sources was found by Vesto Slipher in the Doppler-shifted spectra of spiral nebulae. Slipher found that the Andromeda nebula (M31) was approaching the solar system with a velocity of 300 $km\,s^{-1}$ (Slipher, 1914), and began building a body of evidence for nebulae both receding with great speed and sometimes approaching.

Speculation as to the significance of M31 (and other spiral nebulae) continued into the 1920s, with the Great Debate between Harlow Shapley and Heber Curtis in 1920. Curtis had reasoned that spiral nebulae were some 100 times further away than galactic nebulae, assuming the same intrinsic brightness (Curtis, 1917), and was therefore arguing in favour of the "Island Universe" hypothesis, while Shapley contested that the Milky Way was the full extent of the Universe. While Curtis had several valid lines of argument, and history has proved his proposition correct, Shapley was deemed to have won
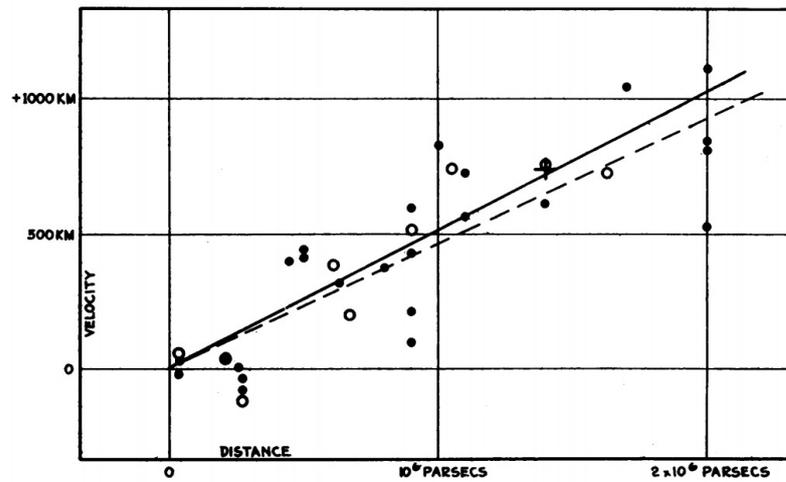
FIGURE 1

Velocity-Distance Relation among Extra-Galactic Nebulae.

FIGURE 1.1: The original velocity–distance plot of Hubble (1929) for 24 "extra-galactic nebulae".

the debate. Within a few years, however, the earlier work of Leavitt & Pickering (1912) on the period–luminosity relation in Cepheid variable stars in the Small Magellanic Cloud had been taken advantage of by Edwin Hubble, who used these variable stars in M31 and M33 to estimate their intrinsic luminosity, and therefore a distance for both nebulae of 285 kpc (Hubble, 1925). While both galaxies are now known to considerably more distant than these estimates, Hubble had nevertheless shown these nebulae to be beyond our galaxy.

1.1.2 Cosmic Expansion

Aware of the work of Slipher in the previous decade, Hubble set about carefully measuring the distance and apparent Doppler shift of as many galaxies as possible. This culminated in the seminal paper, from which we derive Hubble's law (Hubble, 1929):

$$v = H_0 d, \tag{1.1}$$

where the recession velocity, v, of a galaxy is proportional to the distance, d, to that galaxy, with the constant H_0 being known as the Hubble parameter at the present day (see Figure 1.1). This definition allows us to express the more general Hubble parameter as a function of scale factor, $a(t)$: $H = \dot{a}/a$. The law also gives us the motivation for defining redshift, z, which quantifies the stretching of light waves due to the expansion of space (rather than Doppler shift):

$$1 + z = \frac{\lambda_{\text{observed}}}{\lambda_{\text{emitted}}} = \frac{1}{a}, \tag{1.2}$$

where redshift is also related to the recession velocity by $z \simeq v/c$, for $v \ll c$.

The observation that galaxies tend to be moving away from us appears, at first glance, to violate the Copernican principle but, when put in the context of an expanding universe, the linearity of Hubble's law makes it an accurate description of observations from any point in the Universe. In fact, Hubble's value of $H_0 = 530 \, \mathrm{km \, s^{-1} Mpc^{-1}}$ can be inverted to give an estimate of the age of the Universe (~ 2 Gyr), assuming a constant rate of expansion. This observation refuted the previous assumptions of a static universe, which had already been ruled out by general relativity, without Einstein adding a cosmological constant to force the issue (see Section 1.1.3).

Hubble's expanding universe continued to split opinions in cosmology for decades, two possible explanations being cited:

Steady-state theory Hoyle (1948) revised the static universe previously adopted by Einstein and his predecessors given the observed expansion, postulating that new matter is created continuously, filling the space created by expansion, and resulting in an unchanging appearance over time.

"Big Bang" theory Georges Lemaître had proposed an expanding model of the Universe and predicted Hubble's law (Lemaître, 1927), independently deriving the Friedmann equations (see Section 1.1.3) in the process. He later invoked a "primeval atom", to describe an initial dense state of the Universe, before expansion, at the point of a "Big Bang" (a term coined by Hoyle to distinguish the model from his own).

Whilst initially poorly received by cosmologists at the time, the Big Bang theory was advocated by George Gamow, and it was his and subsequent work that predicted that an infinitely hot, dense Big Bang universe would initially be dominated by radiation, and that the first of this radiation to escape the bounds of the dense matter would remain today, highly redshifted into the microwave part of the spectrum (Gamow, 1948; Alpher & Herman, 1948). This cosmic microwave background (CMB) radiation was detected by Penzias and Wilson, originally thought to be noise, but later confirmed to be an isotropic black-body emission with a temperature estimated at 3.5 K (Penzias & Wilson, 1965; Dicke et al., 1965). More recent space- and balloon-based observations have measured the temperature with remarkable precision at ~ 2.73 K. This discovery, resulting in a Nobel Prize, was the decisive blow to Hoyle's steady-state cosmology, and the half-century since has been devoted to parameterizing the expansion of the Universe, and understanding the components and physical mechanisms controlling it.

1.1.3   Robertson-Walker Metric and Friedmann Equations

According to the current paradigm, gravity is described by Einstein's theory of general relativity, with the geometry of space-time determined by the energy/matter content of the Universe through the Einstein field equation:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{1.3}$$

$R_{\mu\nu}$ is the Ricci tensor, describing the local curvature of space-time, R is the curvature scalar, $g_{\mu\nu}$ is the metric, $T_{\mu\nu}$ is the energy–momentum tensor describing the matter content of the Universe, and Λ is the cosmological constant introduced by Einstein to obtain a static universe.

Given the assumptions of isotropy and homogeneity, the most general solution to the Einstein field equation is the Robertson–Walker metric, which is defined in spherical polar coordinates (r, θ, ϕ) by the scale factor $a(t)$ and the curvature k:

$$ds^2 = dt^2 - \frac{a^2(t)}{c^2}\left(\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2\right). \tag{1.4}$$

For simplicity, r is a comoving radial coordinate, such that a fundamental observer (one who is locally at rest with respect to the matter in their vicinity) has a fixed r, even with a changing scale factor. Similarly, $k \neq 0$ is held fixed at $k = \pm 1$ for convenience (r being rescaled accordingly). If $k = +1$, space is positively curved (i.e. elliptic geometry) such that parallel lines converge, corresponding to a 'closed' universe. If $k = -1$, space is negatively curved (i.e. hyperbolic geometry) such that parallel lines diverge, corresponding to an 'open' universe. A flat universe ($k = 0$) is often assumed, but also continues to be favoured by observational evidence.

Under the assumption of no peculiar velocities and a uniform ideal fluid with energy density ρc^2 and pressure p, substituting the Robertson–Walker metric into the field equation, one obtains the Friedmann equations:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3\frac{p}{c^2}\right) + \frac{\Lambda c^2}{3}, \tag{1.5}$$

from the time–time (g_{00}) component, and

$$\left(\frac{\dot{a}}{a}\right)^2 = H^2(t) = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3}, \tag{1.6}$$

which follows from substituting equation 1.5 into the solution for the space–space (g_{ij}) components. We may simplify these equations by considering the cosmological constant as part of the energy content of the Universe, with a 'mass' density $\rho_\Lambda = \Lambda c^2/8\pi G$. At this point it becomes useful to define a critical

density for a spatially flat Universe, above which earlier cosmological models without a Λ component determined that the Universe would eventually begin to contract:

$$\rho_{\mathrm{crit}}(t) = \frac{3H^2(t)}{8\pi G}. \tag{1.7}$$

Cosmic density parameters for matter, radiation and Λ, therefore, can all be described in terms of ρ_{crit}:

$$\Omega_i(t) = \frac{\rho_i(t)}{\rho_{\mathrm{crit}}(t)}. \tag{1.8}$$

The pressure, p, in equation 1.5 can be expressed in terms of ρ by invoking the equation of state w, such that $p = w\rho c^2$, and the density evolves with scale factor as $\rho_{\mathrm{m}} \propto a^{-3(1+w)}$. Matter (pressureless, $w = 0$) therefore has a density that, intuitively, falls off as the inverse cube of a. Radiation, however, with $w = 1/3$, has $\rho_{\mathrm{r}} \propto a^{-4}$, decreasing more rapidly than matter due to the extra energy loss via redshifting. Any period when radiation dominated the energy density of the Universe must eventually give way to a period of matter dominance as a continues to increase. The cosmological constant, on the other hand, has no dependence on scale factor, giving it an equation of state $w = -1$ (i.e. it exerts a negative pressure). As a result, in contrast to ρ_{m} continuing to decrease in an expanding universe, ρ_{Λ} remains constant, and a period of matter-dominance eventually gives way to a Λ-dominated (de Sitter) universe. While $a \propto t^{1/2}$ initially, the transition through matter–radiation equality shifts the evolution to $a \propto t^{2/3}$ (in the simple case where $k = 0$), but a subsequent transition to Λ-dominance causes the Universe to expand exponentially: $a \propto \exp[H_0(t - t_0)]$. At the present day, we appear to be in this transition phase from matter-domination to accelerated expansion due to a cosmological constant (e.g. Riess et al., 1998; Perlmutter et al., 1999).

1.1.4 Λ–CDM Cosmology

Today more than ever, cosmological theories abound, with the invocation of string theory or modified gravity, or dispensing with the Copernican principle, for example, giving rise to many conflicting explanations for some subset of the accumulated observations of the Universe to date. The model that remains the benchmark against which others may be compared, sometimes termed the 'standard model', is the Λ–Cold Dark Matter cosmological model. This is a parameterization of cosmology, resulting from the assumption of homogeneity and isotropy on large scales, using the Robertson–Walker metric described above. The model describes a universe with non-zero cosmological constant (Λ), and a significant contribution from an electromagnetically non-interacting matter component ('dark matter'), with sufficiently low primordial velocities to maintain the small-scale density fluctuations to seed

galaxy formation ('cold').

The effect of Λ comprising a significant fraction of the combined mass density of the Universe is an accelerated expansion, the detection of which earned a Nobel Prize for Adam Riess, Saul Perlmutter and Brian Schmidt. In two independent groups, they used Type Ia supernovae (SNe) as standard candles to extend the Hubble diagram to $z \sim 0.8$ (Riess et al., 1998; Perlmutter et al., 1999), effectively ruling out a $\Lambda = 0$ universe (see Figure 1.2).

The concept of dark matter was first proposed by Fritz Zwicky in 1933, seeking reconciliation between the orbital velocities of galaxies in clusters and the seemingly insufficient mass to maintain their orbits. This missing mass can be detected indirectly, in a number of ways: strong gravitational lensing; acoustic peaks in the CMB (see Section 1.2.5); or the dynamics of the galaxies themselves. For the latter, the virial theorem states that the kinetic energy of a galaxy or cluster is half the gravitational binding energy, with any apparent excess in velocity indicating extra mass. The existence of dark matter was confirmed when Rubin, Ford & Thonnard (1980) demonstrated the flat rotation curves of several galaxies out to large radii, where for a mass concentrated in the central bulge, the expected rotation curve falls off approximately as $r^{1/2}$ beyond this region (see Figure 1.3). The shape of the rotation curves observed may be explained by a spherical 'halo' of non-luminous matter, and density profiles have been established that agree with observations and simulations (e.g. Navarro, Frenk & White, 1997).

The Λ–CDM model reduces cosmology to six parameters, from which other important values may be derived, successfully fitting to observations. These six parameters, equivalent to the following at late times ($z = 0$), can be used in combination to infer others. Using the latest temperature and lensing maps from the *Planck* satellite, in combination with WMAP polarization maps, high-ℓ CMB observations from ACT and SPT, and baryon acoustic oscillation measurements (Planck Collaboration et al., 2013a), the following parameters are found:

Hubble constant ($h = 67.8 \pm 0.77$) The expansion rate of the Universe (equation 1.1) expressed in terms of the dimensionless quantity h: $H_0 = 100h \text{ km s}^{-1}\text{Mpc}^{-1}$.

Baryon density ($\Omega_\text{b} h^2 = 0.0221 \pm 0.0002$) The fraction of critical density of baryonic matter which, in the Big Bang nucleosynthesis (BBN) scenario, can be calculated directly from elemental abundances of the light elements (^{4}He, ^{3}He, D and ^{7}Li), which should be relatively unaffected by subsequent stellar processes.

Dark matter density ($\Omega_\text{c} h^2 = 0.119 \pm 0.002$) The fraction of critical density of non-baryonic cold dark matter ($\Omega_\text{m} = \Omega_\text{c} + \Omega_\text{b}$).
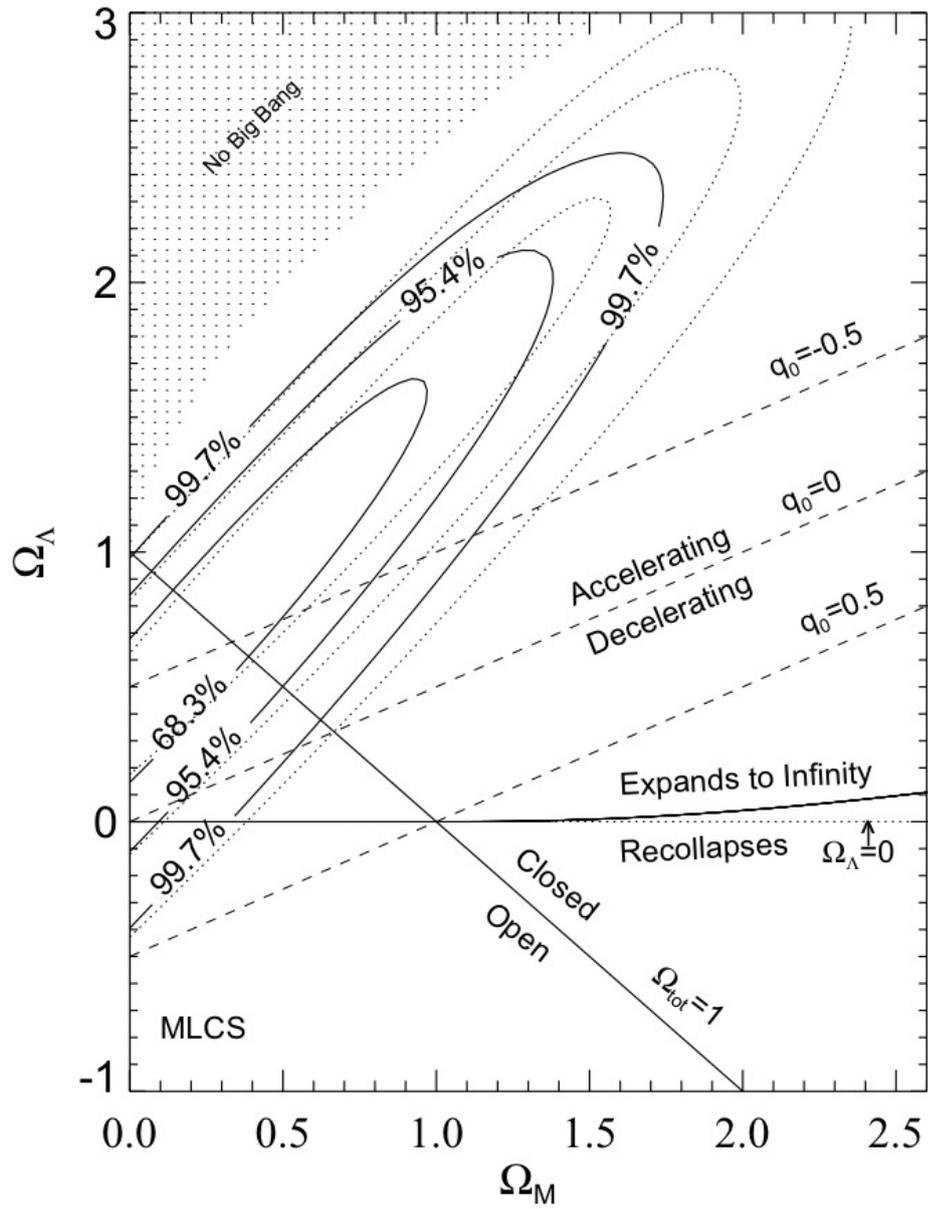
FIGURE 1.2: Joint confidence intervals for $(\Omega_m, \Omega_\Lambda)$ from SNe Ia, as calculated by Riess et al. (1998).
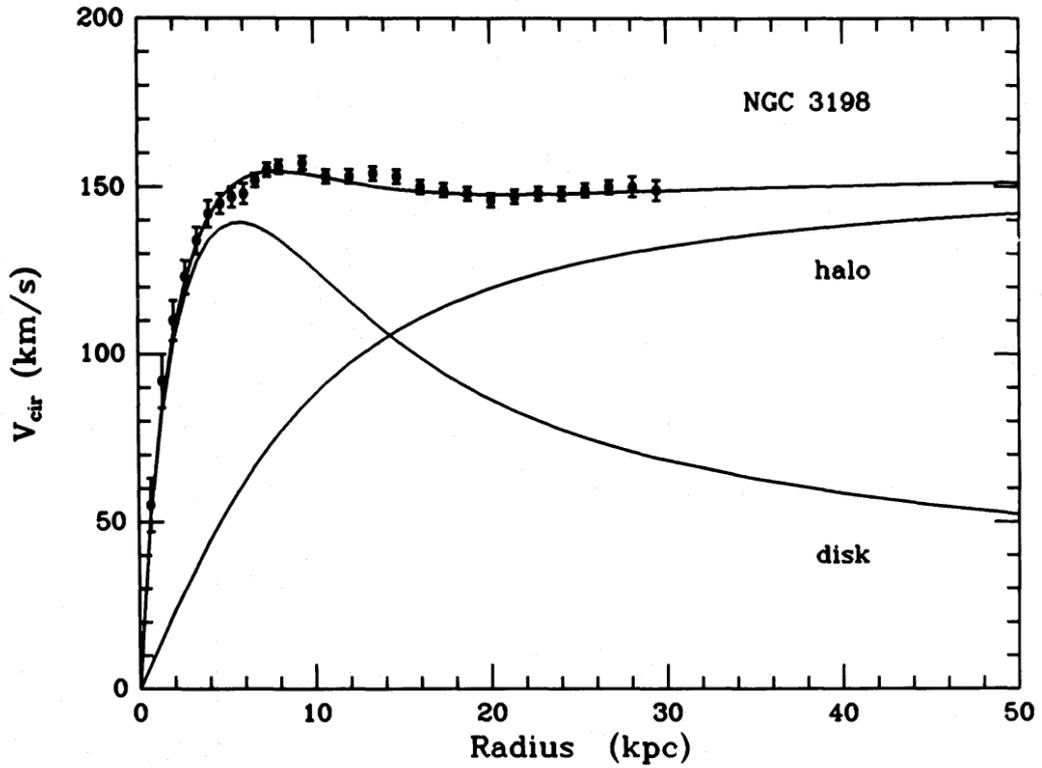
FIGURE 1.3: An observed rotation curve for NGC 3198, with line fits for model mass profiles for a disk and dark matter halo component (van Albada et al., 1985).

Cosmological constant density ($\Omega_\Lambda = 0.692 \pm 0.010$) The fraction of critical density of a cosmological constant ($\Omega_{\text{tot}} = \Omega_{\text{m}} + \Omega_\Lambda = 1 - \Omega_k$).

Fluctuation amplitude at 8 h^{-1}Mpc ($\sigma_8 = 0.862 \pm 0.012$) The normalization of the matter power spectrum (see Section 1.2.1) at 8 h^{-1}Mpc, equivalent to the rms matter density fluctuation in spheres of radius 8 h^{-1}Mpc.

Redshift of reionization ($z_{\text{re}} = 11.3 \pm 1.1$) After recombination when the hot plasma of the early Universe cooled enough to condense into neutral atomic hydrogen, some further time passed before objects began to condense from this neutral phase. The redshift of reionization describes the time at which these newly condensed objects started to emit radiation strong enough to reionize this atomic hydrogen.

There are also several common extensions to the basic Λ–CDM model, which make incremental differences to the main parameter fits for a given experiment, and increase the errors due to the added dimensions in parameter space. These extensions include: N_{eff}, the effective number of neutrino-like species; neutrino mass in eV (otherwise assumed massless); dynamical dark energy, with either a

constant $w \neq 1$, or a simply evolving function of scale factor ($w(a) = w_0 + (1 - a)w_a$).

1.2 Large-Scale Structure

The CMB provides the earliest snapshot of what the nascent Universe looked like, proving it to be remarkably homogeneous, with only very small deviations in temperature across the entire sky. These fractionally hot and cold spots formed the seeds of present day large-scale structure in the form of clusters and voids, respectively. These structures exist, first and foremost, in the dark matter distribution, but this governs the distribution of light-emitting matter in the form of the galaxies we are able to observe.

1.2.1 Direct Clustering Measurements

In order to understand and quantify the evolution of large-scale structure, we must apply ourselves to studying this galaxy distribution, and the relationship between it and the underlying dark matter distribution. There are various methods, each with their own advantages or applications, for quantifying the clustering of galaxies: cluster-finding algorithms (e.g. 'friends-of-friends'; Huchra & Geller, 1982), nearest neighbour measures (e.g. Bahcall & Soneira, 1983), and counts-in-cells (e.g. Magliocchetti et al., 1999; Blake & Wall, 2002b; Yang & Saslaw, 2011), for example. The most common and, in the context of this thesis, most useful methods for quantifying galaxy clustering are described below.

Correlation Function

A simple measure of the clustering of sources in an angular distribution is found in the angular two-point correlation function, $w(\theta)$. It is defined as the excess probability of finding a galaxy at an angular distance θ from another galaxy, as compared with a Poissonian (unclustered) distribution (Peebles, 1980):

$$\delta P = \sigma_0[1 + w(\theta)]\delta\Omega, \tag{1.9}$$

where δP is the probability, σ_0 is the mean surface density, and $\delta\Omega$ is the surface area element. A value of $w(\theta) = 0.01$, therefore, indicates a 1 per cent enhancement of source pairs separated by an angle, θ, due to the tendency for galaxies to either form in the same dark matter haloes, or cluster due to gravitational attraction, this being viewed in projection onto a two-dimensional surface.

A simple estimator is given by:

$$w(\theta) = \frac{n_r(n_r - 1)}{n_d(n_d - 1)} \frac{DD}{RR} - 1, \tag{1.10}$$

where $DD(\theta)$ and $RR(\theta)$ represent the number of galaxy pairs separated by θ in the real data and a randomly distributed catalogue, respectively. Introducing an additional count of the cross-pair separations $DR(\theta)$ allows for reduced variance, such as is found for our chosen estimator by Landy & Szalay (1993):

$$w(\theta) = \frac{n_r(n_r - 1)}{n_d(n_d - 1)} \frac{DD}{RR} - \frac{(n_r - 1)}{2n_d} \frac{DR}{RR} + 1, \tag{1.11}$$

where n_d and n_r are the number of real sources and random sources. The most common alternative estimator is that of Hamilton (1993), but there is little to separate the two in terms of results or reliability (see Kerscher, Szapudi & Szalay, 2000, for an in-depth comparison of various estimators).

Traditionally, $w(\theta)$ has been fitted by a power law (e.g. Peebles, 1980) with a slope of $\alpha \sim 0.8$ commonly found for the clustering of objects of various sizes (Bahcall & Soneira, 1983):

$$w(\theta) = A\theta^{-\alpha}. \tag{1.12}$$

However, the angular correlation function of radio galaxies, the subject of this thesis, more closely follows a double power law form (Blake & Wall 2002a; Section 1.3.5).

The 3-dimensional analogue of $w(\theta)$ is the spatial two-point correlation function $\xi(r)$, which measures the excess probability, due to clustering, of finding a pair of objects separated by $r \to r + \delta r$ as compared with a Poissonian distribution, defined as:

$$\delta P = n[1 + \xi(r)]\delta V, \tag{1.13}$$

where n is the mean number density of objects and δV the volume element. As for $w(\theta)$, the Landy & Szalay estimator may be used to calculate $\xi(r)$ from the pair counts $DD(r)$, $RR(r)$ and $DR(r)$, and is often fitted by an analogous power law, expressed as:

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}, \tag{1.14}$$

where α in the angular case is equivalent to $\gamma - 1$.

Power Spectrum

The angular and spatial correlation functions can also be represented by their spherical harmonic and Fourier transforms, respectively, as the angular and spatial power spectra, theoretically encoding the
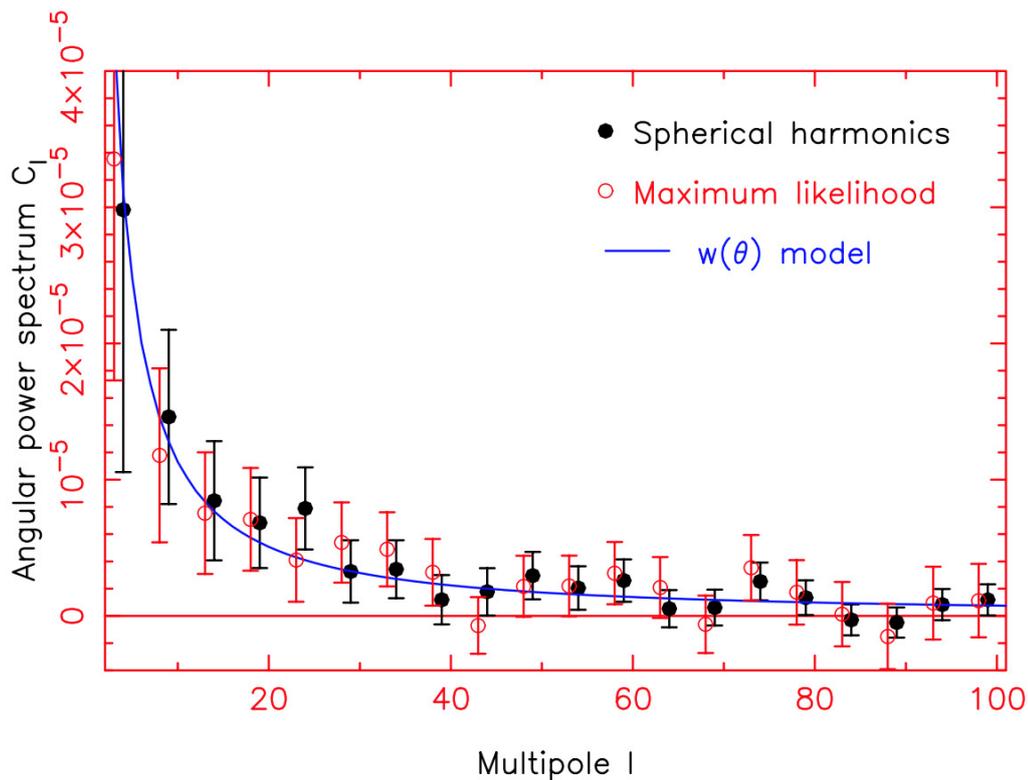
FIGURE 1.4: The angular power spectrum, C_ℓ, by Blake, Mauch & Sadler (2004), for radio sources in NVSS. Spherical harmonic estimation and maximum likelihood estimation methods are compared and both agree well with the C_ℓ inferred from the angular correlation function, $w(\theta)$, measured in Blake & Wall (2002a).

same clustering information. The angular power spectrum, C_ℓ quantifies the amplitude of fluctuations on the angular scale corresponding to ℓ ($\theta \sim 180°/\ell$) by representing the density field as a sum of sinusoidal angular density fluctuations of different wavelengths. It can be related to the angular correlation function by

$$w(\theta) = \frac{1}{4\pi\sigma_0^2} \sum_{\ell=1}^{\infty} (2\ell + 1)C_\ell P_\ell(\cos\theta), \tag{1.15}$$

where σ_0 is the source surface density and P_ℓ is the ℓth Legendre polynomial. The two statistics have different strengths, however, with the highest signal-to-noise ratio for $w(\theta)$ occurring at small angles of up to a few degrees, beyond which Poisson noise dominates. By contrast, C_ℓ has highest signal-to-noise at small ℓ, corresponding to large angular scales, complementing $w(\theta)$. In practice, C_ℓ estimation is complicated by incomplete sky coverage of any given galaxy survey, as each ℓ is dependent on all values of θ. For the 75 per cent sky coverage of the NRAO VLA Sky Survey (NVSS; Section 1.3.4), Blake, Ferreira & Borrill (2004) calculated the angular power spectrum using the spherical harmonic estimation based on that of Peebles (1973), and a maximum likelihood estimation method (see e.g.

Statistic	Dependence	Validity
Spatial correlation function	$\xi(r) \propto r^{-\gamma}$	All r
Angular correlation function	$w(\theta) \propto \theta^{1-\gamma}$	Small θ
Spatial power spectrum	$P(k) \propto k^{\gamma-3}$	All k
Angular power spectrum	$C_\ell \propto \ell^{\gamma-3}$	High ℓ

TABLE 1.1: Functional forms of clustering statistics for pure power-law clustering parametrized by slope γ.

Efstathiou & Moody, 2001; Huterer, Knox & Nichol, 2001), shown in Figure 1.4. Despite the two statistics being appropriate for clustering at different angular scales, Blake, Ferreira & Borrill (2004) show the agreement of the $w(\theta)$ estimation for NVSS in Blake & Wall (2002a) with both estimates of C_ℓ.

The spatial analogue to equation 1.15, relating the spatial correlation function, $\xi(r)$, to the three-dimensional power spectrum, $P(k)$ is given by

$$\xi(r,z) = \frac{1}{2\pi^2} \int_0^\infty P(k,z) \frac{\sin(kr/a)}{(kr/a)} k^2 \mathrm{d}k, \qquad (1.16)$$

where the scale factor $a = (1+z)^{-1}$, and $P(k,z)$ is assumed as a separable function of comoving wavenumber k and redshift z (Baugh & Efstathiou, 1993; Efstathiou & Moody, 2001):

$$P(k,z) = \frac{P(k)}{(1+z)^{\alpha_z}}, \qquad (1.17)$$

for some constant α_z.

The present-day matter power spectrum increases with k initially (less power at the very largest scales), before turning over at $k \sim 0.01$ Mpc^{-1} (corresponding to the horizon scale at matter–radiation equality), and decreasing roughly as a power law $P(k) \propto k^{-3} \ln(k/k_{\mathrm{eq}})^2$ (with some small wiggles; see Section 1.2.5), where k_{eq} is the wavenumber of a mode entering the horizon at the matter–radiation transition. Assuming this power law form for $P(k)$ gives the dependencies described in Table 1.1 for the other related clustering statistics. The shape of this power spectrum is explained by starting with a primordial, featureless power law form of $P(k) \propto k^n$, where $n \sim 1$, predicted from inflation. Before decoupling, overdensities oscillated due to competing radiation pressure and gravitational attraction; only fluctuations on scales larger than the sound horizon were allowed to grow, causing $P(k)$ to grow only for suitably small k. Consequently, the sound horizon expanded, shifting this peak towards smaller k before being frozen at the horizon scale at matter–radiation equality.

1.2.2 Limber Approximation

The angular correlation function can be described as the projection of the spatial correlation function on a sphere. Therefore, if we know the radial distribution of sources $\mathrm{d}n/\mathrm{d}r$, we can project $\xi(r)$ to $w(\theta)$ as follows:

$$w(\theta) = \int_0^\infty \frac{\mathrm{d}n_1}{\mathrm{d}r_1} \mathrm{d}r_1 \int_0^\infty \frac{\mathrm{d}n_2}{\mathrm{d}r_2} \xi(R, \frac{r_1 + r_2}{2}) \mathrm{d}r_2, \tag{1.18}$$

where $R = \sqrt{r_1{}^2 + r_2{}^2 - 2r_1 r_2 \cos\theta}$ is the spatial separation of r_1 and r_2 (the comoving distances to different galaxy populations), assuming a geometrically flat cosmology. This equation gives the angular correlation between two populations with different radial selection functions, and makes the implicit assumption that $\xi(r)$ does not evolve within the overlapping volume of $\mathrm{d}n_1/\mathrm{d}r_1$ and $\mathrm{d}n_2/\mathrm{d}r_2$. This double integral can be simplified by introducing some further assumptions (Limber, 1953; Peebles, 1980; Simon, 2007):

  i. $\mathrm{d}n/\mathrm{d}r$ is a function of only $\bar{r} = (r_1 + r_2)/2$.

  ii. $\xi(r)$ varies significantly only over the separations $\Delta r = r_2 - r_1$.

  iii. Small angle approximation ($\sin\theta \approx \theta$ and $\cos\theta \approx 1$).

The result of these assumptions is the Limber equation:

$$w(\theta) = \int_{-2\bar{r}}^{2\bar{r}} \frac{\mathrm{d}n_1}{\mathrm{d}r_1} \frac{\mathrm{d}n_2}{\mathrm{d}r_2} \int_0^\infty \xi\left(\sqrt{\bar{r}\theta^2 + \Delta r^2}, \bar{r}\right) \mathrm{d}\Delta r. \tag{1.19}$$

Using the distance–redshift relation $\bar{r} = cz/H_0 E(z)$, where $E(z) = \left[\Omega_{m,0}(1 + z)^3 + \Omega_{\Lambda,0}\right]^{\frac{1}{2}}$ for a flat universe, and $\bar{r}$ is the comoving distance at a given redshift, equation 1.19 can be rewritten. Assuming a power law form of $\xi(r)$ as in equation 1.14, the Limber equation becomes:

$$w(\theta) = \frac{H_0}{c} \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{\gamma-1}{2}\right)}{\Gamma\left(\frac{\gamma}{2}\right)} \frac{\int_0^\infty N^2(z) \bar{r}^{1-\gamma}(z) r_0^\gamma(z) E(z) \mathrm{d}z}{\left[\int_0^\infty N^2(z) \mathrm{d}z\right]^2} \theta^{1-\gamma}, \tag{1.20}$$

where it is also assumed that galaxy pairs are drawn from the same selection function $\mathrm{d}N/\mathrm{d}z \equiv N(z)$.

More often than not, the observable quantity is the angular correlation function, where individual redshifts are unavailable with which to calculate r_1 and r_2. In these circumstances, γ is inferred by fitting a power law function to $w(\theta)$ and the spatial clustering information is retrieved through rearranging equation 1.20 to make r_0 the subject (as will be demonstrated in Chapters 4 and 5). This method, however, is only accurate to the extent to which the above assumptions prove reliable, and Simon (2007) investigates the regime of applicability of Limber's equations and the conditions under which it breaks down, finding that this is dependent mainly on the ratio of the width of the galaxy distribution in comoving distance and the mean comoving galaxy distance.

1.2.3 Redshift-Space Distortions

Calculating galaxy distances from their observed redshifts and a cosmological model, the spatial separation between galaxy pairs that we find is given by s (i.e. the distance in redshift space), which may differ from r due to 'redshift-space distortions'. We may directly compute the redshift-space correlation function, $\xi(s)$, but this is systematically different to the true real-space correlation function $\xi(r)$ (likewise for power spectra in real or redshift space)[*]. The difference occurs because measured redshifts are in fact measuring the total apparent velocity of a galaxy, from the combination of Hubble expansion and peculiar velocities induced by the galaxy's local potential, i.e.

$$v_{\text{total}} = v_{\text{H}} + v_{\text{peculiar}}. \tag{1.21}$$

This peculiar velocity is itself composed of two terms: small-scale random motion of galaxies within clusters, and coherent infall, around clusters, of matter streaming from under dense to overdense regions:

$$v_{\text{pec}} = v_{\text{random}} + v_{\text{infall}}. \tag{1.22}$$

This coherent infall is the cause of the Kaiser effect (Kaiser, 1987) – a flattening in the π direction of $\xi(\sigma, \pi)$, where σ and π are the transverse and line-of-sight components, respectively, of the galaxy pair separations. The infall of galaxies reduces their apparent recession velocity when positioned on the far side of a cluster, and boosts their recession velocity when falling in from the near side of a cluster. This is parameterized by β, the infall parameter, which takes into account the large-scale effects of linear redshift-space distortions. Kaiser (1987) showed that, assuming a power law for $\xi(r)$, β can be estimated by

$$\xi(s) = \xi(r) \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 \right), \tag{1.23}$$

in the linear regime, or a more general estimator is given by

$$\xi(\sigma, \pi) = \xi(r) \left[1 + \frac{2(1 - \gamma\mu^2)}{3 - \gamma}\beta + \frac{3 - 6\gamma\mu^2 + \gamma(2 + \gamma)\mu^4}{(3 - \gamma)(5 - \gamma)}\beta^2 \right], \tag{1.24}$$

where μ is the cosine of the angle between r and π (the distance along the line of sight), and γ is the slope of the power law.

At small scales, the random velocities within a cluster (v_{random}) cause an apparent elongation along the line of sight, giving rise to "fingers of God" (see Figure 1.5). The magnitude of this elongation is denoted by $\langle w_z^2 \rangle^{1/2}$, equivalent to the pairwise velocity dispersion σ_{12}, usually expressed as a Gaussian:

$$f(w_z) = \frac{1}{\sqrt{2\pi}\langle w_z^2 \rangle^{1/2}} \exp\left(-\frac{1}{2}\frac{|w_z|}{\langle w_z^2 \rangle^{1/2}} \right). \tag{1.25}$$

---

[*]This section strongly follows Ross et al. (2007).
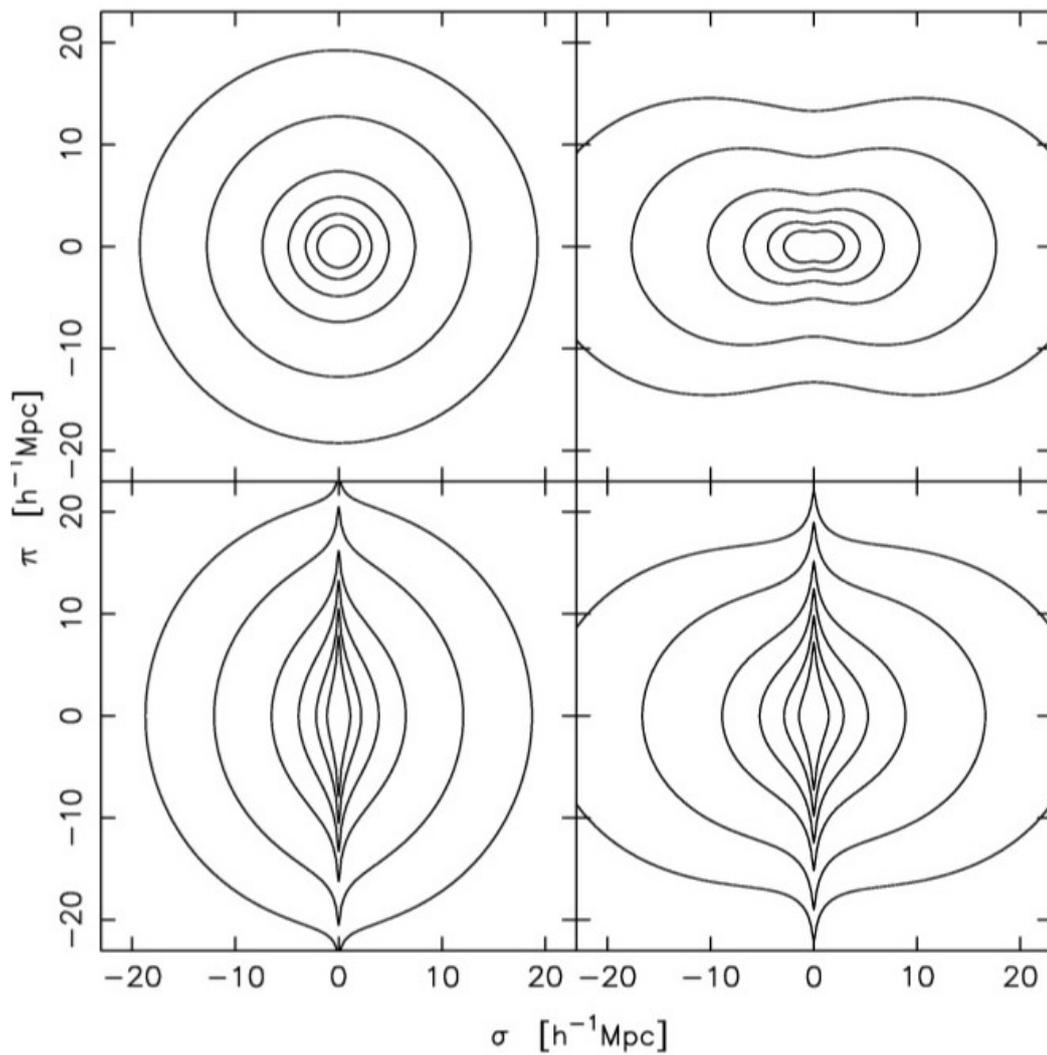
FIGURE 1.5: Model plots of the 2D spatial correlation function, $\xi(\sigma, \pi)$, where σ is the transverse separation, and π is line-of-sight separation (affected by redshift-space effects). The top left panel shows the undistorted model, while the top right panel includes the Kaiser effect of coherent infall, and the bottom left panel shows the elongation due to random pairwise velocities. The bottom right panel shows the combination of both of these effects. (Hawkins et al., 2003)

Combining these small-scale non-linear distortions with the the Kaiser formulae gives an expression for the full model for the observed correlation function $\xi_{\text{obs}}(\sigma, \pi)$:

$$\xi_{\text{obs}}(\sigma, \pi) = \int_{-\infty}^{\infty} \xi\left[\sigma, \pi - w_z(1+z)H(z)\right] f(w_z)\mathrm{d}w_z, \tag{1.26}$$

where ξ is given by equation 1.24.

If the infall parameter β can be calculated from the correlation functions as in the above, it can be used to determine the bias (Section 1.2.4), given an assumed value of Ω_{m} (Kaiser, 1987):

$$\beta \simeq \frac{\Omega_{\text{m}}^{0.6}}{b}, \tag{1.27}$$

where $\Omega_{\text{m}}(z)$, for a flat universe, is given by

$$\Omega_{\text{m}}(z) = \frac{\Omega_{\text{m}}^0(1+z)^3}{\Omega_{\text{m}}^0(1+z)^3 + \Omega_{\Lambda}^0}. \tag{1.28}$$

These redshift-space effects make it difficult to find the real-space clustering parameters r_0 and γ from direct $\xi(s)$ measurements, but they can be found directly through the projected correlation function $\Xi(\sigma)$ (Section 3.3.2) or indirectly through deprojecting the angular correlation function $w(\theta)$ (Section 1.2.2).

1.2.4 Mass Bias

The differences in clustering of different classes of extragalactic objects and the background matter distribution motivates the use of a linear bias parameter, b, as introduced by Kaiser (1984) and Bardeen et al. (1986):

$$b^2(z) = \frac{\xi_{\text{gal}}(r, z)}{\xi_{\text{DM}}(r, z)}, \tag{1.29}$$

where the numerator and denominators are the galaxy and dark matter correlation functions (or may equivalently be replaced by the power spectra), respectively.

The bias parameter can be determined directly if both the real- and redshift-space forms of the correlation function are known. The ratio of the two can be approximated by a quadratic in the infall parameter $\beta \propto 1/b(z)$ (Kaiser, 1987). Therefore, with a good estimate of the real- and redshift-space correlation functions, one may calculate the ratio of the two and hence the bias.

The bias parameter (as a function of redshift) may be defined as equation 1.29 with $r = 8\,h^{-1}\text{Mpc}$, where the numerator can be written

$$\xi_{\text{gal}}(8, z) = \left[\frac{r_0(z)}{8}\right]^{\gamma}. \tag{1.30}$$

and the corresponding function for the denominator is given by Peebles (1980) as

$$\xi_{\mathrm{DM}}(8, z) = \sigma_8^2(z)/J_2,$$ (1.31)

where $J_2 = 72/[(3 - \gamma)(4 - \gamma)(6 - \gamma)2^\gamma]$ and the parameter σ_8^2 is the dark matter density variance in a comoving sphere of radius 8 h^{-1}Mpc. The combination of these equations gives the scale-independent evolution of bias with redshift, given only the correlation length and slope:

$$b(z) = \left[\frac{r_0(z)}{8}\right]^{\gamma/2} \frac{J_2^{1/2}}{\sigma_8 D(z)/D(0)},$$ (1.32)

where $D(z)$ is the cosmological growth function.

Proposed mechanisms for this bias began with the 'high peak' model (Kaiser, 1984), followed by the 'natural bias' model (White et al., 1987; Cole & Kaiser, 1989), which has objects of a given mass collapsing sooner if they lie in a region of a large-scale overdensity, enhancing their abundance with respect to the mean. This is refined using the 'peak–background split' where large-scale density components are added to a pre-existing small-scale density field, perturbing the collapse threshold (Efstathiou et al., 1988; Cole & Kaiser, 1989). Under the assumption of the physical properties of an object being strongly determined by the mass of the parent dark matter halo (e.g. Mo & White, 1996; Matarrese et al., 1997), several models have been devised, using the formalism of Press & Schechter (1974), of bias as a function of halo mass and redshift.

1.2.5   Modern Applications

The above statistics have their own inherent uses in the realm of galaxy formation and evolution; describing the strength of the clustering of observationally disparate galaxy populations may shed light on unrealized relationships between them. Samples with entirely different selection criteria, and redshift distributions, may show similarities in dark matter halo mass, clustering, and evolution. Such findings highlight the time-evolving nature of galaxies themselves, as opposed to the Universe as a whole, and hint at progression through different phases in the life cycle of these sources.

For cosmological applications, however, understanding the clustering properties of a population is a stepping stone towards an understanding of the dark matter distribution from which they are derived. This is an essential part to using observations of galaxies to gain insight into the deeper workings of our cosmological model, using observational phenomena described below.
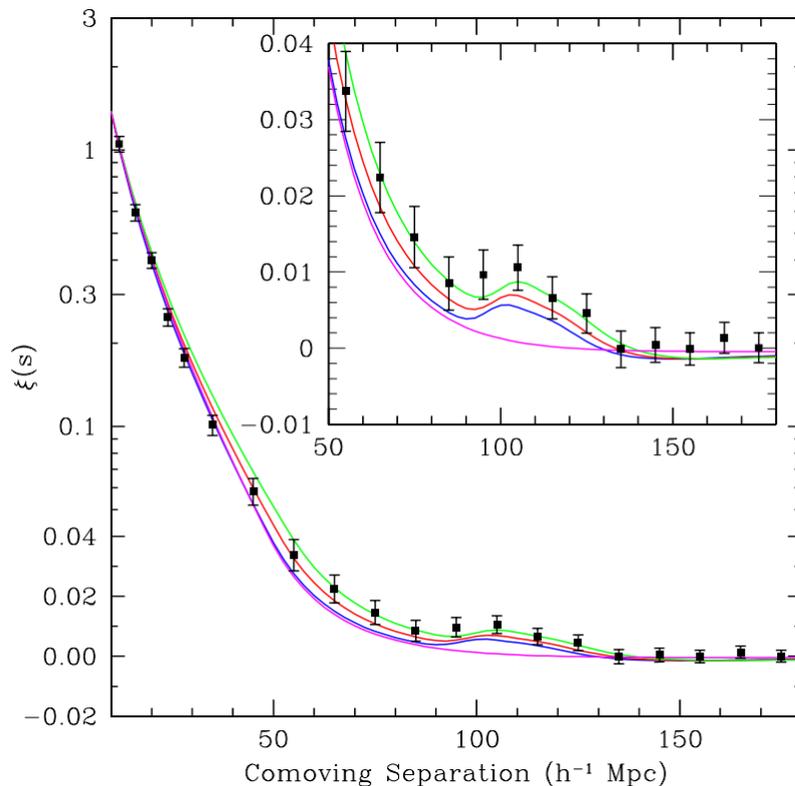
FIGURE 1.6: The characteristic bump in the spatial correlation function - a signature of baryon acoustic oscillations at $\sim 110\ h^{-1}$Mpc (Eisenstein et al., 2005).

Baryon Acoustic Oscillations

The signature of baryon acoustic oscillations (BAOs) can be found directly from the power spectrum or galaxy auto-correlation function. The very early Universe consisted of a tightly-coupled photon-baryon fluid, in which small inhomogeneities gave rise to spherical oscillations driven by the competing outward pressure due to the photons and gravitational force inwards. At the point of photon-baryon decoupling, and recombination of the baryons to form a neutral gas ($z \sim 1000$), the photons stream away (producing the CMB). The original central overdensity in the dark matter remains, along with an overdense, spherical shell of baryons frozen in place. The dark matter overdensities draw in the baryons in the absence of outward photon pressure, and continue to form the seeds of large-scale structure in the subsequent evolution of the Universe, but there persists a slight density enhancement at this characteristic radius (the sound horizon at last scattering). The observational evidence for this phenomenon in galaxy clustering is found in the form of 'wiggles' in the power spectrum, or a single 'bump' in the correlation function (Figure 1.6; e.g. Eisenstein et al., 2005; Percival et al., 2010; Blake et al., 2011); a roughly 1 per cent enhancement of galaxy pairs found at a separation corresponding to the BAO scale.

This scale is intrinsic to the clustering of baryonic matter at any stage of the evolution of the Universe, and expands accordingly.

In much the same way that Type Ia supernovae (SNe) may provide a standard candle, BAOs give us a standard ruler – a distance scale throughout the history of the Universe, which is independent of cosmological model. This provides a route from galaxy clustering to measuring the Hubble constant and other parameters independently from other tests that are used to constrain the current best understanding of cosmology.

Cosmic Magnification

The presence of mass alters the geodesic followed by light rays, causing the deflection of those rays along the line of sight when passing by intervening large-scale structure. This causes distortions and magnification in the observed images of distant astronomical sources and is known as gravitational lensing. The shape distortion of the background galaxies ('cosmic shear') can itself be used to constrain some cosmological parameters (for a review, see Van Waerbeke & Mellier, 2003; Refregier, 2003), as can the magnification of background sources by foreground lensing structures ('cosmic magnification'). This magnification is governed by two effects. The increased flux received from distant sources due to lensing has the effect of bringing into the survey sources that would otherwise have been too faint to be detected. The lensing also stretches the solid angle, reducing the apparent surface density of the lensed background sources. Gravitational lensing should, therefore, leave a cosmic magnification signal in the angular cross-correlation function of two samples of sources with non-overlapping redshift distributions. This signal – the magnification bias – is dependent on the balance struck between the loss of sources due to dilution, and the gain due to flux magnification.

Using the galaxy-quasar cross-correlation function, Scranton et al. (2005) made an 8σ detection of the cosmic magnification signal of quasars lensed by foreground galaxies, both selected from the Sloan Digital Sky Survey (SDSS). They find bright quasars, with steep source counts, exhibit an excess around foreground structure, and faint quasars, with shallow source counts, are in deficit. Since this first successful result, however, Hildebrandt, van Waerbeke & Erben (2009) have made detections with normal galaxies from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS), and Wang et al. (2011) have done likewise in the far-infrared with *Herschel*.

Integrated Sachs-Wolfe (ISW) Effect

Another cosmological probe can reveal itself through the angular cross-correlation function of the CMB and a foreground galaxy survey through which it is observed. The ISW effect (Sachs & Wolfe, 1967) manifests itself in the correlation between between large-scale structure and CMB temperature. In an Einstein-de Sitter universe, the energy gained by a photon falling into a gravitational potential well is exactly cancelled out by the energy lost upon climbing out of the well. In a universe with a dark energy component or modified gravity to the same effect, the local gravitational potential varies with time, and potential wells are stretched throughout the crossing time of the photons. This disparity between potential on entry and exit imparts a net blueshift on the incident photon and, equivalently, an increase in the observed temperature. Specifically, this effect is important at late times ($z \lesssim 1$), once the Universe had begun its transition towards a dominant dark energy component, and accelerated expansion.

The ISW signal is small, compared to the intrinsic temperature anisotropies in the CMB, acting on large scales where cosmic variance most affects CMB uncertainties, and so cross-correlation with local large-scale structure, with extensive sky coverage, is required in order to produce a significant result. Giannantonio et al. (2012) discuss the current state of ISW measurements, and potential problems therewith, as well as producing an update to their own results (Giannantonio et al., 2008). While several detections of the ISW effect had been made previously, cross-correlating the CMB maps from the Wilkinson Microwave Anisotropy Probe (WMAP) with radio, infrared, optical, and X-ray surveys, these were all at low significance and unable to constrain cosmological models. Giannantonio et al. (2008) reached an increased significance of $\sim 4.5\sigma$ by combining surveys to develop a fuller catalogue of local large-scale structure. With the latest CMB measurements from the *Planck* satellite, and corresponding all-sky surveys of massive, low-redshift galaxies (and an understanding of their relationship to the underlying dark matter distribution), the ISW could be a powerful cosmological tool.

1.3 Radio Sources

In this thesis, I use radio sources as cosmological probes, thus here I provide a brief overview of extragalactic radio astronomy: the discovery of radio galaxies, their various classifications, and the current state of large-area surveys and clustering analysis in the radio regime.
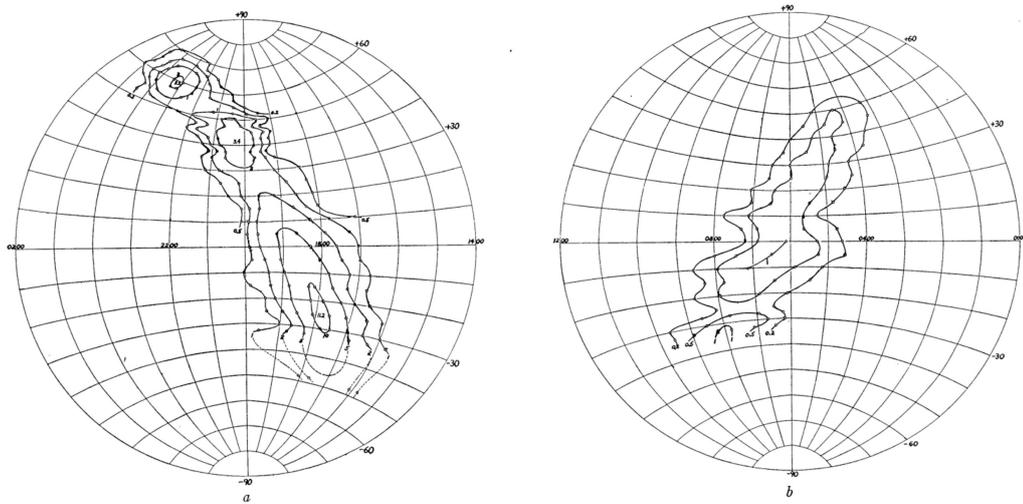
FIGURE 1.7: Reber's 1944 radio sky map at 160 MHz.

1.3.1   Historical Overview

Karl Jansky, upon joining Bell Laboratories in Holmdel, New Jersey in 1928, was tasked with deter-
mining sources of radio interference with a view to a transatlantic radio telephone service. Building an
antenna designed to receive 14.6 m (20.5 MHz) radio waves, mounted on a turntable, by 1932 he had
been able to detect a steady hiss in the radio background, which appeared to reach a daily maximum.
After a time, he was able to conclude that this background was in fact varying with the sidereal day
and rather than being associated with the Earth or Sun, was in fact coming from a well-determined
location in the constellation of Sagittarius (Jansky, 1933a,b). Further observations found radio sources
distributed throughout the Milky Way and, given that the Sun was such a weak emitter, led to the con-
clusion that the noise was from non-stellar sources (Jansky, 1935), also speculating that it was generated
in an ionized interstellar medium (ISM).

    In order to observe Jansky's cosmic radio emission, engineer Grote Reber built a 30-ft parabolic
reflector in his garden in Illinois in 1937, initially attempting observations at 3 GHz before eventually
obtaining good results at 160 MHz (Reber, 1940). His subsequent paper (Reber, 1944) contained a
detailed map of the radio sky he observed (Figure 1.7), showing the principal maximum towards the
Galactic centre as well as two other maxima corresponding to what we now know as Cygnus A and
Cassiopeia A. Following Reber's work, Jan Oort at Leiden Observatory reasoned that efforts should be
put towards finding a spectral line in the radio in order to trace the structure and dynamics of the ISM
using the Doppler shift of this line, unobscured by dust in the Galaxy. Oort's student soon identified

the 21 cm line emitted by atomic hydrogen in transition between parallel and anti-parallel spin states of the proton and electron (van de Hulst, 1945), confirmed independently by Josef Shklovsky, along with several molecular lines (Shklovsky, 1949).

J. S. Hey, working initially on radio interference arising from the Sun, published a 64 MHz radio map confirming the features of Reber's map (Hey, Phillips & Parsons, 1946), before noticing the maximum at Cygnus A showed fluctuations in intensity on a timescale of approximately 10 seconds (Hey, Parsons & Phillips, 1946), implying source dimensions comparable to a star (i.e. less than 10 light-seconds across), rather than diffuse emission from interstellar gas. Controversy followed as to whether this Galactic radiation was the superposition of 'radio stars' distributed throughout the Milky Way, or perhaps thermal bremsstrahlung from electron–ion collisions in a hot interstellar gas requiring temperatures in the region of 150,000 K (Townes, 1947). By 1950, however, it had been conclusively shown that the intensity fluctuations of discrete radio sources were the result of ionospheric scintillation (Ryle, 1950) and therefore imposed no constraint on the source size. Meanwhile, Bolton, Stanley & Slee (1949) had published the first optical identifications of the brightest radio sources: Taurus A (associated with the Crab Nebula), Virgo A (M87) and Centaurus A (NGC4528). Duyvendak (1942) had deduced the Crab Nebula to be the remnant of a supernova from 1054, while M87 and NGC4528 were known to exhibit optical peculiarities.

Alfvén & Herlofson (1950) were the first to suggest that these radio stars were emitting synchrotron radiation, but while their model was unrealistic, a consensus was reached quite rapidly that the nature of the emission was non-thermal in origin (e.g. cosmic ray electrons gyrating in the Galactic magnetic field; Kiepenheuer 1950). Optical observations of the precise positions of radio sources by Baade & Minkowski (1954) established the real significance of radio astronomy in terms of highlighting novel and extreme astrophysical phenomena. Cassiopeia A showed radial velocities varying by $> 4,000$ km s^{-1} from wisp to wisp, evidently the remnant of a supernova, consistent with Shklovsky's earlier suggestion that the synchrotron emission observed in the Crab Nebula was produced by energetic electrons resulting from the supernova. Cygnus A, however, proved to be remarkable in its own way. Faint in the optical, but one of the brightest radio sources in the sky, its spectrum exhibited a recession velocity of 16,000 km s^{-1} ($z = 0.056$), setting it apart from more local, and fainter, extragalactic sources (M87 and NGC4528) or sources in our galaxy associated with supernovae (Crab Nebula and Cassiopeia A).

Radio astronomy had emerged from a terrestrial engineering experiment, first encouraging work on our solar system and radio sources within our galaxy, but the real surprises and subsequent scientific
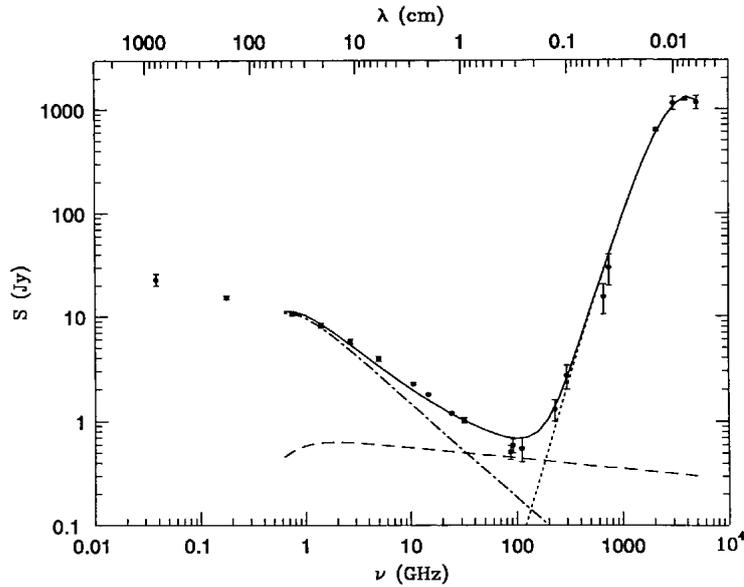
FIGURE 1.8: Radio and infrared spectrum of the starburst galaxy M82. The broken lines show the contribution of separate components: synchrotron (at low frequencies), free-free, and thermal (at high frequencies) radiation (Condon, 1992).

breakthroughs came from those initial few sources which were identified with galaxies beyond the Milky Way, and the two distinct types of galaxies they were found to be. Broadly speaking, these are star-forming galaxies (Section 1.3.2) where the radio emission arises from supernovae and HII regions, or active galaxies (Section 1.3.3) where the energy emitted is derived from accretion onto a central massive black hole.

1.3.2   Star-Forming Galaxies

Most nearby galaxies can be detected in the radio with low luminosities, our Galaxy having a luminosity of $L \sim 3 \times 10^{21}$ W Hz^{-1} at $\nu = 1.4$ GHz (Berkhuijsen, 1984). Condon (1992) presents a review of low-power radio emission from normal galaxies such as ours, resulting from star formation. Figure 1.8 shows the radio and far-infrared (FIR) spectrum of M82, a bright starburst galaxy, showing the contribution made by different emission mechanisms over this frequency range. Due to their power-law spectra ($S \propto \nu^{\alpha}$), these varieties of radio emission can be distinguished observationally by their spectral index α *:

$$\alpha(\nu) = \frac{\partial \log S(\nu)}{\partial \log \nu}. \tag{1.33}$$

At low frequencies ($\nu \lesssim 30$ GHz), the dominant source of radio emission in normal, star-forming

---

*Sign convention for α varies, but any references to it in this thesis will be based on the definition $S \propto \nu^{\alpha}$.

galaxies is non-thermal synchrotron radiation. This is produced by ultra-relativistic electrons ($E \gg m_e c^2$) accelerated through a magnetic field by the energy released from supernovae. These electrons emit most of their energy at a critical frequency ν_c,

$$\left(\frac{\nu_c}{\text{GHz}}\right) \sim 0.016 \left(\frac{B\sin\theta}{\mu\text{G}}\right)\left(\frac{E}{\text{GeV}}\right)^2, \tag{1.34}$$

for electron energy E, moving at a pitch angle θ, in a magnetic field of strength B. For a typical spiral disk field strength $B \sim 5~\mu\text{G}$, therefore, two decades in frequency at 0.1–10 GHz can be generated by electrons with energies in the range 1–10 GeV. As shown in Figure 1.8, synchrotron radiation can be confirmed with observations at multiple low frequencies, highlighting a significant negative spectral index, $\alpha \sim -0.7$.

Free-free (or thermal bremsstrahlung) emission is associated with HII regions - large, low-density clouds of atomic hydrogen which has been partially ionized by recent star formation. The long wavelength emission (dominating the starburst spectrum at ~ 100 GHz in Figure 1.8) is produced when free electrons scatter off ions in the HII regions without being captured, with low-energy radio photons being produced in those weak interactions where the electrons are only deflected slightly. At high frequencies, the optical depth τ of the HII region is very low, and the free-free emission has a spectral index $\alpha \sim -0.1$. At sufficiently low frequencies that the HII region becomes opaque ($\tau \gg 1$), the spectrum approaches that of a black body and the spectral index rises to $\alpha \sim 2$, giving a practical lower frequency limit to the observable range of this contribution to the radio flux. The almost flat spectrum above this critical frequency means thermal bremsstrahlung emission affects the synchrotron sources at lower frequencies as well as the higher frequency thermal emission.

At higher frequencies ($\nu \gtrsim 300$ GHz), we move firmly into the far infrared (FIR) region of the spectrum where thermal radiation overwhelms any other emission mechanism. Here we observe the ultraviolet (UV) light produced by young stars being absorbed and re-emitted by the surrounding gas and dust. The heating of this dust as a result of the star formation results in the in the re-radiation of this light with a black body spectrum extending into the FIR, with a spectral index of $\alpha \sim 2$ (see Figure 1.8). A surprising correlation has been found to exist between this thermal FIR radiation and the non-thermal synchrotron emission found at low frequencies.

The FIR–radio correlation arises due to the common link with luminous, massive stars and their end products – dust, supernovae and cosmic rays. Expressed as a simple model (e.g. Voelk, 1989; Lisenfeld, Voelk & Xu, 1996), the massive stars emitting in the UV, and heating the dust producing the thermal FIR radiation, are also responsible for the supernovae which accelerate electrons later found to
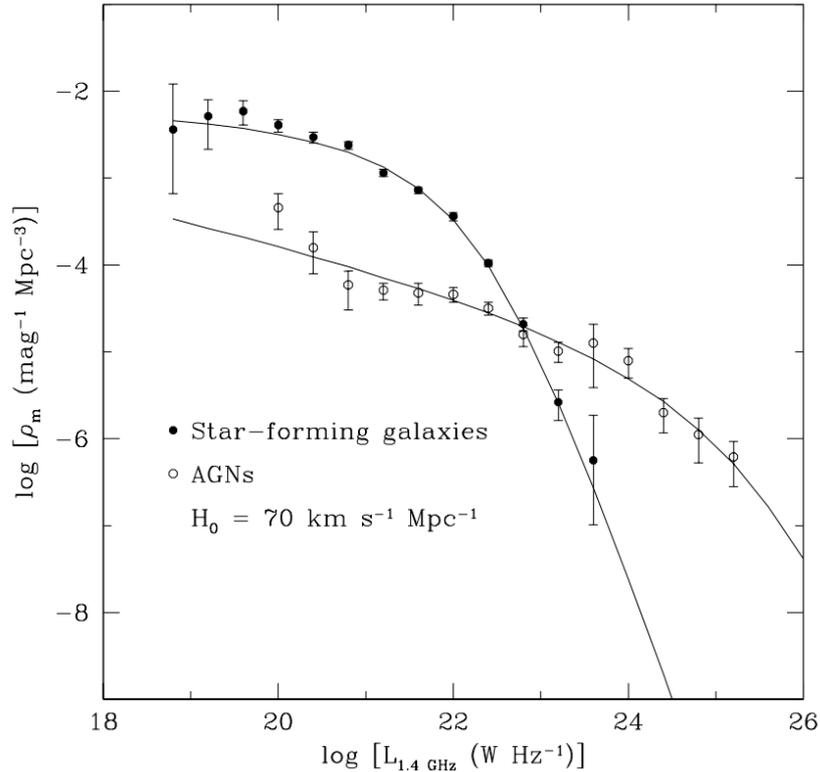
FIGURE 1.9: Luminosity functions of star-forming galaxies and AGN, showing the transition at $L \sim 10^{23}$ $W\,Hz^{-1}$, above which AGN dominate (Condon, Cotton & Broderick, 2002).

emit synchrotron emission in the radio as they move through the galactic magnetic field. The relation between these two star formation tracers is surprisingly tight (de Jong et al., 1985; Helou, Soifer & Rowan-Robinson, 1985; Condon, 1992; Garrett, 2002), and spans many orders of magnitude in luminosity, gas surface density, and photon, cosmic-ray and magnetic energy density. It remains unclear why such a strong correlation would exist, requiring many models to rely on significant fine tuning. Discussion of the possible reasons for this can be found in Lacki, Thompson & Quataert (2010) and Lacki & Thompson (2010) and, in fact, Bell (2003) argues that the linear FIR–radio correlation is a conspiracy, with both underestimating star formation rates.

1.3.3 Active Galactic Nuclei

As was evident from the early optical identifications of Baade & Minkowski (1954), some radio sources are exceptionally powerful, while being associated with relatively faint galaxies in the optical from large distances. Additionally, the interferometry of Jennison & Gupta (1956) identified the strong radio source Cygnus A as a double, with an overall size of more than an arcminute, suggesting similar objects

could be readily discovered at much greater distances. Subsequent advances in radio astronomy led to the compiling of the third Cambridge catalogue (3C; Edge et al. 1959) at 159 MHz, and a revised version at 178 MHz (3CR; Bennett 1962), the latter detecting all sources with a declination $\delta > -5°$ to a flux limit of 9 Jy *. The improved angular resolution of these catalogues allowed identification with optical sources – mostly galaxies, but in some cases star-like objects.

All of these powerful extragalactic radio sources derive their emission from synchrotron radiation, but star-forming galaxies tend to be found at low redshift with relatively low radio power $P_{1.4} \lesssim 10^{23}$ W Hz^{-1} (Mauch & Sadler, 2007). A new mechanism is required to explain the highly extended nearby sources like Cygnus A and Centaurus A, the exceptionally distant and powerful star-like sources, and a range of shapes, sizes, luminosities and optical features in between. These sources where the radio emission cannot be accounted for by star-forming processes generally fall under the category of 'active' galaxies, or active galactic nuclei (AGN).

Emission from 'normal' galaxies is dominated by stars and, with stellar atmospheres generally being in hydrostatic and thermal equilibrium, this predominantly thermal emission is the sum of Planck spectra of the individual stars in the galaxy. Given temperatures in the range of 3,000 K $\lesssim T \lesssim$ 40,000 K, the spectrum of a normal galaxy is largely confined to a wavelength range between $\sim 4,000$ Å and $\sim 20,000$ Å. A galaxy with significant star formation and dust content may extend this range to shorter wavelengths (young, hot stars) as well as longer wavelengths (thermal emission from dust heated by these stars). A small fraction of galaxies, however, have a far broader SED than could be produced by these thermal components alone.

The observed AGN population is classified into subgroups according to observational properties, some of which are largely historical and may obfuscate the underlying physical processes involved. In broad terms, we identify an AGN as having one or more of the following observed properties:

- A compact nucleus much brighter than a similar region in a normal galaxy;

- Non-stellar (or non-thermal) continuum emission;

- Strong emission lines;

- Variability in continuum and/or emission lines on short time scales.

Whilst initially controversial, the prevailing understanding of AGN is currently that they are powered

---

* 1 Jansky (Jy) $\equiv 10^{-26}$ W m^{-2} Hz^{-1}

by accretion onto a central massive black hole (Salpeter, 1964; Zel'dovich & Novikov, 1965; Lynden-Bell, 1969, 1978), upon which a unified model of AGN is based (see Section 1.3.3). Some of the observational subgroups of this AGN paradigm are described below.

Seyfert Galaxies

After the first optical spectrum of an active galaxy was obtained by E. A. Fath in 1908, noting strong emission lines in NGC 1068 later resolved to have widths of several hundred $\mathrm{km\,s^{-1}}$, a study of 6 'spiral nebulae' by Carl Seyfert (Seyfert, 1943) set out the template for a new morphological classification of galaxies. Seyfert galaxies were firstly identified by their appearance – ordinary spiral galaxies with high surface brightness nuclei - but their spectra are also characterized by strong, high-ionization emission lines and are further subdivided into two categories, Seyfert 1 and Seyfert 2, according to the widths of these lines (Khachikian & Weedman, 1974).

In Seyfert 1 galaxies, the forbidden lines are narrow, corresponding to velocities of a few hundred $\mathrm{km\,s^{-1}}$, but the permitted emission lines are very broad, corresponding to velocities of 1,000–5,000 $\mathrm{km\,s^{-1}}$. This is indicative of a high-velocity, high-density ionized gas (broad-line region), where the forbidden transitions are collisionally suppressed, which is physically separate from the low-velocity gas (narrow-line region). Seyfert 2 galaxies differ from Seyfert 1 galaxies in that they exhibit only narrow emission lines for both permitted and forbidden transitions, suggesting an origin from the same region in the host galaxy. A further class of AGN related to Seyferts are low-ionization nuclear emission regions (LINERs). These are similar to Seyfert 2 galaxies, except with forbidden lines arising from less-ionized atoms.

Radio Galaxies

Radio galaxies are characterized by their relatively strong radio emission, distinguishing them from normal galaxies exhibiting weak radio emission ($P_{1.4} \lesssim 10^{23}\,\mathrm{W\,Hz^{-1}}$), largely due to supernova remnants. These tend to be associated with elliptical host galaxies, but otherwise can be categorized analogously to Seyferts as broad-line radio galaxies (BLRGs) and narrow-line radio galaxies (NLRGs), based on their optical spectra. Additionally, radio galaxies are distinguished based on radio morphology and often have extended structures several hundred kiloparsecs or even megaparsecs across. They often take the form of double lobes fed by jet-like structures extending from a compact core.

Fanaroff & Riley (1974) divided these extended radio galaxies into two morphological classes. FRI sources are weaker, with 'hot spots' either side of a bright nucleus which span less than half the overall
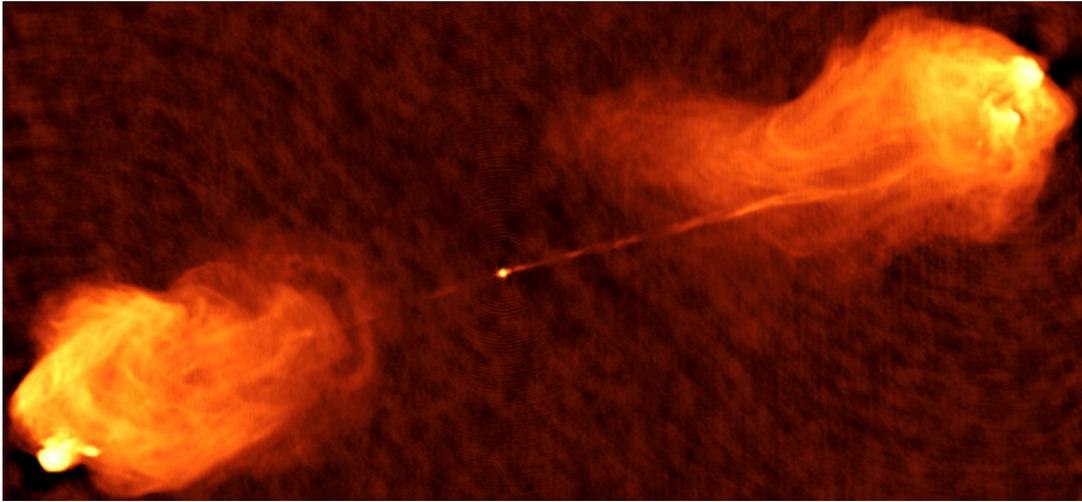
FIGURE 1.10: The radio source Cygnus A, showing double lobe features arising from relativistic jets of electrons, emitted from a central active galactic nucleus, interacting with the magnetic field of the galaxy. Image courtesy of NRAO/AUI.

size of the galaxy, and surface brightness decreasing towards the edges. FRIIs are limb-brightened and more luminous, with the transition in power between the two classes occurring at $P_{1.4} \gtrsim 10^{25}$ W Hz^{-1}. Figure 1.10 shows a radio image of Cygnus A, one of the earliest radio sources discovered, now recognizable as a prototypical FRII class galaxy. The jet structures driving the radio emission in these sources emanate from the AGN at the centre of the host galaxy, with one jet often fainter than the other, or missing entirely from observations. This is because relatively few such sources are observed perpendicular to the jet axis, and relativistic (Doppler) beaming causes an apparent enhancement in brightness of the approaching jet, while the counter-jet appears fainter or absent. The hotspots at either side indicate a surrounding medium into which the collimated jets are moving, resulting in shocks, with the relativistic electrons spreading throughout the lobe structure. The synchrotron radio emission in these regions is produced by these fast-moving electrons as they move through the magnetic field of the galaxy.

Spectral index is used as another common means of subdividing radio galaxies into steep-spectrum or flat-spectrum categories, with the division made at $\alpha \simeq -0.4$. Extended sources tend to have a steep spectrum, while more compact sources have a flat spectrum. This is due to the spectrum of the compact AGN core showing the superposition of emission from different optically thick regions, whereas the extended emission from lobes/jets is optically thin.

Quasars

Early radio surveys such as 3C/3CR found some sources at high Galactic latitude which were coincident with objects that looked like stars on normal photographs, the first such discovery being 3C 48 which was associated with a magnitude 16 'star' (Matthews & Sandage, 1963). These quasi-stellar radio sources (quasars) were found to have confusing spectra with very broad unidentified emission lines, and photometry revealing that they were anomalously blue compared to normal stars. The importance of quasars was soon realized once Schmidt (1963) identified the emission lines in 3C 273 as the hydrogen Balmer series and MgII at a redshift of $z = 0.158$ – one of the largest measured at the time, and an order of magnitude larger than the original Seyfert galaxies. The extreme distances of these sources implied enormous luminosity, with 3C 273 being the brightest known quasar ($B = 13.1$ mag), corresponding to about 100 times as luminous as normal galaxies like the Milky Way or M31.

While their discovery was based on their radio emission, the majority of quasi-stellar objects (QSOs) are not, in fact, powerful radio emitters (Sramek & Weedman, 1980; Condon et al., 1981). The terms 'quasar' and 'QSO' are now used interchangeably and are subdivided into radio-loud quasars (RLQs) and radio-quiet quasars (RQQs) – an observational distinction with no physical differences. Likewise, there appear to be no fundamental differences between quasars and Seyfert 1 galaxies, other than the AGN luminosity which dominates the quasar host galaxy, although Seyferts tend to be found in spiral galaxies.

Further subclasses of quasars are BL Lac objects and optically violent variables (OVVs), collectively known as 'blazars'. While most AGN exhibit variability in their continuum emission across all wavelengths, from X-ray to radio, OVVs are characterized by strong and rapid optical variability on time scales of less than one day, suggesting the regions responsible for this emission are less than 1 light-day across. This variation is observed at other wavelengths, with the time scale of variability typically decreasing with wavelength. OVVs are also characterized by their relatively strongly polarized optical emission compared to ordinary quasars.

BL Lac objects, named for their prototypical source BL Lacertae (originally thought to be a variable star), are similar to OVVs in that they are strong radio emitters, highly variable in optical and X-ray emission, and strongly polarized in the radio and optical. They differ from OVVs in their lack of emission or absorption lines in their spectra. This may be due to intrinsic variability in the sources though, with emission lines being visible only when the continuum is relatively faint.
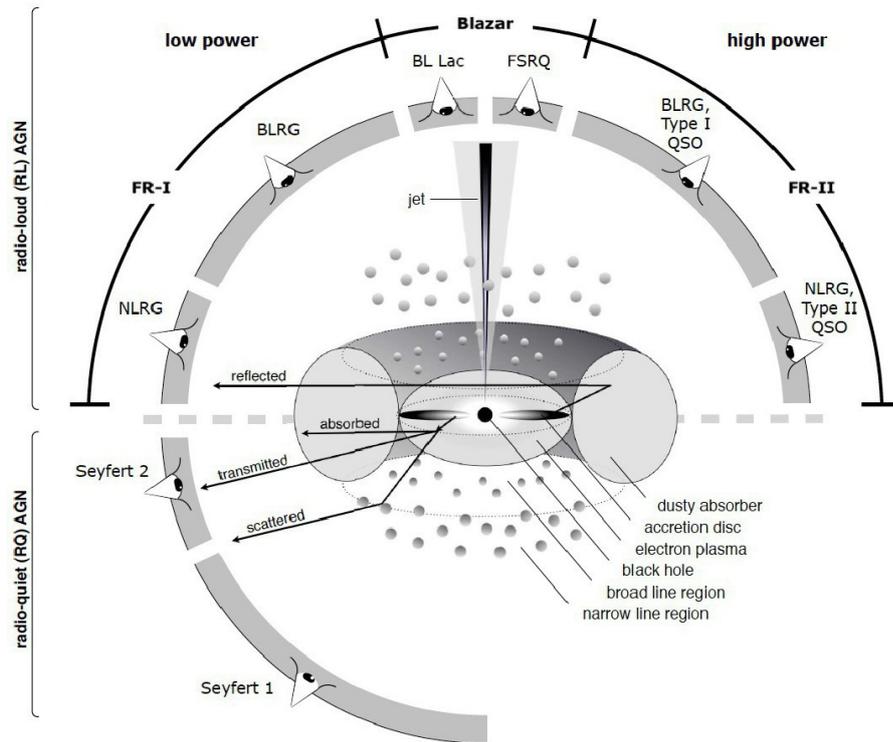
FIGURE 1.11: A schematic diagram of the archetypal AGN, showing the physical components and orientation giving rise to particular observed AGN types (Beckmann & Shrader, 2012).

The Unified Model of AGN

The above descriptions of the independently discovered classes of objects known as Seyfert galaxies, radio galaxies, quasars and blazars are written from the perspective of an established consensus as to their common physical origin, but this relationship between them was not immediately apparent. The first suggestion that various types of AGN observations were a result of different viewing angles was made by Scheuer & Readhead (1979), suggesting RLQs were an aligned version of RQQs. While this theory was not robust, Orr & Browne (1982) proposed the unification of core-dominated and lobe-dominated quasars such that flat-spectrum quasars are aligned versions of steep-spectrum quasars. Subsequently, a full picture of a unified AGN model has been built on similar arguments where observational differences are attributed to intrinsic luminosity differences or orientation effects. A detailed historical and physical explanation of this model is given in review papers by Antonucci (1993) and Urry & Padovani (1995), but a severely abridged version is given here.

Figure 1.11 shows a schematic interpretation of the unified model of AGN as it is currently understood, with the physical components of the AGN labeled, and lines of sight marked with the corresponding AGN class observed. The power source is the supermassive black hole at the centre of the

galaxy, accompanied by a thin, hot accretion disk occupying in the region of 1 parsec around the black hole. Vast amounts of gravitational energy are released as matter falls into the black hole, and the extremes of density, temperature and velocity give rise to the majority of continuum emission from the X-rays to optical wavelengths. Surrounding the accretion disk is the broad line region (BLR), which consists of the high-velocity dense gas responsible for broad emission lines from permitted transitions (cf. Seyfert 1s, Type 1 QSOs, BLRGs). Beyond the BLR is a torus of cold, dusty material, which is necessary to obscure the central region of the AGN in some cases, surrounded by clouds of low-density gas described as the narrow line region (NLR) with a scale of around 100 parsecs. When the viewing angle is suitably close to the plane of the disk, the torus obscures the accretion disk and BLR, and we observe Seyfert 2s and NLRGs.

Radio morphology is explained by both luminosity and viewing angle, with FRII galaxies considered higher-power FRIs, both generating powerful collimated jets of material originating from the central black hole and accretion disk. The double lobe features generated by optically thin synchrotron emission of the accelerated electrons perpendicular to the AGN disk and torus are associated with steep-spectrum radio emission. This distinction arises as the more powerful FRII sources emit highly energetic jets which are able to extend beyond the host galaxy, allowing interaction with the intergalactic medium (IGM) while maintaining relativistic energies. As the viewing angle moves away from the plane of the torus, the nucleus may come into view with its own flat-spectrum radio emission, due to the superposition of many optically thick synchrotron spectra, at the base of the AGN jets. Increasing the angle further, the same extended radio source would appear first as a quasar and then a blazar when viewed close to and along the jet axis, respectively.

An orientation-based model, however, is unable to explain the lack of radio emission in many AGN, with such powerful emission in others. Radio-loud QSOs are found to have black hole masses greater than $10^8 \, M_\odot$ (McLure & Jarvis, 2004), while radio-quiet QSOs may have black hole masses as low as $10^7 \, M_\odot$, suggesting this mass to be a possible limiting factor in radio emission. Alternatively, the smaller numbers of radio-loud AGN than radio-quiet may lend credit to the suggestion that as part of the life cycle of an AGN, a phase of radio loudness occurs.

1.3.4 Radio Surveys

While early radio surveys such as 3C and its successors proved useful in firmly establishing the field of radio astronomy, and highlighting the variety to be found in radio-emitting, extragalactic sources, large steps forward in terms of sensitivity were required to build radio catalogues which would include more

than only the most powerful sources and sample higher-redshift populations. Single-dish facilities (such as Jodrell Bank Observatory in the UK, and the Green Bank Telescope at the National Radio Astronomy Observatory (NRAO) in the USA) have performed admirably in the task of producing large radio surveys of tens of thousands of sources, but nowadays it is large radio interferometer arrays which carry the majority of the workload.

The noise level, σ, achieved by an interferometer (with detection limit $S_{\mathrm{lim}} = 5\sigma$) is given by:

$$\sigma = \frac{\sqrt{2}k_B T_{\mathrm{sys}}}{A_{\mathrm{eff}}\eta_q}\sqrt{\frac{1}{N_{\mathrm{base}}t\Delta\nu}}, \tag{1.35}$$

where T_{sys} is the system temperature, A_{eff} is the effective antenna area, η_q is the sampling efficiency, $N_{\mathrm{base}} = N(N-1)/2$ is the number of baselines (for N antennas), t is the integration time, and $\Delta\nu$ is the bandwidth. While a greater collecting area is clearly ideal, there is a practical limit to the size of a single dish, but also of great importance is the angular resolution of the survey. Long radio wavelengths require a large telescope diameter, which is facilitated by multiple antennas extending to long baselines, the resolution (in arcseconds) being given by:

$$\theta = 2.5 \times 10^5 \times \frac{\lambda}{D}, \tag{1.36}$$

where the λ is the wavelength observed, and D is the telescope diameter. Even for a 305m single dish such as at Arecibo, this corresponds to a resolution of almost $3'$ at 1.4 GHz, whereas the Very Large Array (VLA) in its extended (A) configuration has a resolution of $1.5''$.

Angular resolution is improved for a given instrument observing at higher radio frequencies, but at the expense of field of view and therefore survey speed (and brightness sensitivity), with the number of pointings required being inversely proportional to the field of view. Furthermore, there is some frequency dependence of the type of sources that will dominate a given survey, with low-frequency surveys being biased towards steep-spectrum or 'lobe-dominated' sources, and high-frequency surveys being more 'core-dominated' with flatter radio spectra.

Finally, the 'wide and shallow' versus 'narrow and deep' debate over maximising the number of sources detected hinges on the slope of the source counts. If the integral source counts (number of sources with a flux density greater than S) scale as $S^{-\beta}$, the number of sources detected in a given area scales with the integration time as $t^{\beta/2}$. Therefore, $\beta < 2$ implies that increasing area is more efficient than increasing integration time (and depth), while $\beta > 2$ favours depth over area for increasing source detections. At radio frequencies, the differential source count slope is decidedly greater than -3 ($\beta < 2$), making area a priority. With sky coverage and some complementarity in resolution and frequency being of such importance, the following are some of the major radio surveys of recent years:
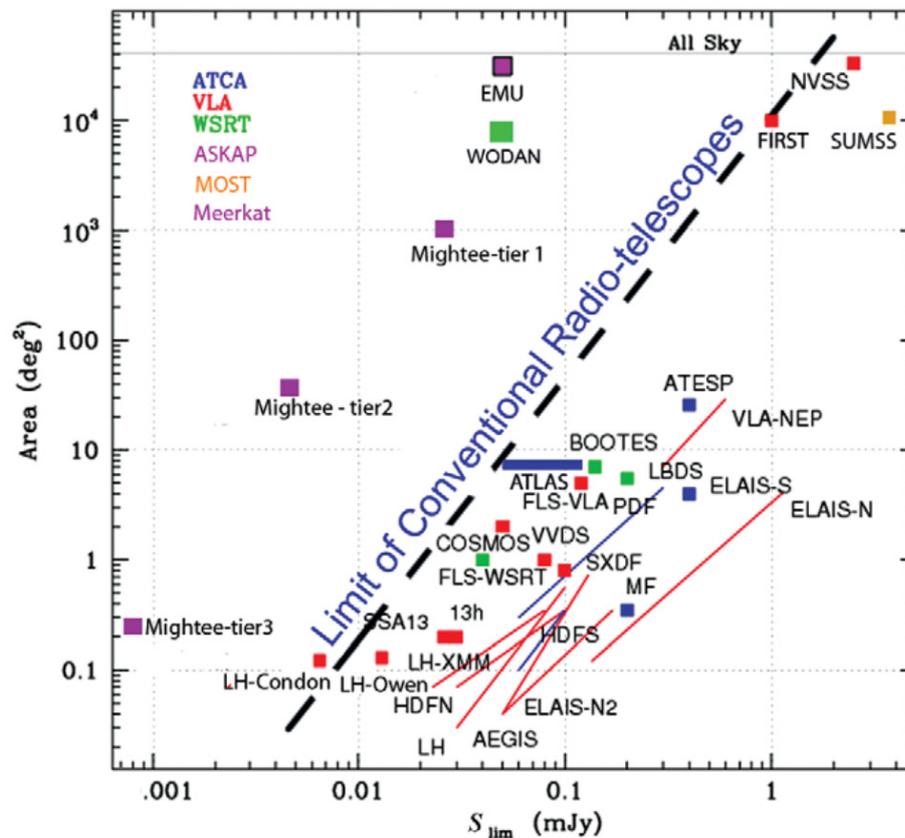
FIGURE 1.12: Survey area versus sensitivity (at 1.4 GHz) for a variety of existing radio surveys, and the forthcoming EMU, WODAN and MIGHTEE surveys with the ASKAP, WSRT and MeerKAT facilities, respectively (Norris et al., 2013).

FIRST The Faint Images of the Radio Sky at Twenty-cm survey (FIRST; Becker, White & Helfand, 1995) was carried out at an observing frequency of 1.4 GHz ($\lambda = 21$ cm) with the VLA in B configuration. The most recent catalogue contains 946,464 sources covering over 10,000 square degrees ($\sim 8,500$ in the northern galactic cap and $\sim 1,500$ in the south) with an angular resolution of $5.4''$ (FWHM) to a completeness of 95 per cent at $S > 2$ mJy.

NVSS The NRAO VLA Sky Survey (NVSS; Condon et al., 1998) is a similar survey to FIRST, using the VLA in the more compact D and DnC configurations ($45''$ resolution), both following from two competing proposals in 1990. Two surveys were recommended simultaneously, using the same pointing grids, but with NVSS having greater sky coverage ($\delta > -40°$), and increased sensitivity to low surface brightness sources, at the expense of coarser resolution, and a slightly higher detection limit. The catalogue contains almost 2 million discrete sources with flux densities greater than $S \approx 2.5$ mJy.

WENSS The Westerbork Northern Sky Survey (WENSS; Rengelink et al., 1997) is a low-frequency survey carried out on the Westerbork Synthesis Radio Telescope (WSRT) at 325 MHz ($\lambda = 92$ cm). With a 5σ depth of 18 mJy and angular resolution of $54''$, it complements the higher-frequency NVSS, surveying the entire sky north of $\delta = 28.5°$, identifying 229,420 sources.

SUMSS The Sydney University Molonglo Sky Survey (SUMSS; Bock, Large & Sadler, 1999), using the Molonglo Observatory Synthesis Telescope (MOST), covers the 25 per cent of the sky south of $\delta = -30°$, at the intermediate frequency of 843 MHz ($\lambda = 36$ cm), identifying 211,063 discrete sources, as of 2008. The complementary sky coverage, similar depth (~ 6 mJy) and resolution ($45''$) to NVSS allows the two surveys to be combined, mapping the entire radio sky.

New and Future Facilities

With the Square Kilometre Array (SKA) on the horizon, increased focus has been placed on synergy between different large surveys for the sake of efficiency as well as combined scientific goals. In Australia and South Africa particularly, new telescopes are being built while others such as the VLA have been upgraded. The Low Frequency Array (LOFAR; van Haarlem et al., 2013), for example, has begun observations of the entire northern hemisphere ($\delta < +30°$) with an array based in the Netherlands and international baselines extending through Germany, France, the UK and Sweden. As part of a three-tiered survey, the large-area survey of the entire hemisphere will reach an rms detection limit of 2, 1 and 0.07 mJy at 15–40, 40–65 and 120–180 MHz, respectively, anticipating 30 million radio sources.

The WSRT, updated with APERTIF phased array feeds (PAFs), gives us the Westerbork Observations of the Deep APERTIF Northern sky (WODAN), which will survey 25 per cent of the sky ($\delta > +30°$) at 1.4 GHz with $15''$ resolution and an rms sensitivity of 10 μJy (although source confusion may require raising this to 20 μJy). The Australian SKA Pathfinder (ASKAP; Johnston et al., 2007), consisting of 36 12 m antennas distributed over a region 6 km in diameter, will yield the Evolutionary Map of the Universe (EMU; Norris et al., 2011) survey. This will perfectly complement WODAN, surveying the remaining southern sky ($\delta < +30°$) with $10''$ resolution and 10 μJy sensitivity. These two surveys combined will observe the entire sky, producing a catalogue of almost 100 million sources, with 45 times the sensitivity and 3–4 times the resolution of NVSS.

The South African SKA precursor telescope is MeerKAT, which will consist of 64 13.5 m dishes. Phase 1 of MeerKAT will have a maximum baseline of 8 km, with the additions in Phase 2 extending this to 20 km, improving the resolution at 1.4 GHz from $6.5''$ to $2.6''$. The MeerKAT International GHz Tiered Extragalactic Exploration survey (MIGHTEE; Jarvis, 2012), will be in three tiers aimed

at deeper higher resolution observations below the confusion limit of WODAN and EMU. The largest area part of the survey will cover 1000 square degrees to an rms depth of 5 μJy, and the deeper surveys will reach 1 and 0.1 μJy over 35 and 1 square degree, respectively.

Figure 1.12 shows how the area and depth of the next generation of radio surveys compares with what is currently available. This figure is taken from Norris et al. (2013) who provide a detailed discussion of radio continuum surveys in advance of the completion of the SKA. Specific discussion of the applications of some of these surveys to cosmology is given by Raccanelli et al. (2012).

1.3.5 Clustering of Radio Sources

Extragalactic radio sources make useful probes of large-scale structure, being readily detectable up to high redshifts ($z \sim 6$). Being unaffected by dust extinction, radio surveys are able to provide unbiased samples of larger volumes than would be available to an optical survey. Unfortunately, difficulty in optically identifying and obtaining redshifts for these individual sources (see e.g. McAlpine et al., 2012) confines many clustering analyses to viewing spatial correlations in angular projection. It had been thought that the broad redshift distribution typical of radio surveys could make detection of this clustering difficult, as the majority of close pairs of sources are widely separated in the line-of-sight direction, diluting any genuine clustering signal. Indeed, no evidence for clustering was found by Webster (1976) in the angular power spectrum of the 4C and GB radio surveys, or by Masson (1979) in the angular correlation function of the 6C survey.

Marginal detections of angular clustering were subsequently made in the Green Bank 4.85 GHz Sky Survey at $S > 35$ mJy (Kooiman, Burns & Klypin, 1995) and the Parkes–MIT–NRAO survey (Loan, Wall & Lahav, 1997) at $S > 50$ mJy, as well as confirmation of the spatial correlation of bright radio sources (e.g. Peacock & Nicholson, 1991), suggesting that large enough surveys with faint enough flux limits would uncover this signal. The FIRST survey provided such improvements, allowing Cress et al. (1996) to use a catalogue of $\sim 138,000$ sources to measure $w(\theta)$ with high signal-to-noise, inferring a correlation length of $r_0 \sim 10\ h^{-1}$Mpc, confirming that even this faint radio population was more strongly clustered than an optically selected sample ($r_0 \sim 5\ h^{-1}$Mpc; Peebles, 1980). Blake & Wall (2002a,b) performed a similar analysis of the NVSS and full FIRST catalogues, acheiving unprecedented signal-to-noise and tackling the unique problem of radio sources resolved into multiple components, resulting in a second, steeper power law form of $w(\theta)$ at small separations ($\theta \lesssim 0.2°$). Wilman et al. (2003) showed the importance of these large radio surveys, being able only to place an upper limit on the cosmological clustering of radio sources in the Bootes Deep Field (5.3 square

degrees), even as deep as $S_{1.4} = 0.2$ mJy, with the signal being dominated by the size-distribution of multi-component radio galaxies.

Soon followed a comprehensive confirmation and discussion of these results in NVSS and FIRST by Overzier et al. (2003). Comparing with numerous non-radio-selected studies, they conclude, for example, that powerful FRII galaxies probe significantly more massive structures than radio galaxies of average power at $z \sim 1$, and that the similar clustering properties of these powerful sources and extremely red objects (EROs; Daddi et al., 2001) suggests they may be the same systems seen at different evolutionary stages.

While measuring the angular correlation function of SUMSS (at 843 MHz), Blake, Mauch & Sadler (2004) repeated the angular correlation function analysis of NVSS at higher frequency (1400 MHz) and compared with WENSS (325 MHz), finding no evidence for a dependence on observed frequency or flux density (transformed to 1.4 GHz equivalent using $S_\nu \propto \nu^{-0.8}$). Fitting only to the large scales ($\theta > 0.2°$), where the signal is cosmological in origin, they consistently find a slope $\alpha > 1$ (1.05 for NVSS, 1.2 for SUMSS/WENSS), in contrast with the canonical value of 0.8 found for normal galaxies, and indeed reported in the previous NVSS analyses. This steeper slope agreed with that found in the FIRST survey by Cress et al. (1996), where the effect of multi-component sources was mitigated by collapsing pairs within $72''$ of one another, and is not surprising given the known dependency of α on galaxy type; its steepest value is found in elliptical galaxies such as radio AGN (Norberg et al., 2002).

In order to measure the spatial clustering properties of these radio galaxies directly, precise distances to the sources are required, necessitating the use of corresponding deep optical data, whereby host galaxies can be identified and spectroscopic redshifts obtained. This and the faintness of many radio galaxies in the optical somewhat negates the advantages of big surveys like NVSS and FIRST. As such, early studies of the three-dimensional clustering of radio galaxies were limited by small sample sizes. Magliocchetti et al. (2004) used a sample of 820 objects (at $z < 0.3$, with $S_{1.4} > 1$ mJy and $b_J < 19.45$) taken from the FIRST and 2 degree Field (2dF) Galaxy Redshift surveys. They found a real-space correlation length for the 526 AGN-fuelled sources in this sample of $r_0 = 7.6 \pm 0.8 \, h^{-1}$Mpc, while Brand et al. (2005) found $r_0 = 6.1 \pm 1.1 \, h^{-1}$Mpc for a similar sample in the Texas-Oxford NVSS Structure (TONS) survey ($z < 0.5$, $S_{1.4} > 3$ mJy, $R \lesssim 19.5$). These results confirm the clustering strength of radio galaxies as somewhere between that of normal galaxies and rich clusters of galaxies, although the relatively bright optical limits biases these samples to low redshift (and low-luminosity) radio galaxies. It is also worth noting that the mixed populations observed (low-redshift star-forming galaxies, and high-redshift AGN) undoubtedly confuse the interpretation of their clustering.

The short supply of radio sources with known redshifts has been partially overcome by cross-correlation radio-loud subsamples of optical surveys with the parent catalogues (see Chapter 5 for use of this tactic). For example, Wake et al. (2008) examine the radio clustering through the spatial cross-correlation of a radio-loud sample of luminous red galaxies (LRGs) in the 2SLAQ LRG survey at $z \simeq 0.55$, and Fine et al. (2011) extend a similar analysis to LRG samples from SDSS and AAΩ at $z \simeq 0.35$ and 0.68, respectively. Both studies find a radio-loud sample to be more strongly clustered than the broader LRG population, estimating a halo mass of 9–10×10^{13} h^{-1}Mpc, considerably more massive than optical AGN/QSOs.

1.4   Aims of this Thesis

In this thesis, I explore the clustering strength of radio sources, probing fainter and/or more distance samples, which have eluded prior investigation due to the various technological, observational, and statistical limitations of the surveys. Experience of directly observing powerful radio sources with a long-slit optical spectrograph (Chapter 2) has served to emphasize the importance of spectroscopic redshift surveys in producing large numbers of accurate redshifts to make any significant impact on adding a third dimension to ever-increasing radio catalogues (cf. Section 1.3.4). The Galaxy and Mass Assembly (GAMA) is one such redshift survey and is used in Chapter 3 to directly calculate the spatial clustering of a sample of FIRST radio sources at $S > 1$ mJy at $z \simeq 0.34$. Subsequent chapters look to improving on this work, first making use of photometric redshifts to add to the proportion of radio sources for which we can attribute optical counterparts, before also employing the latest model redshift distributions to infer the clustering of those radio sources without optical counterparts out to $z \simeq 1.54$ (Chapter 4). The disparity between radio sources and corresponding redshifts (spectroscopic or otherwise) can be addressed by moving to deeper, narrower surveys with a smaller number of radio sources, but also deep enough optical photometry to successfully cross-match them more effectively. The statistical sacrifice in terms of sample size is mitigated well by cross-correlating the sparse, strongly clustered radio sources with a densely-populated, infrared-selected survey with intermediate clustering strength (Chapter 5). This method extends the redshift of the most distance sample to $z \simeq 2.2$, and also maintains sufficient signal-to-noise to use limits on radio power to observe the large increase in clustering (and, to a degree, its evolution) of an AGN-only sample. This moves beyond the $z \approx 1$ regime which has been the limit of previous observations of the bias of radio AGN.

Chapter 2 describes the observations carried out with the ESO New Technology Telescope (NTT)

in Chile, with the aim of obtaining optical spectra to detect emission lines and find redshifts for a sample of $S > 100$ mJy radio sources in the *Herschel* 9h and 15h fields. In Chapter 3, the GAMA survey is cross-matched to gives a sample of FIRST radio sources with spectroscopic redshifts, and the redshift-space and projected correlation functions, and linear bias are calculated for this sample. Chapter 4 extends the area slightly and adds photometric redshifts from SDSS/UKIDSS to increase the number of cross-matched radio sources. The angular correlation function of several samples are calculated and the bias inferred by Limber inversion. Using a model redshift distribution from the Square Kilometre Array Design Study (SKADS) simulations of Wilman et al. (2008), the bias of the unmatched FIRST sources is found at $z \simeq 1.54$. Chapter 5 uses deeper VLA radio data ($S > 90$ μJy) and photometric redshifts from VIDEO/CFHTLS over an area of 1 square degree to compute the angular cross-correlation function of the radio sources with a $K_{\rm s} < 23.5$ sample from VIDEO. Disentangling this from the intrinsic clustering of the VIDEO sources, I find a radio source bias up to $z \simeq 2.2$, and also for separate samples of AGN, with star-forming galaxies removed. Chapter 6 summarizes the conclusions drawn over the course of the work described, with mention given to the relevance to galaxy evolution, possible applications and intended further work on the subject.

2

Optical Spectroscopy of $S_{1.4} > 100$ mJy Radio Sources with the NTT

"Too many of our children

Haven't known a true dark night;

Will they ever see the beauty

We're losing to the light?"

Sarah Williams, *'Blinded by the Light'*

2.1 Introduction

The radial component of the position of a distant object is inferred from its cosmological redshift (Section 1.1.2), induced by the expansion of the Universe. The light observed from a distant galaxy appears to us at longer wavelengths than in the rest frame of that galaxy. The most accurate determination of the exact redshift, z, comes from directly observing the spectrum of an extragalactic source and measuring a consistent multiplicative shift, relative to the rest frame, of various emission (or absorption) features. The rest-frame wavelengths of these emission lines are known to a high degree of accuracy which, depending on the resolution of the observed spectra, can be conferred onto the measured spectroscopic redshifts.

This chapter describes the observations and subsequent data reduction to obtain redshifts from long-slit spectroscopy of a sample of optically-faint, radio-loud galaxies, using the New Technology Telescope (NTT) in Chile. These spectra were taken over the course of a 5-night run on the nights of the 20th–24th March, 2012.

2.1.1 Observing Proposal

The proposal to the European Southern Observatory (ESO), requesting to take the observations described in this chapter, was entitled "The optically faint radio source population in the *Herschel*-ATLAS–GAMA fields". The *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010) observes 550 square degrees at five infrared and sub-mm wavebands (100, 160, 250, 350, and 500 μm), which allows the determination of the full FIR spectral energy distribution (SED) for dust-obscured star formation at high redshift. The Galaxy And Mass Assembly (GAMA; Driver et al., 2011) is a spectroscopic survey covering 144 square degrees in three equatorial fields, overlapping with three H-ATLAS fields, centred at 9h, 12h, and 14h30m in right ascension, which reaches an optical limit of $r \sim 19.8$. As such, the proposed observations were to individually target radio-loud sources fainter in the optical than the GAMA limit, in order to obtain redshifts for a complete sample of radio AGN, and enable a statistically significant study of their dust properties with H-ATLAS. FIR observations of radio sources can also be used to separate star-forming galaxies from AGN, which will enable clustering studies to distinguish the bias of radio-emitting AGN from that of the more weakly clustered normal galaxies.

Current semi-analytic models invoke AGN-driven feedback to prevent the overproduction of stars in the most massive galaxies (e.g. Croton et al., 2006), separating this into "quasar-mode" and "radio-mode" feedback, which can be interpreted as the efficient accretion of cold gas (such as from mergers), and the less efficient accretion of warm gas from the hot interstellar medium (e.g. Hardcastle, Evans & Croston, 2007). However, *Herschel* provides the previously missing picture of obscured star formation in these galaxies. The Submillimetre Common-User Bolometer Array (SCUBA) instrument on the James Clerk Maxwell Telescope (JCMT) in Hawaii has been used to establish a link between star formation and some classes of AGN (e.g. Archibald et al., 2001; Priddey et al., 2003; Rawlings et al., 2004), but the five-band H-ATLAS observations fully map the galaxy SEDs at FIR wavelengths, giving far more accurate FIR luminosities, and covering the peak of dust emission at $z > 1$. This allows the measurement of dust temperature, dust emissivity, and star formation rate in the AGN host galaxies over cosmic time.

The immediate aims of the proposed observations were to obtain spectroscopic redshifts for all radio-loud AGN, with radio flux density $S_{1.4} > 100$ mJy but optically fainter than the GAMA limit of $r \sim 19.8$, in the equatorial H-ATLAS/GAMA fields. Hardcastle et al. (2010) have shown that more powerful radio sources have higher FIR emission and therefore higher star formation rates (although this is in a different regime to the FIR–radio correlation discussed in Section 1.3.2). This study was limited, by lacking spectroscopy and insufficient depth, to $z < 0.8$, while the spectra obtained using the NTT will target more distant AGN, reaching epochs when the Universe was most active, when the most massive galaxies are thought to go through their major episodes of star formation. To obtain photometric redshifts for these distant, optically-faint AGN is only possible with the deepest degree-scale fields, where there are very few powerful radio AGN (~ 6 per square degree; e.g. Seymour et al., 2011).

The broader science goals of the proposal are beyond the scope of this chapter, but can be summarized as follows:

- Measuring the evolution of obscured star formation as a function of redshift and radio luminosity for powerful radio-loud sources. The redshift evolution of the FIR luminosity of similarly powerful AGN can be probed with an estimated 40 sources at $z > 2$ and $L_{1.4} > 10^{26.5}$ W Hz^{-1}, while an expected 30 sources at $1.2 < z < 2.2$, spanning two decades in radio luminosity, can be used to investigate the level of star formation in the host galaxies of radio-loud AGN, and the connection to accretion activity.

- Classifying high-excitation and low-excitation radio galaxies (HERGs and LERGs). This classification is thought to be linked to accretion mode (e.g. Hardcastle, Evans & Croston, 2007), and can be used to determine whether star-formation is enhanced in cold-mode accretors, as would be expected if such sources are triggered by major galaxy mergers.

- Investigating the link between FIR luminosity and radio source size, expected for coeval star formation and AGN activity (e.g. Willott et al., 2002). The FIR luminosity is related to the star formation rate, while the radio source size can be used to estimate the time since triggering of the radio source. Given a typical expansion velocity of radio jets of $\sim 0.1c$ (Owsianik & Conway, 1998), and a radio survey resolution of $\sim 5''$, the minimum age at which extended sources are resolved is ~ 10 Myr since triggering. Extended sources expected at arcminute angular scales can extend this investigation to ~ 100 Myr into the radio AGN phase.

- Compare the cold dust masses of the targeted radio sources with radio-loud and radio-quiet

quasars from SDSS/2SLAQ surveys over these fields (Serjeant et al., 2010; Bonfield et al., 2011), in order to test unified schemes of AGN.

2.1.2 New Technology Telescope

The NTT, located on the ESO site at La Silla in Chile, saw first light in 1989 (subsequently being upgraded in 1997). It is a 3.6 m Ritchey-Chrétien telescope on an alt–az mounting, with two Nasmyth foci, hosting two instruments: SofI, a large-field infrared (0.9–2.5 μm) spectrograph and imaging camera ($R = 600$–2200); and the second version of the ESO Faint Object Spectrograph and Camera (EFOSC2), an optical camera and low-resolution spectrograph, with multi-object spectroscopy, (spectro)polarimetry, and coronography modes. The NTT boasts the first use of full active optics, with both primary and secondary mirrors controlled by actuators to preserve the shape of the optics, ensuring observations reach the ambient seeing (although not improving upon it, as in the case of full adaptive optics). This feature, and the dome design itself, with optimized air flow and carefully regulated temperature throughout, gave the telescope its name, and helped to perfect the design of the four Unit Telescopes (UTs) of the Very Large Telescope (VLT), which saw first light at the Cerro Paranal site about ten years after the NTT.

The EFOSC2 instrument has a wavelength range of operation of 305–1100 nm, and a field of view of $4.1' \times 4.1'$. The CCD detector is 2048×2048 pixels, each pixel being 15×15 μm in physical size, representing $0.12'' \times 0.12''$ on the sky. For long-slit spectroscopy, the spectral resolution depends on the slit width and grism used; 16 grisms are available, with grating densities between 100 and 1070 mm^{-1}, tailored to low- and high-resolution spectra, respectively, and each with their own optimal wavelength range.

2.2 Target Selection

The radio sources observed were drawn from the FIRST survey (see Section 1.3.4), which was carried out at 1.4 GHz, covering a large area, including the three H-ATLAS/GAMA fields, with a relatively high resolution of $5.4''$ required to accurately locate the position of the optically faint host galaxy. The complicated morphology of radio sources makes this task especially difficult, and I was forced to visually inspect the radio images of each target in advance.

Priority was given to flux density, with the brightest sources in the radio being observed first. Next priority was in observing compact, isolated sources, for which the host galaxy position is most likely

to be coincident with the radio emission. This compactness was defined, somewhat arbitrarily, using the area of the elliptical Gaussian models used for the sources given in the form of major and minor axes (maj $\times$ min $<$ $(5.4'')^2$ defining an unresolved source, smaller than the beam size). Many of these apparently compact sources, however, are part of a multiple-component source, extending 10s of arcseconds across (100s of kpc at $z \gtrsim 1$). To account for this, groups of detections within $100''$ of one another were matched and their FIRST images inspected to assess any such double or triple sources and, where possible, deduce a host galaxy position. In some cases, there was clear radio emission from the core, and in others we cross-matched with SDSS to confirm the presence of an optical counterpart. Many such multiple sources were morphologically ambiguous, and an observation was not attempted. Figure 2.1 shows radio images of two of our targets taken from the FIRST survey, highlighting the range between compact and highly extended radio sources.

The positions on the sky of the three H-ATLAS/GAMA fields at the time of observing were such that the minimum airmass was achieved observing the 9h field at the start of the night, and the 15h field at the end of the night, with the 12 hr field being optimal in the middle. Given the number of targets, it was decided to favour a more complete sample of two out of three fields, at the expense of the 12h field. As a result, approximately half of each night was devoted to the field closer to the zenith at the time.

The target criteria, and the order with which observations proceeded, with the above considerations in mind, is summarized as follows:

1. FIRST sources within H-ATLAS/GAMA 9-hour and 15-hour fields ($\sim$ 144 square degrees).

2. Detections with an integrated flux-density above 100 mJy (or multiple sources within $100''$ whose total integrated flux-densities exceed this limit).

3. Sources without spectroscopic redshifts from GAMA (i.e. $r \gtrsim 19.8$) or observations in a previous observing run at the William Herschel Telescope (WHT) in La Palma (February 2011).

4. Radio images for multiple-component sources inspected by eye, prior to observation; objects with ambiguous morphologies, or estimated host galaxy positions unconfirmed by SDSS were discarded.

5. Targets observed in descending order of flux density, with the extended sources being observed after all of the compact sources, and the 9h field being observed for the first half of each night before switching to the 15h field.
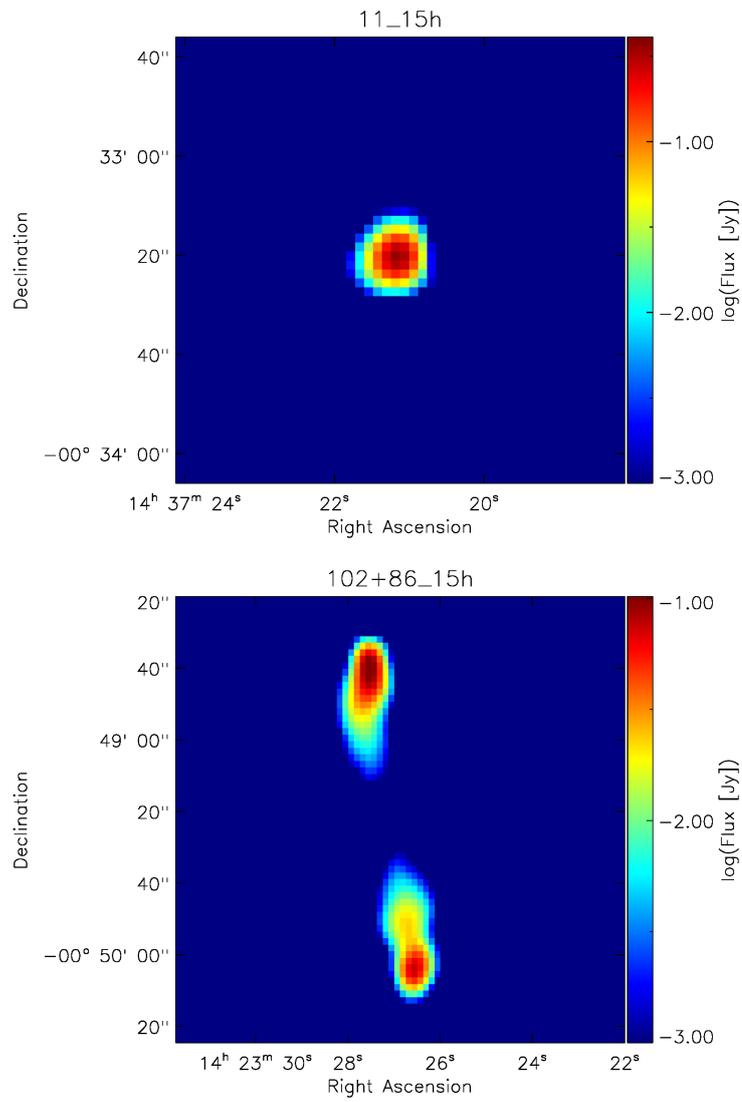
FIGURE 2.1: Example FIRST images of two targets: a compact source ('11_15h', top panel) at $z = 0.966$, and a large double-lobed source ('102+86_15h', bottom panel) at $z = 0.793$. The positions of the two discrete radio sources in '102+86_15h' are $82''$ apart, equating to a physical separation of ~ 600 kpc.

2.3 Observations

The EFOSC2 CCD has 2048×2048 pixels, but these pixels can be binned in a 2×1 or 2×2 scheme, improving the signal-to-noise ratio at the expense of resolution in the spatial and/or spectral direction. For the purposes of these observations, maximizing the signal-to-noise of optically-faint radio sources took priority over resolution in both directions and the 2×2 binning was used, yielding 1024×1024 pixel images with a pixel scale of 0.24″/px. This binning also expedites the readout time, reducing the overheads of each exposure.

Spectral resolution depends on both the CCD binning and the grism used to disperse the incident light, with a finer grating spreading the different wavelengths further apart, giving higher resolution over a narrower range of wavelengths. Again, with signal-to-noise being the priority, as well as detecting a range of emission lines at a range of redshifts, grism #13 was used, with a range of 3685–9315 Å. This broad wavelength range, combined with the pixel binning, results in a dispersion of 5.44 Å/px (dependent on slit width), and covers many common emission lines expected in AGN at moderate redshift, with the Ly α line at 1216 Å in the rest frame being observable at high redshift ($z > 2.03$).

2.3.1 Bias Frames

A digital signal is read out from a CCD by charge transfer, between neighbouring pixels, to a single readout pixel. Bias noise is a background level of noise induced by this readout process, resulting in pixels (especially those near to the readout pixel) retaining some charge and appearing 'brighter' than their neighbours. This can cause a gradient across the CCD, but can be corrected for by taking several zero second bias exposures, with any electronic bias being subtracted from all subsequent images using the average bias frame. At the start of each night of observations, 20 bias frames were taken.

2.3.2 Flat Fields

Correcting for any uneven illumination of the detector (e.g. as a result of vignetting, due to the optical properties of the detector) requires calibration exposures to be taken of a 'flat field'. This generally must be uniform, diffuse and out of focus, such that no structure is spatially resolved in the images. Likewise, for spectroscopic flats, the spectrum of the flat field should be featureless. There are two types of flat fields commonly used, both of which were taken at the start and end of each night of observations:

- Dome flats are obtained by observing an illuminated white area within the telescope dome. A quartz lamp is used, producing a black-body spectrum, free of emission lines. This allows a

continuum to be fitted and divided from the flat to produce a normalized flat-field image for scientific purposes. This method suffers, however, from a paucity of counts at the blue end of the spectrum, which increases the noise in this region of the normalized flats.

- Sky flats are obtained by observing the sky itself during twilight, which gives a spectrum far richer in blue photons, due to Rayleigh scattering. The problem with sky flats is that they are dominated by strong line features, making them unsuitable for flat-fielding in the spectral direction. Instead, they are useful for constructing the slit function, tracing the illumination of the slit in the spatial direction.

A set of dome flats ($5{\times}15$ s and $5{\times}62$ s exposures, with a $5''$ and $2''$ slit, respectively) were taken before and after the observations each night, and several twilight sky flats were taken when possible (at the start or end of the night).

2.3.3 Wavelength Calibration

Helium and Argon (He-Ar) lamps are used for wavelength calibration of the slit spectra. These produce a number of emission lines across the full wavelength range of the grism being used, which can be identified in the calibration images to determine the non-uniform mapping between pixels lines in the dispersion direction along the CCD, and the corresponding observed wavelength. These calibration spectra were taken in exposures of 2.5 s, with a $2''$ and $5''$ slit, at the start and end of each night of observations.

2.3.4 Flux Calibration

Observing a 'standard star' with an accurately-known spectrum provides a flux calibration, relating the CCD pixel counts to a flux value. The standard used for these observations was a 13th-magnitude B star, GD108 ($\alpha = $ 10h:00m:47.33s, $\delta = -$07d:33m:31.2s), the spectrum of which is shown in Figure 2.2. Immediately prior to beginning scientific observations each night, three ~ 50 s exposures were taken of the standard star.

2.3.5 Target Spectra

Given the optically faint nature of our targets, a nearby bright source is required in order to align the slit, acquiring the target source. As both sources must lie along the slit of the spectrograph, the guide star chosen for each target dictates the rotation offset angle (ROA; clockwise from horizontal).
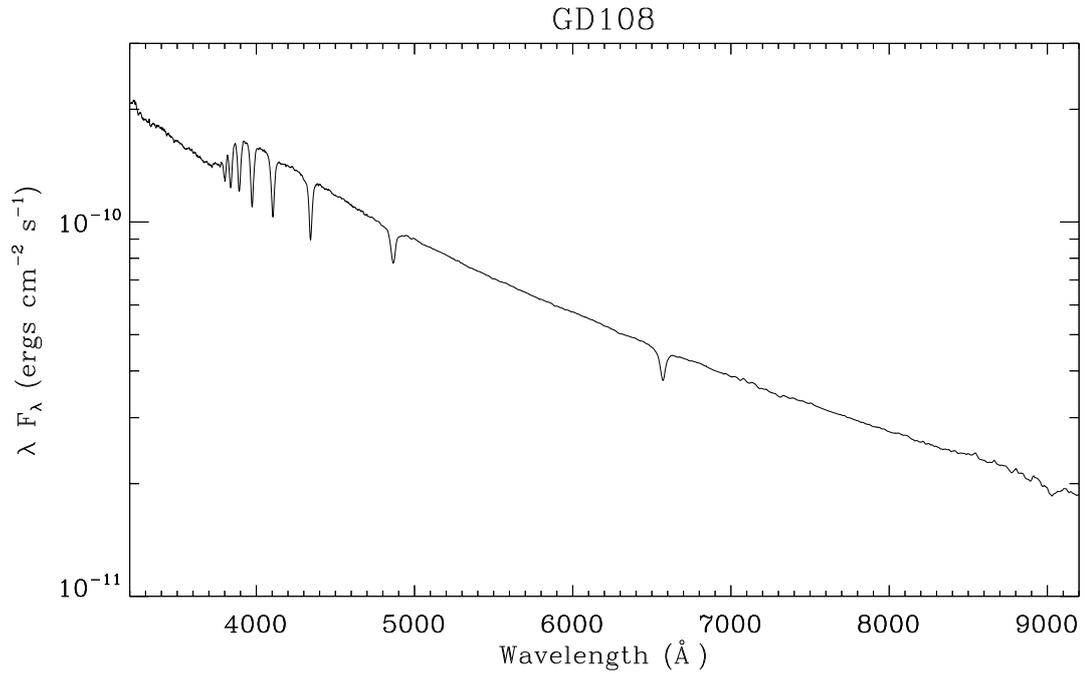
FIGURE 2.2: Optical spectrum of standard star GD108, from Oke (1990).
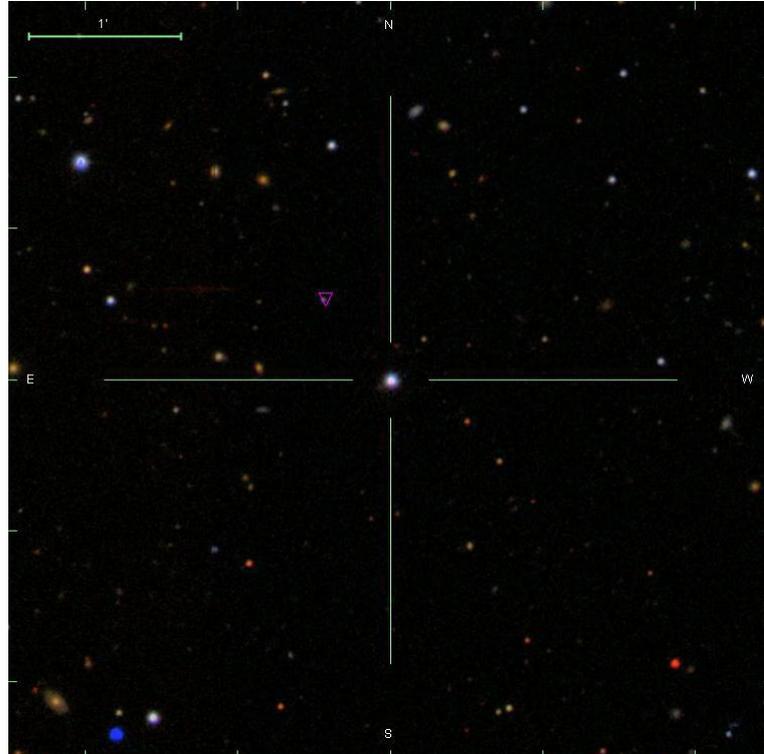
FIGURE 2.3: Example finding chart from the SDSS DR7 images, centred on a magnitude 15.4 star. The triangle marks the location of the radio-loud AGN target (designated '11_15h') at a distance of 41″ to the North-East. The instrument is rotated such that both sources are aligned along the slit.

For convenience, the guide stars were placed roughly in the centre of the images (Column ~ 550), requiring that they be less than $\sim 2'$ from the target (although the position can be shifted in order to accommodate slightly greater separations). As there may be many such bright sources within practical range of the target, in the case of double/triple sources, where there may be some ambiguity in the exact location of the host galaxy, the guide stars were chosen such that the slit was closely aligned to the position angle of the radio sources. An example finding chart is shown in Figure 2.3, with the guide star in the centre of the image and the target coordinates indicated by a purple triangle, $38.6°$ East of North from the guide star (ROA=$128.6°$). In this case, an optical source is visible in the SDSS image, potentially corresponding to the host galaxy of the radio source.

Targets were observed in up to 3×600 s exposures, allowing the resulting images to be averaged to remove spurious counts due to incident cosmic rays. This also allowed time to inspect the first image, identifying any problems with the observations, or strong emission lines, before deciding whether to proceed with the second and third exposures. Several of the sources for which redshifts were obtained required only one or two exposures, saving time in moving onto the next target. As a results, 17 targets were observed (8+9, in the 9h and 15h fields) on each of the 5 nights on the telescope, corresponding to 79 distinct host galaxies being targeted.

2.4 Data Reduction

The reduction of the raw images from the telescope was carried out using the Image Reduction and Analysis Facility (IRAF) software. Figure 2.4 shows an example of the raw and processed images of '11_15h' (see Figure 2.3), and the resulting extracted spectrum, with emission lines labeled.

The reduction process can be broken down into the following steps (and IRAF commands):

- Bias frames are combined (using the `imcombine` command) to create a single high signal-to-noise master bias, which is subtracted from all images (using `imarith`).

- Dome flats are combined to create a master lamp flat. Using the `response` command, a smooth function (high order polynomial) is fitted interactively to the lamp spectrum and the response function image is formed by dividing by this fit.

- Twilight sky flats are combined to create a master sky flat, which is then divided by the lamp response function. The `illumination` command is then used to divide this image into ~ 10 bins in the dispersion direction, fitting a slit profile for each bin before interpolating between the
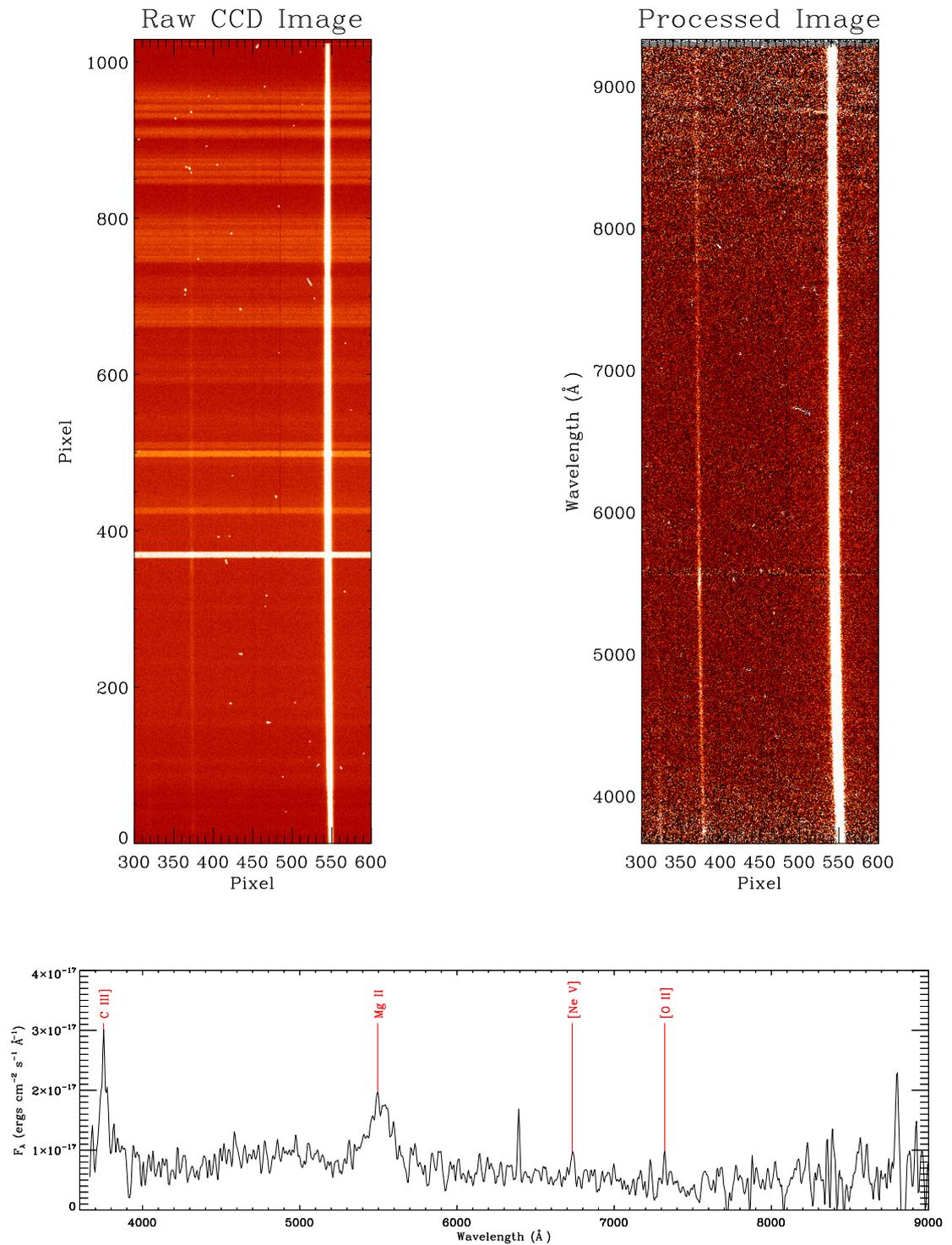
FIGURE 2.4: Example sections of a raw exposure image (top left panel) and corresponding processed, calibrated final image (top right panel), showing the guide star at pixel column ~ 550 and the faint target galaxy at pixel column ~ 375. The bottom panel shows the extracted spectrum of the target galaxy with identified emission lines redshifted by a factor of $(1 + z) = 1.966$.

average profiles for each bin. The output slit profile image is then multiplied by the response function image to give a final flat field calibration image. All other images are then divided by this flat field calibration.

- Emission lines in the He-Ar lamp spectra are identified by the `identify` and `reidentify` commands, with a few strong features being user-identified before others are identified from a line list. The `fitcoords` command then fits a coordinate function mapping pixel coordinates to wavelength. These commands are repeated to apply the same coordinate mapping to the standard star spectrum.

- The `transform` command inverts the coordinate function and transforms the coordinates of all of the bias-subtracted, flat-fielded images from pixel line values to wavelengths in Å.

- The flux calibration is carried out by subtracting the background from the combined standard star images (with the `background` command), and fitting the observed spectrum to calibration fluxes (see Figure 2.2) using the `standard` command. This is done interactively, avoiding fitting across absorption features, allowing the `sensfunc` command to use the fits to accurately determine the detector sensitivity function, relating the pixel counts to a flux value. The target source spectra are then similarly background-subtracted before `fluxcalib` is used to apply the flux calibration to the pixel counts, yielding new pixel values in units of $\mathrm{ergs\,cm^{-2}\,s^{-1}\,\mathring{A}^{-1}}$.

- Finally, multiple exposures (2 or 3, depending on the presence of strong emission lines in the first image, which dictated whether further observations were carried out) for each target are combined (taking the average pixel values) to a single image per source, and apertures are defined for each of the images using the `apsum` command to extract a spectrum.

2.5 Redshifts

Powerful radio galaxies tend to exhibit strong emission lines (e.g. Hine & Longair, 1979), despite often having relatively faint optical continuum emission. The optical spectra of AGN generally are dominated by these strong emission lines, arising from hot gas surrounding the central black hole, with these features often broadened by the large orbital velocity in this region (see Section 1.3.3). Francis et al. (1991) obtained a mean QSO spectrum from over 700 QSOs in the Large Bright Quasar Survey, shown in Figure 2.5, which highlights many of the common emission lines in the rest frame optical and
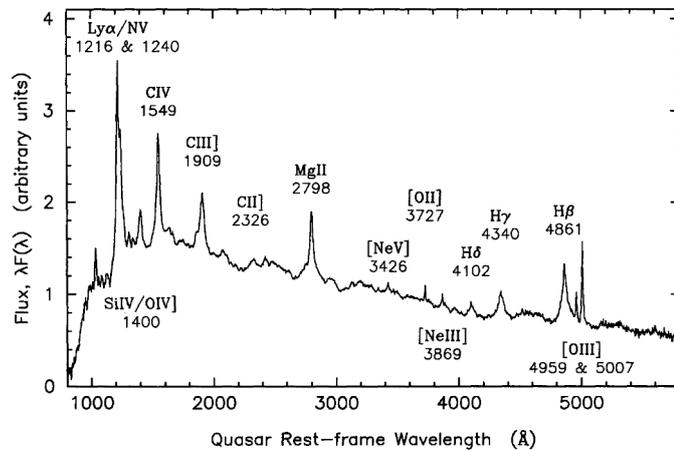
FIGURE 2.5: A composite spectrum (with principal features identified) of the rest-frame ultraviolet and optical region from 718 individual spectra of high luminosity quasars in the Large Bright Quasar Survey (Francis et al., 1991).

Emission Line	Rest-frame wavelength (Å)
O VI	1035
Ly α	1216
N V	1240
Si IV	1400
C IV	1549
He II	1640
C III]	1909
C II]	2326
[Ne IV]	2424
Mg II	2798
[Ne V]	3346, 3426
[O II]	3727
[Ne III]	3869, 3967
H δ	4101
H γ	4340
He II	4686
H β	4861
[O III]	4959, 5007
H α	6563

TABLE 2.1: Common emission lines used to identify features in our galaxy spectra and therefore to determine their redshifts. Square brackets indicate forbidden ([ ]) and semi-forbidden (]) lines.

ultraviolet spectrum. A list of such emission lines found in the spectra of our targets is given in Table 2.1.

A single emission line is generally insufficient to identify the redshift of a galaxy as the origin of a given feature in isolation cannot be identified, and therefore its rest frame wavelength is unknown. However, the redshifting of the rest-frame spectrum (stretching by a factor of $1 + z$) preserves the wavelength *ratio* between features, allowing a pair of emission lines to be identified. Once the lines are identified, the redshift is given by $1 + z = (\lambda_{\mathrm{obs}}/\lambda_{\mathrm{rest}})$, from the observed wavelength λ_{obs}, and the rest frame wavelength λ_{rest} (given in Table 2.1). The example images and spectrum shown in Figure 2.4 show two strong emission lines observed at 3751 Å and 5495 Å, with a ratio of 1.465. Comparing this ratio with various pairs of lines above, the two lines were identified as being C III] 1909 and Mg II 2798, respectively, with a ratio of 1.466. Knowing the rest-frame wavelengths at which these emission lines were emitted, the redshift could be calculated at $z = 1.965$ (1.964 for the peak wavelength of the broader Mg II line). By way of added confirmation, this redshift placed the [Ne V] 3426 and [O II] 3727 lines at the location of two smaller peaks at longer wavelengths.

The reduced efficiency of the detector/grism combination at the red end of the spectrum, combined with the complicated atmospheric absorption (cf. striped appearance of the raw CCD image in Figure 2.4), contrives to produce a very noisy spectrum at longer wavelengths, making these emission lines harder to detect as redshift increases. For a given radio flux limit, the strength of the optical emission diminishes with increasing distance, as well as more lines becoming redshifted into the noisier region of our observed spectra. At $z \gtrsim 2$, however, the Ly α line in the rest frame ultraviolet region is redshifted within the shorter wavelength range of our optical spectra. Figure 2.6 shows a histogram of the 30 redshifts (out of 79 targets), with a mean of $z = 1.2$. The 6 sources with $z > 2$ were detected with Ly α emission.

Table 2.2 shows the redshift and line identifications of the successful observations. The line width is estimated by the full width at half maximum (FWHM) of the best-fitting Gaussian profile for each line. The width is given as a velocity (Δv), derived from the the FWHM in Å ($\Delta\lambda_{\mathrm{obs}}$):

$$\Delta v = c\frac{\Delta\lambda_{\mathrm{obs}}}{\lambda_{\mathrm{obs}}}. \tag{2.1}$$

The observed sources are also broadly categorized, based on their spectra, as quasars (Q) or narrow-line radio galaxies (G). Historically, this categorization was made based on the visibility of their optical hosts, radio galaxies being clearly visible, while the quasar emission is dominated by an unresolved nucleus. A physical distinction can now be made based on the emission line properties of the sources:
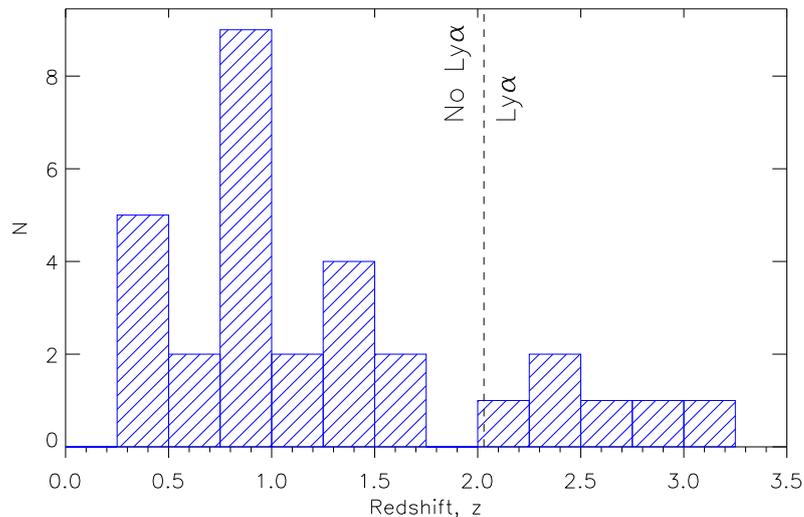
FIGURE 2.6: Histogram of the redshifts extracted from NTT spectra. The vertical line marks the lower redshift limit at which the rest-frame Lyα line enters the observable range.

classical radio galaxies have spectra with narrow permitted lines (FWHM < 2000 km s^{-1}) and strong forbidden lines; quasars have broad permitted lines and weak forbidden lines. All 30 spectra are shown in Figures 2.7, with emission lines and redshifts indicated.

2.6 Conclusions

A total of 79 host galaxies were targeted over the course of 5 nights of observations. Despite up to 1800 s of exposures for each object, and the 2×2 CCD binning for improved signal-to-noise, most of these were still too faint for emission lines to be detected, with 30 redshifts being retrieved from the data (38 per cent).

Of the 30 successful targets, 14 had SDSS/UKIDSS objects within 5$''$ with photometric redshifts estimated from 9 wavebands ($ugrizYJHK$). Figure 2.8 compares these with the spectroscopic redshifts presented here, showing that the photometric method has failed to correctly place the high-redshift sources. Even at $z < 0.8$, where the 1:1 correspondence between redshifts roughly holds, z_{phot} tends to err on the low side of z_{spec}. Sources categorized unambiguously as quasars or narrow-line RGs, are plotted in blue or red, respectively, with ambiguous sources plotted in green. Open circles represent the two targets whose host galaxy position was inferred from the positions of multiple radio components (24 such sources were targeted, with 6 redshifts found). Both of these photometric redshifts had large associated errors; one was $< 1\sigma$ from z_{spec}, and the other was clearly dubious, having an error larger

than the best fit redshift.

Due to the faint optical emission of the sources being targeted, and not aided by opting for a lower-resolution observing mode, some lines identified in the spectra are scarcely detected above the noise level of the image (and in the case of 27_9h, no robust redshift can be truly confirmed). As a result, many of the fainter emission lines presented must have this taken into account when interpreting the Gaussian profile fits performed in IRAF. Considering the observed spectra of these 30 objects, 12 are confidently identified as QSOs (40 per cent), 13 are narrow-line RGs, and the remainder are uncertain.

The broader science that can be studied using these (and similar) observations, outlined in Section 2.1.1, are beyond the scope of this chapter, but, in any case, these spectra will contribute to the ongoing building of a library of redshifts for radio sources in the H–ATLAS/GAMA fields. Furthermore, spectroscopic follow-up of LOFAR radio sources is one of the primary science drivers of the WEAVE multi-object spectrograph (Dalton et al., 2012) on the William Herschel Telescope in La Palma, and will benefit from the experience and knowledge of spectroscopic success rates from individual observing runs targeting radio AGN, such as presented here.

Source name	z	Line	λ_{rest} (Å)	λ_{obs} (Å)	FWHM (km s^{-1})	Galaxy or Quasar
4_9h	1.465	C IV	1549	3818	2000	G/Q
		He II	1640	4043	2000	
		C III]	1909	4700	2300	
		[Ne IV]	2424	5984	1200	
5_9h	2.155	Ly α	1216	3836	2200	G/Q
		N V	1240	3907	1200	
		C IV	1549	4902	1900	
		He II	1640	5179	1600	
		C II]	2326	7331	700	
7_9h	0.752	C II]	2326	4073	2000	G
		Mg II	2798	4902	1300	
		[O II]	3727	6533	1500	
		[Ne III]	3869	6780	2400	
		[O III]	4959	8680	1800	
		[O III]	5007	8776	1700	
9_9h	1.37	C IV	1549	3678	1200	Q
		He II	1640	3889	1200	
		C III]	1909	4528	2400	
		C II]	2326	5508	2000	
		Mg II	2798	6628	2900	
11_15h	0.966	C III]	2326	3751	4000	Q
		Mg II	2798	5505	>10000	
		[Ne V]	3426	6734	1900	
		[O II]	3727	7313	2100	

TABLE 2.2: Measurements obtained from 30 NTT spectra with line emission. The source names are our own, taken from the index of each source in our catalogue of potential targets, and the GAMA field in which they are located. "$a+b$" indicates two bright radio sources considered to be a double, while "$c+$" indicates a single bright source, with one or more fainter companions contributing to the targeted position. Observed wavelengths and line widths are estimated using the mean and FWHM of the best-fitting Gaussian to each line.

Source name	z	Line	λ_{rest} (Å)	λ_{obs} (Å)	FWHM (km s^{-1})	Galaxy or Quasar
14_15h	1.22	C III]	1909	4225	1900	Q
		C II]	2326	5177	>10000	
		Mg II	2798	6225	>10000	
19_9h	1.083	[Ne IV]?	2424	5059	5400?	G?
		[O II]	3727	7762	2000	
27_9h	2.75?	O VI?	1035	3876	2200	G?
		Ly α?	1216	4550	2000	
		Si IV?	1400	5257	2300	
		C IV?	1549	5834	2400	
		C III]?	1909	7153	1700	
30_9h	0.960	C II]	2326	4551	3700	G
		[Ne IV]	2424	4759	2800	
		[O II]	3727	7309	1400	
		[Ne III]	3869	7587	1800	
32_9h	0.403	[O II]	3727	5230	1800	G
		H β	4861	6820	500	
		[O III]	4959	6949	800	
		[O III]	5007	7026	1500	
37_9h	0.724	[O II]	3727	6434	1800	G
		He II	4686	8076	1600	
		[O III]	4959	8536	1800	
		[O III]	5007	8631	1800	
42_9h	2.60	O IV	1035	3724	2100	Q
		Ly α	1216	4385	2600	
		N V	1240	4461	2700	
		C IV	1549	5577	2700	
		C III]	1909	6852	1900	
		C II]	2326	8351	1400	
44_9h	1.72	O VI	1035	3809	4800	Q
		Si IV	1216	4210	7000	
		C III]	1240	5200	7600	
47_15h	0.84	C II]	2326	4292	>10000	Q
		Mg II	2798	5136	>10000	
		[O II]	3727	6846	2200	
		[Ne III]	3869	7136	2500	
		[Ne III]	3967	7298	1600	
49_15h	0.745	C II]	2326	4045	2100	Q
		Mg II	2798	4887	3400	
		[O II]	3727	6502	1100	
		[O III]	4959	8657	1100	
		[O III]	5007	8742	1100	
50_15h	0.339	[O II]	3727	4990	1300	G
		H β	4861	6492	1800	
		[O III]	4959	6640	1800	
		[O III]	5007	6709	1700	

TABLE 2.2: *continued*

Source name	z	Line	λ_{rest} (Å)	λ_{obs} (Å)	FWHM (km s^{-1})	Galaxy or Quasar
51+_15h	2.371	Ly α	1216	4099	2900	Q
		C IV	1549	5223	2400	
56_15h	0.300	[O II]	3727	4844	1700	G
		[O III]	4959	6434	1400	
		[O III]	5007	6500	1700	
		H α	4861	8536	2200	
64_9h	0.899	[Ne V]	3426	6514	2000	G
		[O II]	3727	7075	1400	
		[Ne III]	3869	7348	1200	
		[Ne III]	3967	7541	1600	
65_15h	2.399	Ly α	1216	4134	1900	G
		He II	1640	5574	1000	
67+_9h	0.843	Mg II	2798	5155	7700	Q
		[O II]	3727	6868	2000	
		[Ne III]	3869	7143	2500	
		[O III]	4959	9143	1000	
		[O III]	5007	9225	1100	
68_9h	0.908	[O II]	3727	7111	1600	G
		[Ne III]	3869	7379	1800	
70+80_9h	1.48	C III]	1909	4720	3000	Q
		Mg II	2798	6951	3600	
71_9h	0.431	[O II]	3727	5338	1200	G
		[O III]	4959	7094	1100	
		[O III]	5007	7167	1000	
73_15h	1.61	He II	1640	4268	1700	G/Q
		C III]	1909	4997	2300	
		C II]	2326	6055	2600	
		[Ne IV]	2424	6346	900	
		Mg II	2798	7335	2400	
75_9h	3.07	O VI	1035	4224	2300	Q
		Ly α	1216	4950	1900	
		N V	1240	4993	1800	
		Si IV	1400	5725	4300	
		C IV	1549	6308	6000	
76_9h	0.792	Mg II	2798	5008	3400	Q
		[O II]	3727	6683	2000	
		H δ	4101	7351	1800	
102+86_15h	0.793	C II]	2326	4175	1200	G
		[O II]	3727	6681	1300	
		He II	4658	8396	900	
		[O III]	4959	8892	1000	
		[O III]	5007	8974	1400	

TABLE 2.2: *continued*

Source name	z	Line	λ_{rest} (Å)	λ_{obs} (Å)	FWHM (km s^{-1})	Galaxy or Quasar
119+_15h	0.34	[O II]	3727	4994	2500	G
		[O III]	4959	6627	1000	
		[O III]	5007	6702	1100	
		H α	6563	8825	1800	
123+_15h	1.37	C III]	1909	4539	2600	G
		Mg II	2798	6639	2200	
		[Ne V]	3346	7898	1400	
		[Ne V]	3426	8129	1400	
		[O II]	3727	8839	1200	

TABLE 2.2: *continued*
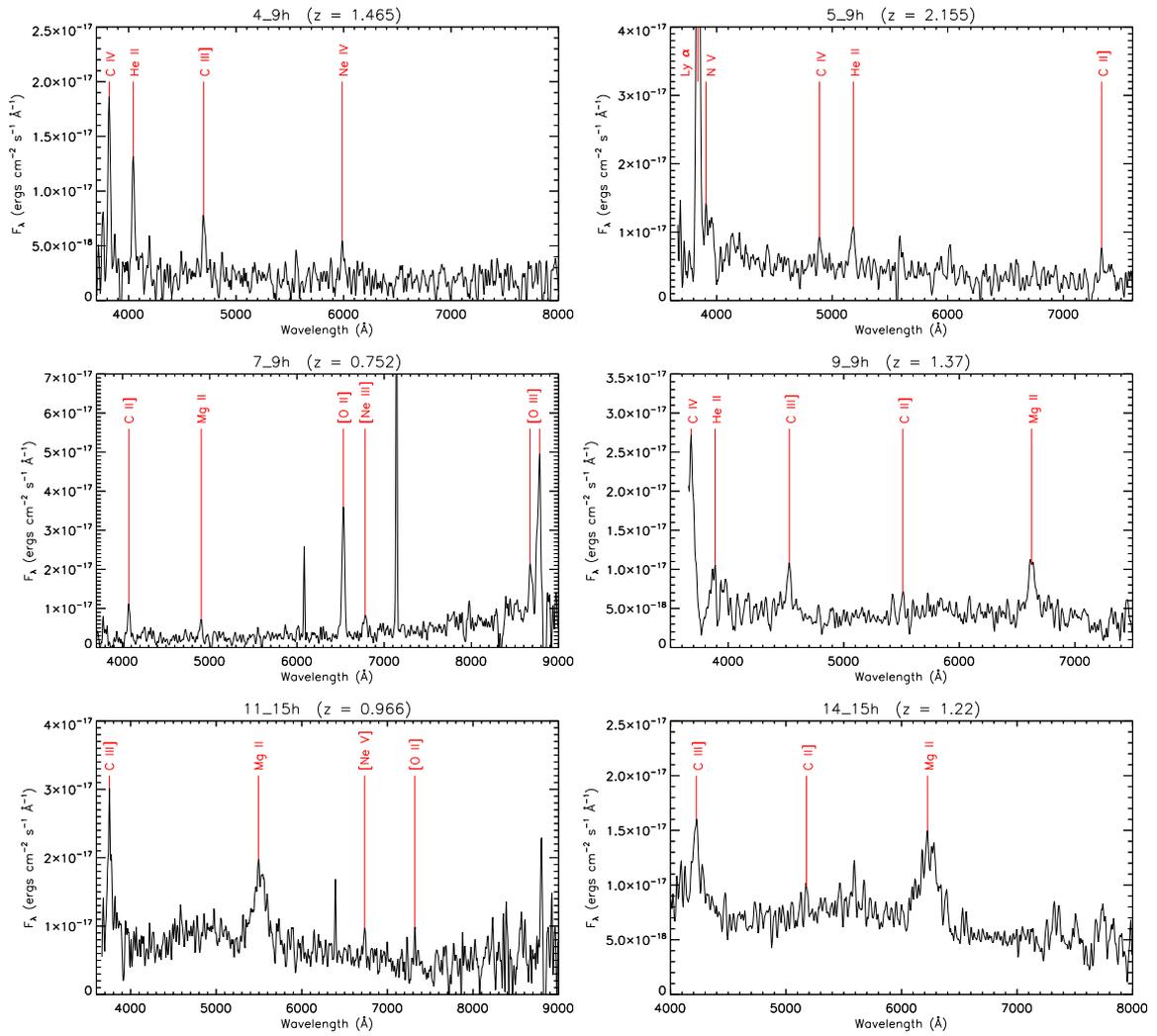
FIGURE 2.7: Observed spectra of 30 targets observed with EFOSC2 on the NTT for which redshifts were able to be determined, with emission lines (listed in Table 2.2) identified. Speculative line identifications are shown for 27_9h, which had faint continuum emission with no strongly-detected lines.
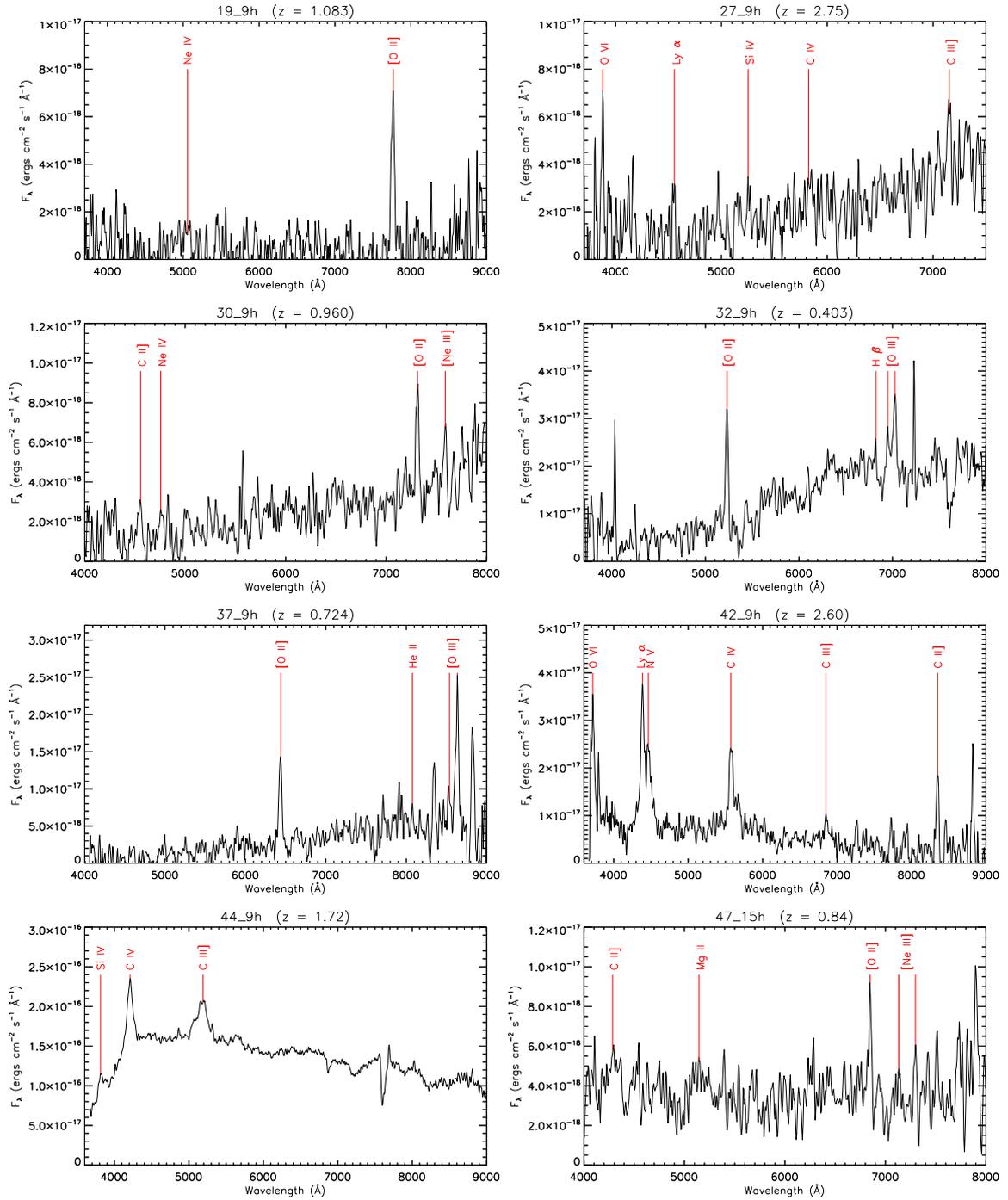
FIGURE 2.7: *continued*
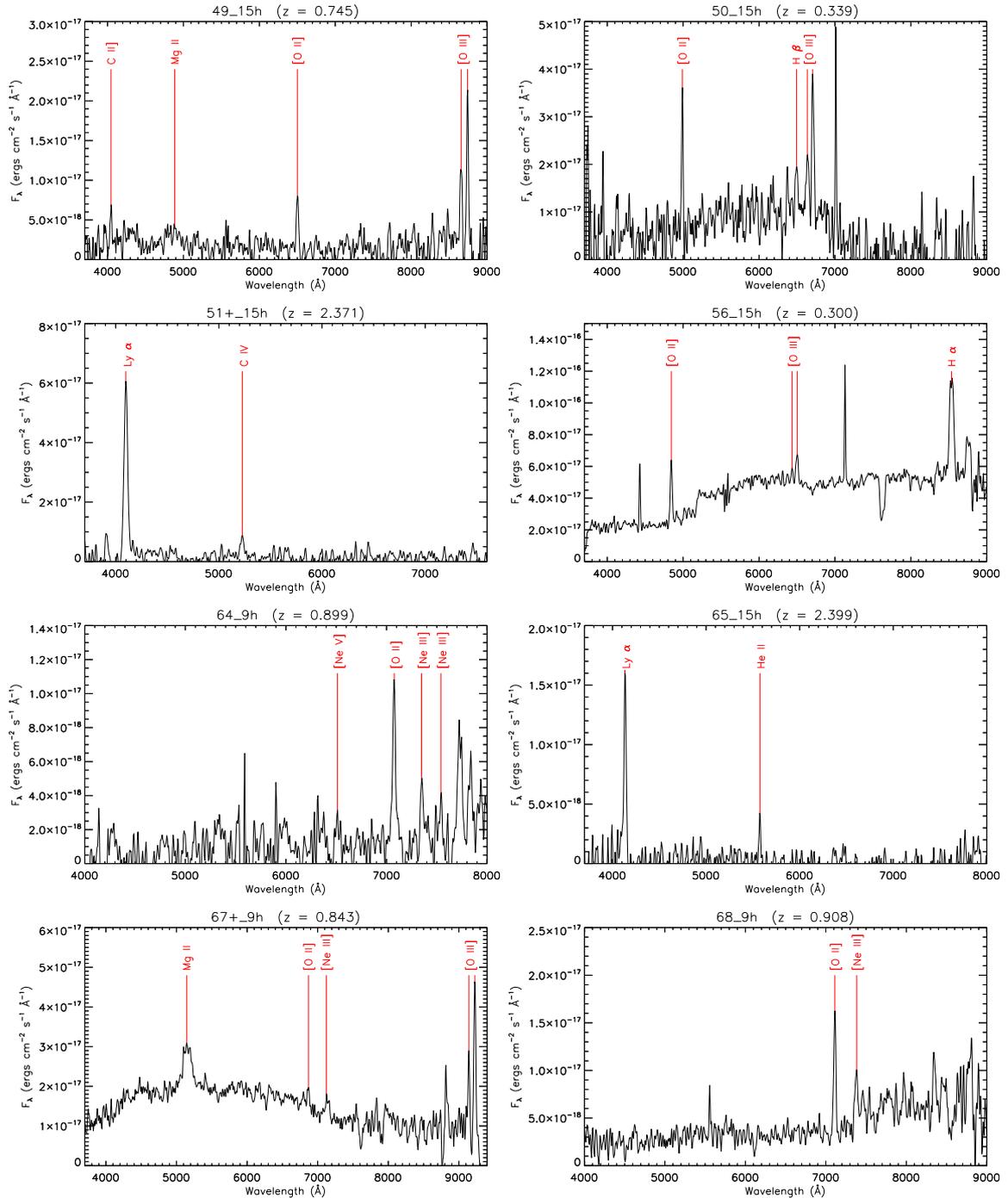
FIGURE 2.7: *continued*
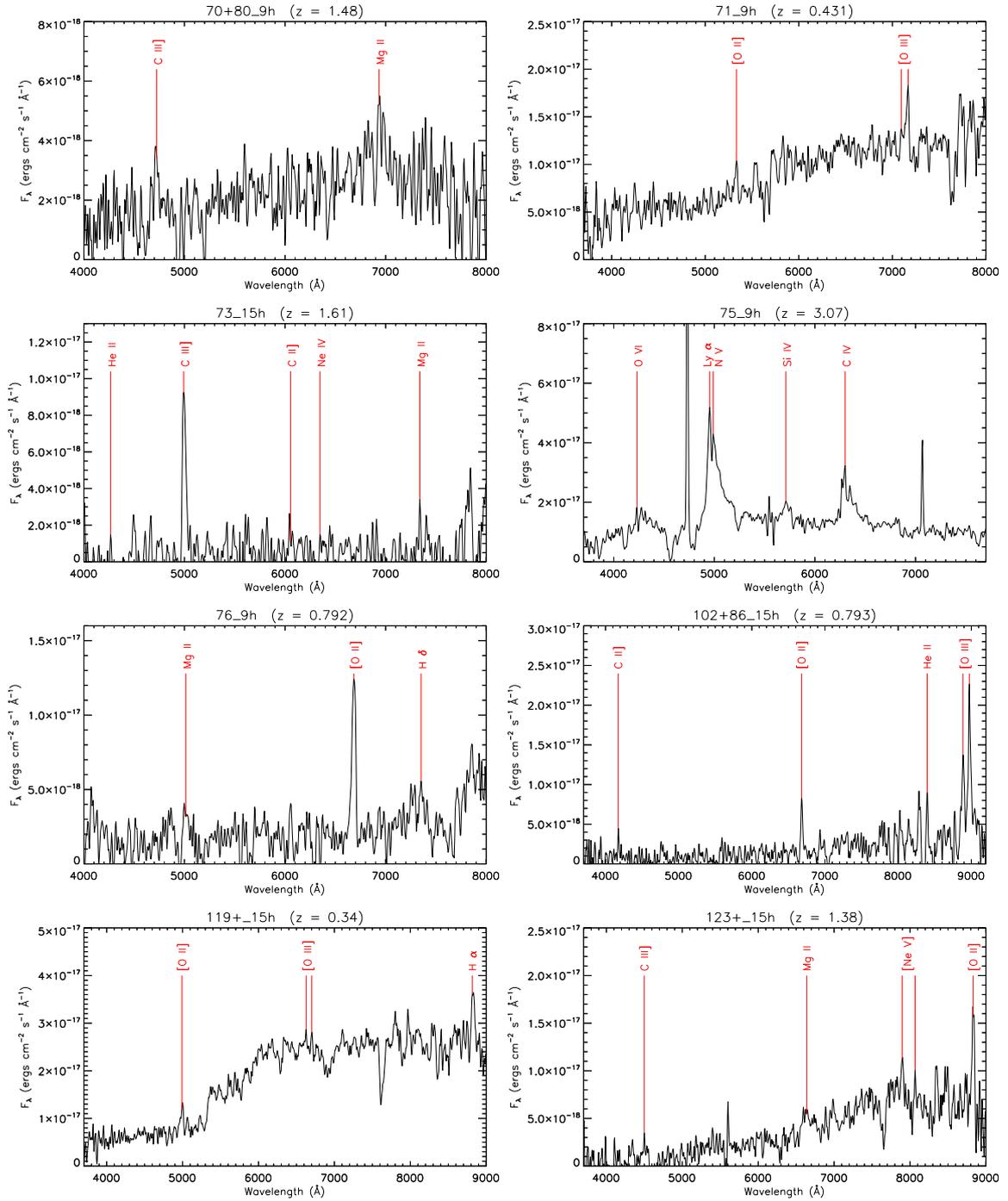
FIGURE 2.7: *continued*
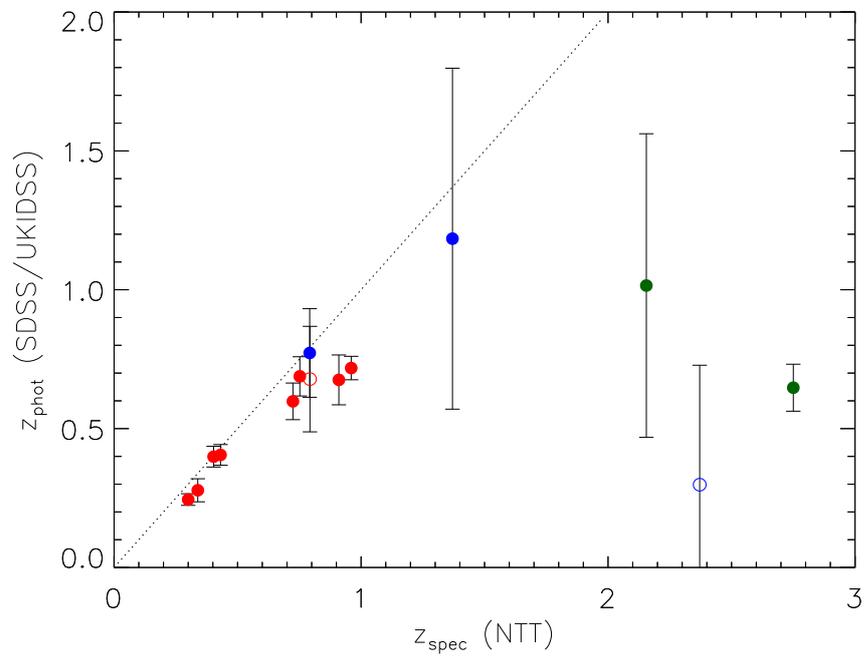
FIGURE 2.8: A comparison of our observed spectroscopic redshifts with the photometric redshifts of sources within 5″ of the target position. Narrow-line radio galaxies are plotted in red, with quasars in blue and ambiguous sources in green. The two open circles represent host galaxies of multiple-component radio sources.

3

Spatial Clustering of FIRST Radio Sources at
$z \sim 0.34$ in GAMA[†]

3.1 Introduction

In recent decades, measurements of the CMB have proven the very early Universe to be remarkably isotropic with tiny under- and overdensities that have grown to form the vast and intricate structures we see today. Galaxies and clusters observed today are far removed from their almost homogeneous beginnings, and we require large numbers of them to piece together a statistical picture on cosmological scales. Clustering measures on large scales can be used to investigate not only the relationships between galaxy populations found by various techniques probing different epochs and masses, but also broader cosmological phenomena such as BAO, cosmic magnification, or the ISW effect (described in Section 1.2.5).

Cosmological applications require information about the gravitating mass distribution in the Universe, which, in a ΛCDM cosmology, is strongly tied to the dark matter distribution. Direct observations tell us only about the baryonic matter, from which we must infer the dark matter distribution. Various

---

[†]The results presented in this chapter have been accepted for publication in Monthly Notices of the Royal Astronomical Society (Lindsay et al., 2014).

tools exist for measuring the clustering signal of an observed source catalogue, such as nearest neigh-bour measures (e.g. Bahcall & Soneira, 1983), counts-in-cells (e.g. Magliocchetti et al., 1999; Blake & Wall, 2002b; Yang & Saslaw, 2011), correlation functions (Section 1.2.1; e.g. Groth & Peebles, 1977; Bahcall & Soneira, 1983; Blake & Wall, 2002a,b; Croom et al., 2005), and power spectra (Section 1.2.1; e.g. Cole et al., 2005; Percival et al., 2007; Komatsu et al., 2011). Due to its relative simplicity to calculate, and relation to its Fourier transform (the power spectrum), the two-point spatial correla-tion function has become a standard in quantifying cosmological structure. A means by which we can quantify the extent to which the observable and dark matter are tied using the correlation function is through the bias parameter $b(z)$ (see Section 1.2.4). This relates the spatial correlation function of an observed galaxy population to that of the underlying dark matter. The bias quantifies the difference in the clustering of the dark matter haloes acting solely under gravity, and of galaxies inhabiting those haloes, with other effects making their structure more or less diffuse. This is heavily dependent on the galaxy masses and the epoch under consideration (e.g. Seljak & Warren, 2004).

Local galaxies with masses comparable to the Milky Way are a relatively unbiased tracer of mass, with a present day correlation length (the clustering scale at which the correlation function falls below unity; see Section 3.3) of $r_0 = 5.4 \ h^{-1}$Mpc found in the early CfA redshift survey (Davis & Peebles, 1983). While the power law slope is consistently found to be ~ 2 in most studies, large variations in r_0 are found depending on the populations and respective epochs being observed, from $\sim 5 \ h^{-1}$Mpc for local galaxies to $\sim 25 \ h^{-1}$Mpc for Abell clusters (e.g. Bahcall & Soneira, 1983). At the lower end of the clustering scale, Saunders, Rowan-Robinson & Lawrence (1992) find $r_0 \sim 4 \ h^{-1}$Mpc for IRAS starburst galaxies and, using the 2dF QSO Redshift Survey, Croom et al. (2001) show a scale, roughly constant with redshift, of $\sim 5 \ h^{-1}$Mpc for quasars at $0 \lesssim z \lesssim 2.5$ (see also Croom et al., 2005; Ross et al., 2009; Ivashchenko, Zhdanov & Tugay, 2010). Likewise, Kovač et al. (2007) find that $z \sim 4.5$ Ly α emitters exhibit a relatively short clustering length of $4.6 \ h^{-1}$Mpc, consistent with their being progenitors of Milky Way type local galaxies (see Nilsson & Meisenheimer, 2009, and references therein), and note the similarity with Lyman break galaxies (LBGs) at $z \sim 3.8$ and $z \sim 4.9$, suggesting that the two populations reside in the same host haloes but with a relatively low duty cycle. Local $L \gtrsim L_*$ ellipticals, however, are a highly clustered population with various authors finding $r_0 \sim 7$–12 h^{-1}Mpc (e.g. Guzzo et al., 1997; Willmer, da Costa & Pellegrini, 1998; Norberg et al., 2002). Similarly high clustering lengths are found for extremely red objects (EROs) at $z \sim 1$ (e.g. Daddi et al., 2001; McCarthy et al., 2001; Roche et al., 2002) and distant red galaxies (DRGs) at $z \sim 1$–2 (Grazian et al., 2006; Foucaud et al., 2007), both posited as progenitors of the local bright ellipticals. Indeed, similar

results for radio galaxies (e.g. Cress et al., 1996; Overzier et al., 2003) give weight to the suggestion by Willott, Rawlings & Blundell (2001) that EROs and radio galaxies are identical and seen at different evolutionary stages, based on their findings of ERO-like hosts of radio galaxies in the 7C Redshift Survey.

Radio surveys are ideal for the purpose of carrying out such large-scale statistical measurements for a number of reasons. From a logistical point of view, radio wavelengths are useful because they occupy a uniquely broad atmospheric window in the electromagnetic spectrum, not being significantly absorbed by Earth's atmosphere. This allows observations to be carried out from terrestrial telescopes, reducing costs compared with launching and operating space telescopes. In a wide radio survey with relatively shallow flux-density limit ($S_{1.4} \gtrsim 1$ mJy), extragalactic sources dominate in the form of synchrotron radiation in active galaxies with supermassive black holes at their centre. By observing such powerful sources, and with the long radio wavelengths being immune to absorption by dust, the ISM and our own atmosphere, ground based telescopes are able to observe the whole sky relatively unobscured and to very high redshifts.

Of particular significance for the science discussed here is that these radio-loud AGN predominantly reside in massive elliptical galaxies (e.g. Jarvis et al., 2001; Willott et al., 2003; McLure et al., 2004; Herbert et al., 2011) and, as such, provide good tracers of large-scale clustering and the underlying dark matter distribution. Radio surveys, such as the 87GB and Parkes-MIT-NRAO surveys at 4.85 GHz, WENSS at 325 MHz, and FIRST and NVSS at 1.4 GHz, have been exploited with this aim for some time (e.g. Cress et al., 1996; Blake & Wall, 2002a; Overzier et al., 2003; Best et al., 2005; Fine et al., 2011), but their potential uses are limited without additional data at other wavelengths. While the large redshift range of radio surveys can be advantageous over more targeted optical surveys, this can also wash out the clustering signal due to the superposition of the influence of differently behaving subpopulations (e.g. Wilman et al., 2008). This effect can be alleviated by the use of complementary optical surveys to aid the separation of radio sources by simple redshift binning or other methods.

As technology advances, observations at radio wavelengths will be able to constrain cosmological parameters with significantly more precision than the previous generation of surveys (Raccanelli et al., 2012; Camera et al., 2012), as well as provide the data for much more galactic and extragalactic science (e.g. Norris et al., 2011). Moving forward, the large numbers of sources observed in various overlapping surveys will make multi-wavelength studies not only more common, but essential to reliably identify objects by their emission across the entire spectrum and delineate the clustering properties of various subpopulations of any given individual survey.

Chapter 2 highlighted the inherent difficulties in targeting optical host galaxies of radio AGN, with many having very faint optical emission, and ambiguity in host galaxy position due to the prevalence of extended and/or multiple-component radio sources. As a result, a sample of radio sources with spectroscopic redshifts is defined in this chapter by taking an existing optically-selected spectroscopic survey (GAMA) and cross-matching with radio sources from FIRST. This does not provide a complete sample of redshifts for any given radio flux limit, but the depth and large numbers of sources in both catalogues allows for a significant detection of the spatial clustering properties of the cross-matched sources, with reasonable errors on the inferred parameters.

The outline of this chapter is as follows: Section 3.2 describes the surveys and catalogues used; Section 3.3 describes direct measurements of the spatial clustering properties of the cross-matched sample of radio sources with GAMA redshifts through the spatial correlation function and projected correlation function; Section 3.4 describes the linear bias inferred from the correlation function results, which are discussed in Section 3.5, and Section 3.6 summarizes our results and conclusions.

3.2 Data

To infer the spatial clustering parameters of the radio galaxy population, redshifts of the sources are required. We use spectroscopic redshifts from the GAMA survey, cross-matching with radio sources from the FIRST survey to assign optical counterparts and redshifts to each radio source.

3.2.1 Optical Catalogues

The Galaxy And Mass Assembly (GAMA; Driver et al. 2011) survey has been in operation since 2008, using the 3.9 m Anglo-Australian Telescope (AAT) and the AAOmega spectrograph to build a > 98 per cent complete redshift catalogue of $\sim 140{,}000$ galaxies to a depth of $r < 19.4$ or 19.8. Two gratings, with central wavelengths of 4800 Å and 7250 Å are used, giving continuous wavelength coverage over the range 3720–8850 Å, with resolution ~ 3.5 Å (in the blue channel) and ~ 5.5 Å (in the red channel). Spectroscopic redshifts are found in real time using the RUNZ code developed for the 2dFGRS (Colless et al., 2001) and given a quality flag Q of 0–4 where $Q \geq 3$ is a probable or certain redshift worthy of publication (Hopkins et al., 2013). Corresponding photometry is also available from UKIDSS and SDSS (Hill et al., 2011). The redshift distribution of GAMA galaxies, and a comparison with photometric redshifts from SDSS and UKIDSS photometry, is shown in Figure 3.1.

GAMA is divided over three $12° \times 4°$ equatorial regions centred at $\alpha = 135°$ (9h), $180°$ (12h) and
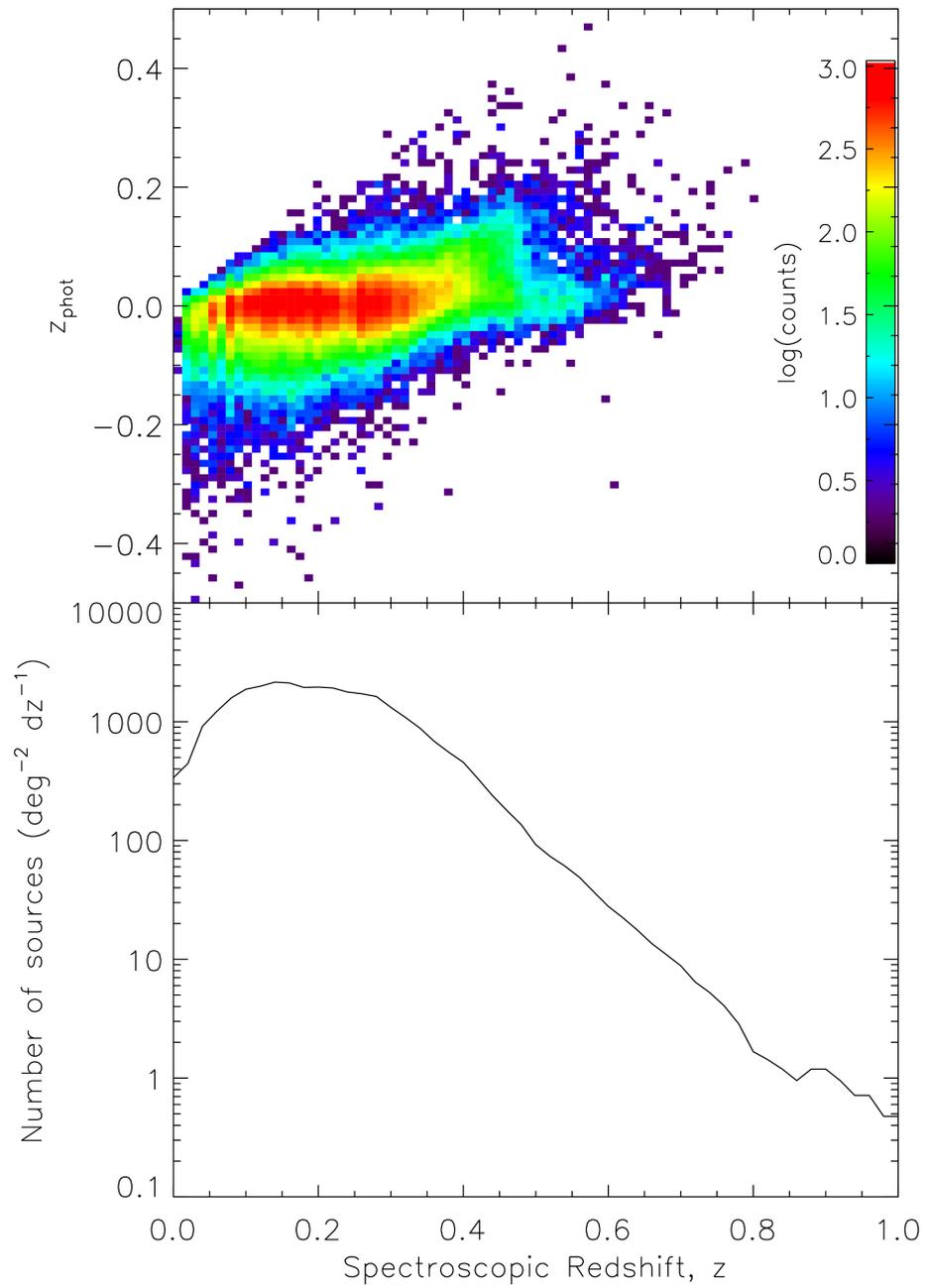
FIGURE 3.1: Redshift distribution of the full GAMA spectroscopic catalogue (galaxies only).

217.5° (14h30), covering 144 square degrees. These were chosen to be large enough in area to fully sample $\sim 100\ h^{-1}$Mpc structures at $z \sim 0.2$, and for their overlap with existing surveys. The data used for this thesis come from the first 3 years of observations (GAMA Phase I; Driver et al. 2011).

The overlap of the GAMA survey area with other surveys allows for the collection of sufficient photometry to be able to generate a large contribution to the redshift catalogue using photometric redshifts. These will serve, in this chapter, to prevent radio sources being mistakenly associated with a GAMA galaxy in the event that a more likely match is to be found in a photometric catalogue. The photometric matches will be used in their own right in Chapter 4. The Sloan Digital Sky Survey (SDSS; York et al., 2000) has $\sim 10,000$ square degrees of sky coverage in 5 bands (*ugriz*), with the southernmost stripes covering the 3 GAMA fields. The UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007) uses the 3.8 m UKIRT consisting of several subsurveys, including the Large Area Survey (LAS). UKIDSS LAS complements the SDSS, covering over 2,000 square degrees in the *YJHK* bands within the SDSS regions and overlapping with GAMA. We therefore have photometry in up to 9 optical and infrared bands with which to obtain photometric redshifts for those galaxies without spectroscopy (further details are given in Smith et al., 2011). For those sources detected in the SDSS imaging data but without GAMA spectroscopic redshifts, we impose a magnitude limit of $r < 22$, i.e. within the limit of the SDSS imaging data.

3.2.2 Radio Catalogue

The FIRST survey (Becker, White & Helfand, 1995) was carried out at an observing frequency of 1.4 GHz with the Very Large Array (VLA) in B configuration. The most recent catalogue[*] contains 946,464 sources covering over 10,000 square degrees ($\sim 8,500$ in the northern galactic cap and $\sim 1,500$ in the south) with an angular resolution of $5.4''$ (FWHM) to a completeness of 95 per cent at $S > 2$ mJy. A number of sidelobes are spuriously counted in the raw catalogue and an oblique decision-tree program, developed by the FIRST survey team (White et al., 1997), finds probabilities of each catalogue entry representing extended activity from a nearby bright source. Those entries with a sidelobe probability of > 0.1 have been excluded for the purposes of this analysis (< 20 per cent of the catalogue), leaving 723,934 sources above 1 mJy.

The NRAO VLA Sky Survey (NVSS; Condon et al. 1998) is a similar survey using the VLA in the more compact D and DnC configurations. This gives more accurate integrated fluxes for extended sources due to the shorter minumum baselines, as well as greater sky coverage compared with FIRST.

---

[*]*http://sundog.stsci.edu/first/catalogs/readme_12feb16.html*

Neither of these factors is of significant advantage for the purposes of this work, which requires only a known flux limit (with FIRST being marginally deeper) and which focuses on a particular area on the sky covered by both surveys. The $45''$ resolution of NVSS prohibits meaningful cross-matching to much more densely populated optical data, making FIRST the more appropriate choice.

3.2.3 SKADS Simulated Skies

Supplementary mock radio catalogues are provided by the SKA Design Study (SKADS) in the form of SKADS Simulated Skies* (S^3), the collective name of a range of simulations developed to reflect the radio continuum data we hope to obtain with the SKA, aiding the design process. The S^3–Semi-empirical Extragalactic simulation (Wilman et al., 2008, 2010) describes extragalactic radio continuum sources over a ~ 400-square-degree field ($-10° < \alpha < 10°$ and $-10° < \delta < 10°$) out to $z \sim 20$. Galaxies are broadly divided morphologically as defined by `sftype` and `agntype` flags. The categories used are: quiescent star-forming galaxies (`sftype=1`), starburst galaxies (`sftype=2`), radio-quiet AGN (`agntype=1`), Fanaroff-Riley Class I (FRI; `agntype=2`) and Fanaroff-Riley Class II (FRII; `agntype=3`) AGN. A summary of the relevant details of the simulations is given here, but a far more in-depth description is found in Wilman et al. (2008).

The SKADS simulation is built upon a present-day dark matter density field, $\delta\rho/\rho$, defined on a cuboid grid of $5\ h^{-1}$Mpc comoving cells in a $550 \times 550 \times 1550$ array, with an imaginary observer placed at the centre of one of the smaller faces looking into the array over a solid angle $\Delta\Omega$. This cell size is a compromise between computational considerations, as well as being large enough to keep $\delta\rho/\rho$ within the linear or quasi-linear regime, while small enough to have adequate mass resolution to resolve clusters of mass $10^{14}\ h^{-1} M_\odot$. Comoving distances are inverted to yield redshifts† in each cell (denoted by suffix i), with observed frame SEDs of each source type (denoted by suffix j) compared with the flux density limits of the simulation to derive a minimum luminosity, down to which the luminosity function is integrated to determine the space density, $\phi(L > L^i_{\min,j}, z_i)$. Large-scale structure is then introduced by the exponential term in the following definition of the expected number of sources in the cell:

$$n^i_{\exp,j} = A e^{b(z_i)D(z_i)\delta\rho/\rho} \phi(L > L^i_{\min,j}, z_i)\Delta V, \qquad (3.1)$$

where $D(z)$ is the linear growth factor for dark matter fluctuations, $b(z)$ is the bias of the population

---

http://s-cubed.physics.ox.ac.uk

†A spatially flat cosmology was used, with $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\mathrm{M}} = 0.3$, $\Omega_\Lambda = 0.7$, $f_{\mathrm{baryon}} = 0.16$, $\sigma_8 = 0.74$ and $n_{\mathrm{scalar}} = 1$.
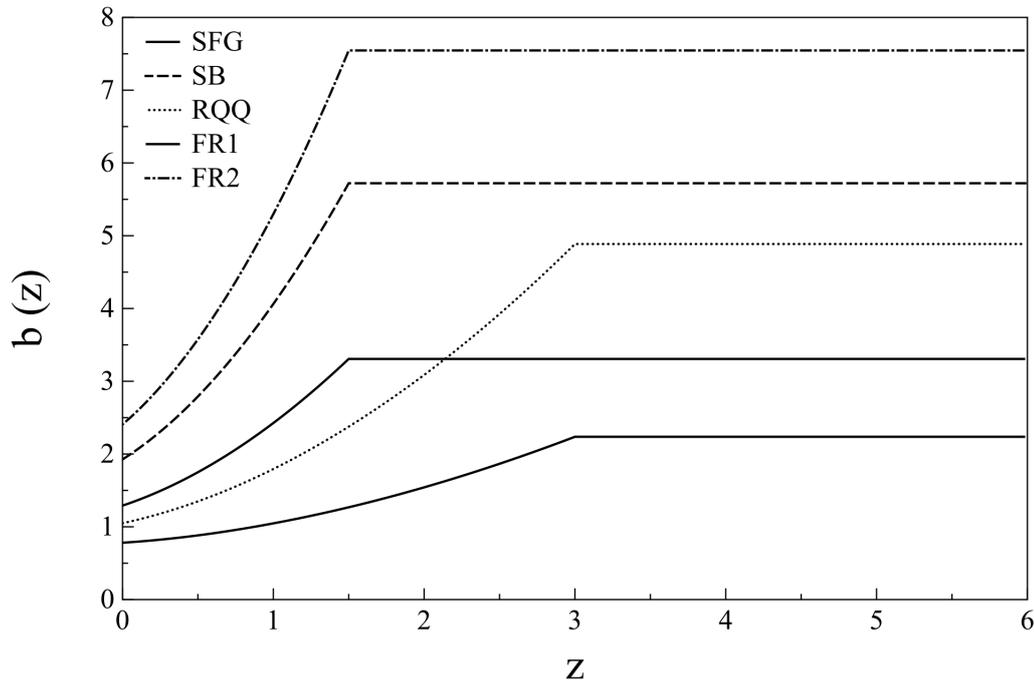
FIGURE 3.2: Bias models assumed by Wilman et al. (2008) for the five radio source populations in the SKADS simulations. (Raccanelli et al., 2012).

of interest, and A is a normalization factor which ensures agreement with the luminosity function when averaged over the largest scales.

Being semi-empirical, the simulations do follow observations with the breaking up of simulated galaxies into different populations, and the use of their individual luminosity functions and clustering properties (to the extent to which these are known).

Radio-quiet AGN are simulated, reflecting the fact that classical radio-loud AGN constitute only a small fraction of the AGN population. Jarvis & Rawlings (2004) used a more complete census of AGN activity provided by the hard X-ray AGN luminosity function (Ueda et al., 2003, HXLF), combined with a radio–X-ray luminosity relation, to show that radio-quiet AGN make a significant contribution to the upturn in 1.4 GHz source counts below 1 mJy (White et al., 2014). Classical double-lobed radio-loud AGN are described using the 151 MHz luminosity function derived by Willott, Rawlings & Blundell (2001), namely their 'model C' luminosity function consisting of low- and high-luminosity components with different functional forms and redshift evolutions. These two components naturally form the basis of the separation of low-luminosity FRI and high-luminosity FRII sources, although it is noted that the luminosity at which the low- and high-luminosity components contribute equally to the luminosity function is about 1 dex higher than the traditional FRI/FRII break luminosity. Similarly,

the luminosity function for star-forming galaxies is taken from the 1.4 GHz IRAS 2 Jy sample by Yun, Reddy & Condon (2001), modelled as the sum of two Schechter functions. For the purpose of assigning morphologies and SEDs, the low- and high-luminosity components of the luminosity function are identified with normal/quiescent late-type galaxies and starburst galaxies, respectively – a separation borne out through further analysis by Takeuchi, Yoshikawa & Ishii (2003). Both components of the luminosity function were ascribed a form of cosmological evolution which takes the form of pure luminosity evolution, in an Einstein-de Sitter universe, of the form $(1+z)^{3.1}$ out to $z = 1.5$, with no further evolution thereafter.

The simulation lacks the mass resolution to directly resolve galaxy- and group-sized haloes, precluding the halo occupation distribution (HOD) function technique used to populate dark matter haloes in N-body simulations. For each source population, a separate bias $b(z)$ is computed using the formalism of Mo & White (1996), with the assumption of a single effective dark halo mass which reflects its clustering. This assumption breaks down at high redshifts as $b(z)$ increases, potentially leading to excessively strong clustering due to the exponent in equation 3.2.3. To circumvent this, the bias of each population is held constant beyond a certain redshift, z_{cut}. The halo mass for radio-quiet AGN is informed by the findings of e.g. Croom et al. (2005) that the clustering of quasars in the 2dFQZ survey is well described by a constant halo mass of $3 \times 10^{12}\ h^{-1} M_{\odot}$ to the redshift limit of the survey ($z = 2.5$), motivating a choice of $z_{\mathrm{cut}} = 3$. Halo masses for the radio-loud AGN populations have been fixed using the discussion of Overzier et al. (2003), using the angular clustering of NVSS and FIRST sources. FRIs are assigned halo masses of $10^{13}\ h^{-1} M_{\odot}$, reflecting the clustering of low-luminosity radio sources and $L \sim L^{\star}$ early-type galaxies, and FRIIs are assigned halo masses of $10^{14}\ h^{-1} M_{\odot}$, which reproduces the clustering of the powerful radio sources and the clustering evolution of $z = 0$ $L \gtrsim L^{\star}$ early-type galaxies, $z \sim 1$ EROs and $z \sim 0.55$ LRGs. The resulting $b(z)$ increases very rapidly with redshift, and given the lack of clustering constraints at higher redshifts, the bias is held fixed from beyond $z_{\mathrm{cut}} = 1.5$ for these sources. For normal star-forming galaxies, $M_{\mathrm{halo}} = 10^{11}\ h^{-1} M_{\odot}$ and $z_{\mathrm{cut}} = 3$, reproducing the $z = 0$ clustering of IRAS galaxies, $L \sim 0.5 L^{\star}$ late-type galaxies and LBGs at $z \sim 3$. For the starbursts, $M_{\mathrm{halo}} = 5 \times 10^{13}\ h^{-1} M_{\odot}$, consistent with the claim of Swinbank et al. (2006) that submillimetre galaxies are progenitors of $L > 3 L_{\mathrm{K}}^{\star}$ early-type galaxies at $z = 0$, and with the clustering of mid-IR samples at $z =$ 1.5–3 Farrah et al. (2006). Given the group/cluster-sized halo mass for these sources, as for the radio-loud AGN, the bias is held constant beyond $z_{\mathrm{cut}} = 1.5$

Redshifts are quoted both with and without taking peculiar velocities into account (i.e. ignoring or including observed redshift-space distortions), as are line-of-sight distances. The database includes

flux densities down to a limit of 10 nJy for 5 radio frequencies (151 MHz –18 GHz) and 10 IR bands (24 – 1200 μm). The simulated redshift distributions are found to agree with deep field studies down to 100 μJy (e.g. Simpson et al., 2012), and the source counts are consistent with the deepest JVLA observations (Condon et al., 2012).

The SKADS catalogue is used to find an approximate redshift distribution for a given radio flux limit, and to find approximate, relative numbers of different galaxy populations for a given redshift distribution. The redshift distributions allow for some three-dimensional clustering information to be drawn even from those radio sources we are not able to match to an optical counterpart (see Chapter 4). The number counts of different types of galaxy are based on separate observationally-derived luminosity functions so we are able to construct a mock population distribution based on the redshifts of our samples. This gives us a means of predicting a value expected for the bias, given the halo masses of respective galaxy types (see Wilman et al., 2008; Raccanelli et al., 2012, for more details).

3.2.4 Collapsing of Multi-Component Radio Sources

In order to address the inevitable issue of extended radio sources resulting in multiple detections for one host galaxy, of which perhaps none corresponds to the core itself (and therefore any associated optical source), we have followed Cress et al. (1996) in applying a collapsing radius of $72''$ ($0.02°$) to the FIRST catalogue. Any FIRST sources within this radius of one another are grouped and combined to form a single entry positioned at the flux-weighted average coordinates of the group and attributed with their summed flux-density (an underestimate, not accounting for diffuse emission resolved out between sources, but this is not a problem due to using the lowest flux limit to define our sample in the first place). This precludes probing the correlation function to angular scales smaller than $72''$, but ensures that only extremely extended sources will be mistakenly treated as independent galaxies. The procedure does, however, suffer from introducing far greater positional uncertainty to the assumed core position given the inherent variation in morphology in the observed objects. Asymmetric multiple-component galaxies or independent galaxies with small angular separations by chance, in particular, will interfere with the reliability of the collapsed catalogue. From the original FIRST catalogue, 227,606 sources are collapsed in 98,225 distinct groups, leaving a > 1 mJy catalogue of 585,473 sources, with 10,399 of these being located within the 3 GAMA fields.

Returning to the option of supplementing FIRST with NVSS data to better deal with extended sources, when matching collapsed FIRST sources or NVSS sources directly to the optical catalogues, we find no evidence of a significant difference between the results.
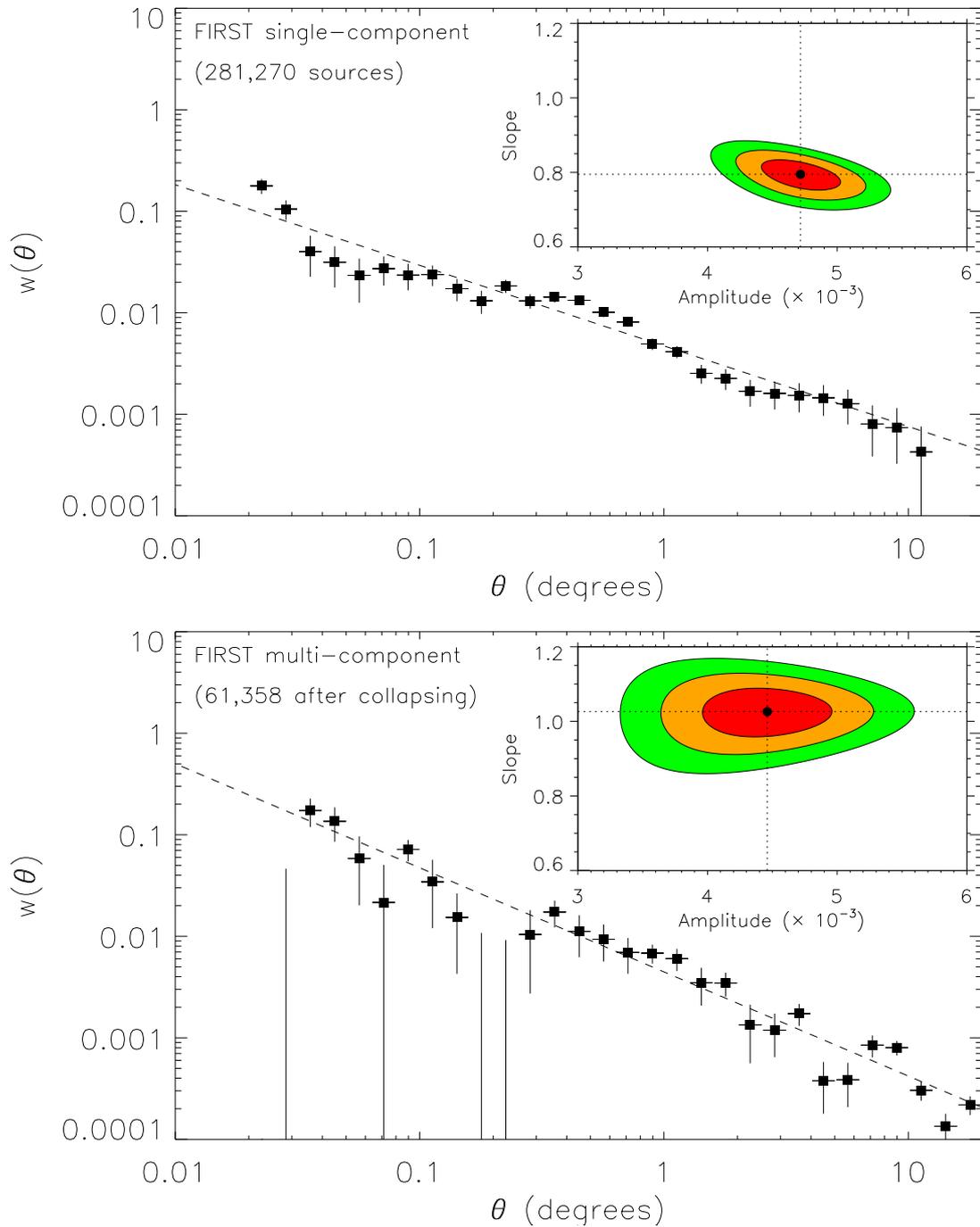
FIGURE 3.3: Angular correlation functions for a 115° × 64° area of the FIRST survey at 1 mJy, with contours at $\Delta\chi^2 = 2.30$, 6.18 and 11.83 for the power-law parameter fits. The top panel corresponds to those FIRST sources that have no neighbours within the 72″collapsing radius adopted. The bottom panel corresponds to the remaining (assumed) extended sources post-collapsing.

Figure 3.3 shows the angular correlation function (discussed in Section 1.2.1, and further in Section 4.3) of a large area ($115° \times 64°$) of the FIRST survey, separately for single sources and collapsed extended sources. The single sources demonstrate a well-constrained power law form while the collapsed sources show more scatter around a power-law fit, owing to their smaller numbers. However, the results are still broadly compatible, and the signal of the combined, collapsed catalogue is dominated by the well-behaved single source power law. The two samples would not be expected to have the same $w(\theta)$, however, as the collapsed sources are dominated by extended FRI and FRII AGN, in contrast to the largely compact single-component sources, which are a combination of compact AGN and low-redshift star-forming galaxies.

3.2.5 Optical Identification of FIRST Sources

An optical catalogue somewhat larger than the GAMA survey area is used for the purpose of cross-matching with the radio sources (see Table 4.1), making use of photometric redshifts determined using up to 9 SDSS/UKIDSS-LAS bands ($ugrizYJHK$). This allows for greater source counts and improved statistics for the analysis, as well as expanding to a region completely containing the GAMA fields as well as other surveys like *Herschel*-ATLAS. The expanded region comprises three $14° \times 5°$ fields, subsuming the three $12° \times 4°$ GAMA fields.

Of the ~ 6 million optical/IR sources in this area with at least a photometric redshift, approximately half are removed on the basis of the star–galaxy separation technique by Baldry et al. (2010), leaving only the likely extragalactic sources to be matched to the radio. This separation is based on $J - K$ and $g - i$ colours rather than simply removing point-like sources and therefore fewer quasars are mistakenly discarded, provided they lie away from the stellar locus. As a second measure, a minimum redshift criterion of $z > 0.002$ was added to filter out the nearby objects (some with significantly negative redshifts), which are assumed to be stellar in origin or to correspond to highly extended extragalactic sources, although we note that this may also filter out a small number of real low-redshift extragalactic radio sources.

For the purpose of radio-optical cross-matching, given the positional accuracies of the catalogues ($< 1''$ for FIRST, $\sim 0.3''$ for GAMA, $\sim 0.1''$ for SDSS) a simple nearest-neighbour match can be reliable. Sullivan et al. (2004) find this method to produce very similar catalogues to the likelihood ratio protocol of Sutherland & Saunders (1992) in their work. The likelihood ratio technique is often used to identify radio sources (Gonzalez-Solares et al., 2005; Ciliegi et al., 2005; Afonso et al., 2006; McAlpine et al., 2012), but here we instead use the simpler method tested by Sullivan et al. (2004) and adopted
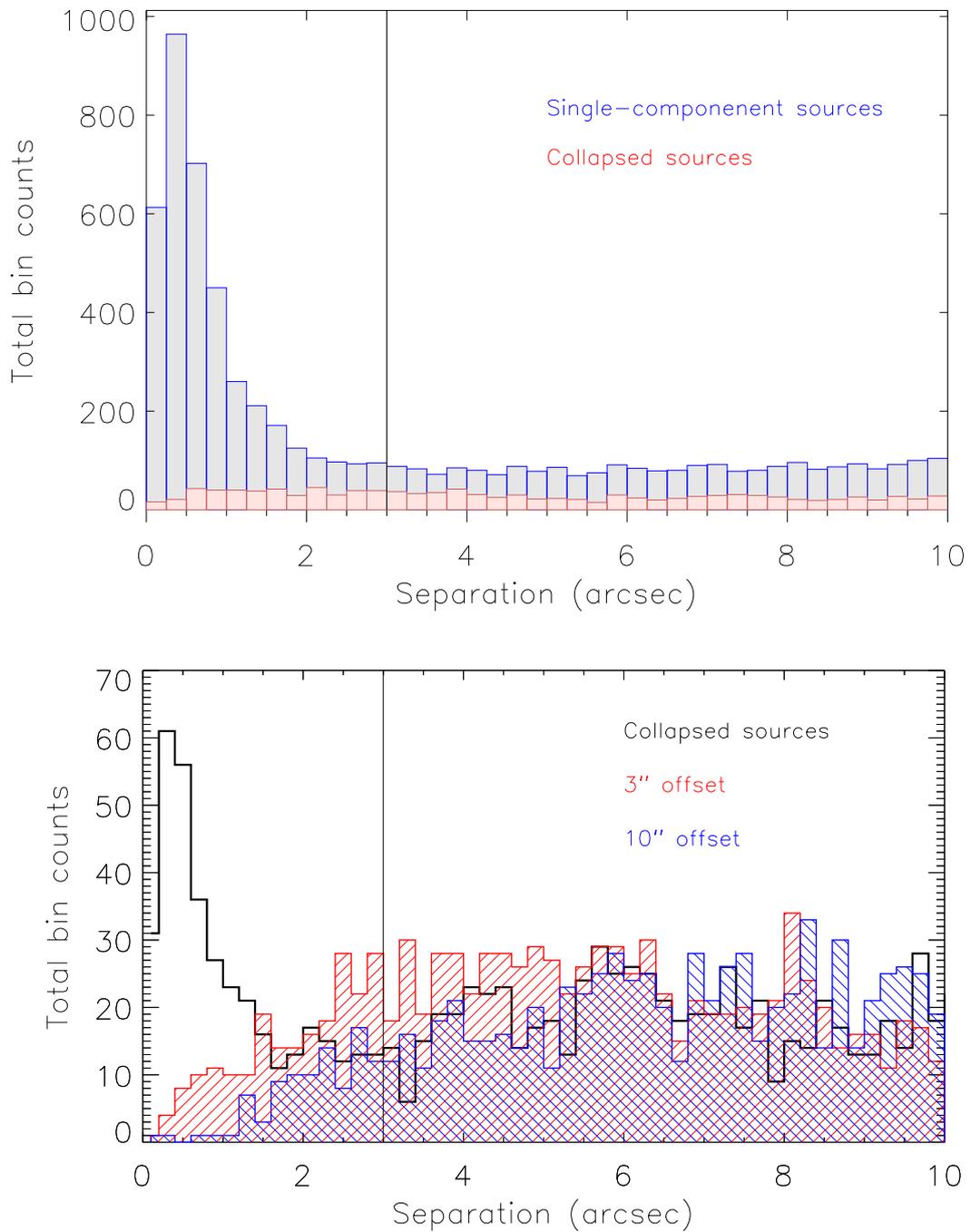
FIGURE 3.4: Top panel: Distribution of offsets between FIRST sources and the nearest GAMA/SDSS source, with single-component sources (blue) placed on top of the collapsed sources (red). The vertical solid line indicates the proposed separation cut-off of 3″ to define a good radio-optical cross-match, below which counts rise above the background level. Collapsed sources account for ∼ 10 per cent of matches within the cut-off. Bottom panel: Distribution of separations for the collapsed radio sources (black). The filled histograms show the result of assigning a random Gaussian offset of $\sigma = 3″$ (red) and 10″ (blue) before the cross-matching, showing a strong decline in matches within 2″.

by e.g. El Bouchefry & Cress (2007). The 10,399 radio sources within the three GAMA fields were all found to have optical counterparts (photometric or spectroscopic) within $1'$. By inspection of the separation distribution shown in Figure 3.4, any separations of more than a few arcseconds are likely to be random. A compromise must be reached between including the greatest number of our radio sample and ensuring reliable optical cross-IDs. The expected contamination of the remaining objects by chance proximity of optical sources can be expressed through the simple equation:

$$P_c = \pi \times r_s^2 \times \sigma_{\text{opt}}, \tag{3.2}$$

where r_s is the search radius cut-off and $\sigma_{\text{opt}} = 9.23 \times 10^3 \text{ deg}^{-2}$, the surface density of the optical/IR galaxy catalogue (which is a function of flux, with source density increasing to fainter flux limits). This gives a 0.9 per cent contamination rate for a $2''$ cut-off and 2.0 per cent for a $3''$ cut-off. In similar matching procedures, Prandoni et al. (2001) used a $3''$ cut-off (2 per cent contamination rate) for the ATESP–EIS (Australia Telescope ESO Slice Project and ESO Imaging Survey), Magliocchetti & Maddox (2002) used a $2''$ cut-off (5 per cent contamination) for the APM(Automatic Plate Measuring)– FIRST matches, and El Bouchefry & Cress (2007) used a $2''$ cut-off (3.5 per cent contamination rate) for the FIRST–NDWFS (NOAO Deep Wide Field Survey). We place our separation limit at $3''$, below which the distribution in Figure 3.4 becomes visibly more dense.

While positional coincidence is appropriate for defining positive identifications of single, isolated radio sources, it is not necessarily so for the collapsed multiple sources, as this introduces uncertainty into the assumed position of the optical core of the host galaxy (as was apparent in the work of Chapter 2). These sources are in the minority, but to assess the reliability of their matches, we repeat the cross-matching process with the radio positions randomly displaced by $3''$ and $10''$ Gaussian distributions. Figure 3.4 (bottom) shows the resulting separation histograms. While all three iterations of the procedure agree above the $3''$ matching cut-off, the significant peak at low separations is completely lost after shifting the radio positions. There will almost certainly be a minority of collapsed sources for which we do not accurately find a radio core position, but this test shows that for those where the collapsing is successful, reliable optical matches are being found.

3.2.6 Completeness

While the FIRST radio catalogue contains a large number of ~ 1 mJy sources, it is incomplete below 2–3 mJy. Discarding these fainter radio sources would significantly affect the size of our samples and the statistical significance of our results, but the effects of the incompleteness at the mJy level must
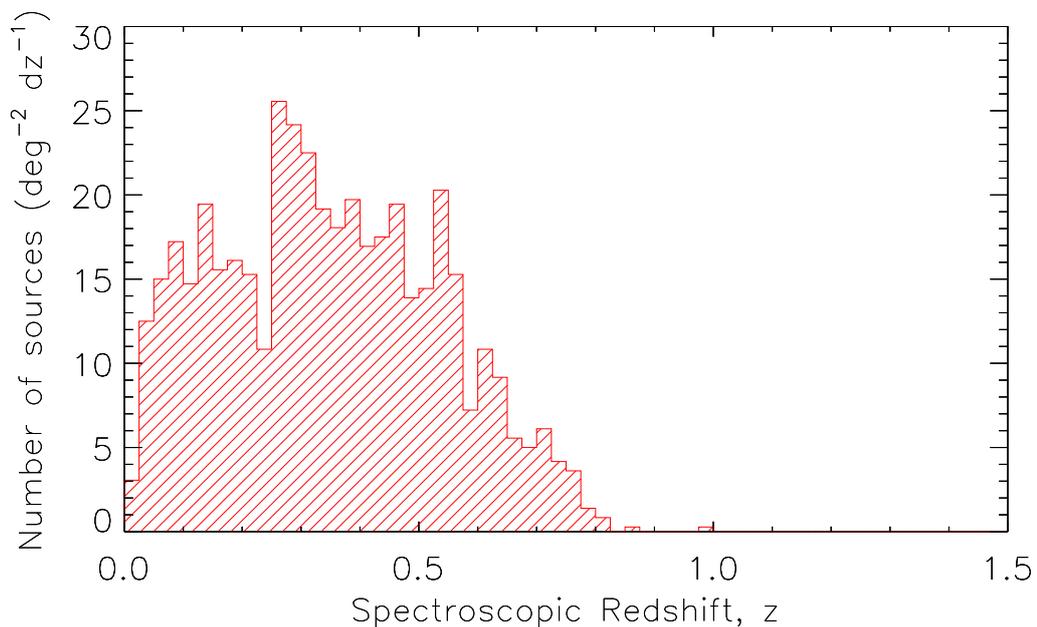
FIGURE 3.5: Redshift distribution of FIRST sources with an optical counterpart within $3''$ in the GAMA spectroscopic catalogue.

be corrected for. To account for the noise variations across the survey area we use an rms noise map of the FIRST survey and their source detection criterion of $S - 0.25(\mathrm{mJy}) > 5 \times$ rms to determine whether a radio source randomly drawn from the SKADS simulated catalogues would be detected in the FIRST data over the GAMA fields. We use only those sources from SKADS with $S_{1.4} > 0.5 \,\mathrm{mJy}$, and if the SKADS radio source fulfils the FIRST survey detection criteria then that source is retained, whereas if the source falls below the detection threshold then the source is omitted. This process not only provides an estimate of the completeness of the survey but also allows us to compile a catalogue of sources, with both flux-density and redshift information, that is subsequently used as our random source catalogue for calculating the spatial correlation function. This ensures that we fully account for any noise variation and incompleteness across the survey area. The model redshift distribution of radio sources in our survey regions is thus modified according to these noise variations, and it is this modified redshift distribution that is used throughout in determining the spatial correlation function.

3.3 Spatial Correlation Function

The clustering properties of a sample of galaxies whose positions are known in three dimensions can be analysed via the spatial two-point correlation function, $\xi(r)$ (defined in Section 1.2.1). The simplest

estimator defines $\xi(r) = DD/RR - 1$, where $DD(r)$ and $RR(r)$ represent the number of galaxy pairs separated by r in the real data and a random catalogue, respectively. Our chosen estimator, by Landy & Szalay (1993), includes the cross-pair counts ($DR(r)$):

$$\xi(r) = \frac{n_r(n_r - 1)}{n_d(n_d - 1)} \frac{DD}{RR} - \frac{(n_r - 1)}{2n_d} \frac{DR}{RR} + 1, \tag{3.3}$$

where n_d and n_r are the number of real sources and random sources.

3.3.1 Random Catalogues

A random catalogue is defined over a geometrically identical area to that of the data, randomly distributed in angular position such that any selected region is equally likely to contain a source as any other region of equal area (subject to subsequent modulation based on the noise properties of the survey data; see Section 3.2.6). For rectangular areas such as the GAMA fields, the right ascension coordinates are uniformly distributed. The declination coordinates are close to uniformly distributed due to the GAMA fields lying on the equator, but in order to account for the spherical geometry, random source positions are generated uniformly in $\sin(\delta)$, to compensate for the decreasing angular separations of lines of longitude with increasing declination. The random catalogue represents a mock sample of galaxies with the same selection, but in the absence of clustering. As such, the angular positions are randomly distributed, but the redshift distribution is preserved by randomly sampling the redshifts of the real sources, and adding some scatter to produce a smooth, continuous distribution of redshifts.

By averaging over several random data sets and using $\overline{DR}$ and $\overline{RR}$, or by using a more densely populated random catalogue, we may assume the statistical error in the random sets to be negligible. The random catalogues themselves have been autocorrelated finding no significant deviation from zero even at extremes of separation where bin counts are lowest. The uncertainty on ξ, therefore, is often given by the Poisson error due to the DD counts alone

$$\Delta\xi = \frac{1 + \xi(r)}{\sqrt{DD}}. \tag{3.4}$$

However, the errors in the correlation function depend on the DD counts beyond simple Poisson variance; adjacent bins are correlated, with each object contributing to counts across a range of separation bins. The errors are therefore calculated somewhat more rigorously using a bootstrap resampling technique (Ling, Barrow & Frenk, 1986) whereby several data catalogues are constructed by randomly sampling (with replacement) the original set of objects. As such, in any given set, some sources are counted twice or more and some not at all. The resulting binned DD counts should give a mean approximately equal to the original data but allow us to calculate a variance for each bin, and therefore

$\xi(r)$ values. Cress et al. (1996), for example, found errors in the angular correlation function, $w(\theta)$, for the early FIRST survey with a bootstrap resampling method. They found the Poisson error estimates to be too small by a factor of 2 on small scales ($\sim 3'$) and more than an order of magnitude for larger scales ($\sim 5°$).

The restricted survey area from which we can measure $\xi(r)$ also results in a negative offset in the observed correlation function, known as the *integral constraint*. Expressed mathematically, the relation between the observed correlation function $\xi_{\mathrm{obs}}(r)$ and the genuine function $\xi(r)$ is

$$\xi_{\mathrm{obs}}(r) = \xi(r) - \sigma^2, \tag{3.5}$$

where σ^2 represents the integral constraint (Groth & Peebles, 1977) which can be approximated, following Roche & Eales (1999), by

$$\sigma^2 = \frac{\sum RR(r)\xi(r)}{\sum RR(r)}. \tag{3.6}$$

Calculating spatial separations between galaxies requires information about the underlying cosmology and is dependent on the dynamics of the clustered systems. At face value, the only extra data required for our catalogue is a line-of-sight distance, which we calculate based on observed redshifts, given by

$$\chi(z) = \frac{c}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')}, \tag{3.7}$$

where H_0 is the Hubble constant, and $E(z)$ is the function used to describe the cosmological expansion history:

$$E(z) = \left[\Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}\right]^{\frac{1}{2}}. \tag{3.8}$$

Given the comoving distances to two objects, χ_1 and χ_2, and their angular separation, θ, their comoving spatial separation is given by

$$r = \left(\chi_1{}^2 + \chi_2{}^2 - 2\chi_1\chi_2\cos\theta\right)^{\frac{1}{2}}, \tag{3.9}$$

assuming a flat cosmology (see Liske 2000 for more general expressions).

An epoch-dependent form of the spatial correlation function is assumed (see e.g. de Zotti et al., 1990; Overzier et al., 2003, and references therein):

$$\xi(r,z) = \left(\frac{r_0}{r}\right)^\gamma \times (1+z)^{\gamma-(3+\epsilon)}, \tag{3.10}$$

where r is in comoving units and ϵ parameterizes the clustering model being assumed. Overzier et al. (2003) and Kim et al. (2011) offer 3 main models: *stable clustering* (where clusters have fixed physical

size; $\epsilon = 0$); *comoving clustering* (where clusters have fixed comoving size; $\epsilon = \gamma - 3$); or *linear clustering* (growth under linear perturbation theory; $\epsilon = \gamma - 1$). For a typical slope found in the literature of $\gamma \sim 2$, these clustering models are stable, decaying, and growing, respectively. Other authors, such as Elyiv et al. (2012), also apply $\epsilon = -3$. Given that $\epsilon > 0$ (or < 0) implies growing (or decaying) clustering, this final ϵ value implies a more rapid clustering decay than the other models. In this thesis we adopt $\epsilon = \gamma - 3$, showing some results with $\epsilon = \gamma - 1$ to provide a conservative range of values at high redshift.

The spatial two-point correlation function is usually fitted by a single power law over a significant range of separations:

$$\xi(r, z) = \left(\frac{r}{r_0(z)}\right)^{-\gamma},$$

(3.11)

where $r_0(z) = r_0(1 + z)^{1 - \frac{3+\epsilon}{\gamma}}$, incorporating the clustering evolution described.

3.3.2 Projected Correlation Function

By calculating galaxy distances from their redshifts and a cosmological model as in equations 3.7 and 3.9, we actually find the *redshift-space* correlation function, $\xi(s)$, which is systematically different from the true *real-space* correlation function $\xi(r)$. The difference occurs because measured redshifts are not entirely the result of cosmological expansion but they are also affected by radial peculiar velocity components of the source distorting the observer's view (the Kaiser effect; Kaiser (1987) and "Fingers of God"). These effects make it difficult to find the real-space clustering parameters r_0 and γ from direct $\xi(s)$ measurements. These can be found directly through the projected correlation function $\Xi(\sigma)$ or indirectly through deprojecting the angular correlation function $w(\theta)$ (see Section 4.4).

Redshift-space distortions affect our measurements of galaxy separations only in the direction of their lines of sight. This can be ameliorated by calculating the correlation function as a function of the line-of-sight separation, π, and the transverse separation, σ. Integrating over the π coordinate gives the projected correlation function, which is a function of the ostensibly real-space σ coordinate:

$$\Xi(\sigma) = 2 \int_0^\infty \xi(\sigma, \pi) \mathrm{d}\pi.$$

(3.12)

For practical purposes, we must impose our own upper limit on π for the integration in order to strike a compromise between capturing the full clustering signal and avoiding the introduction of excessive noise at the highest separations. Hawkins et al. (2003), for example, find their results insensitive to $\pi_{\mathrm{max}} > 60 \ h^{-1}$Mpc. Figure 3.6 shows how sensitive the results are to this choice in π_{max}. We use a
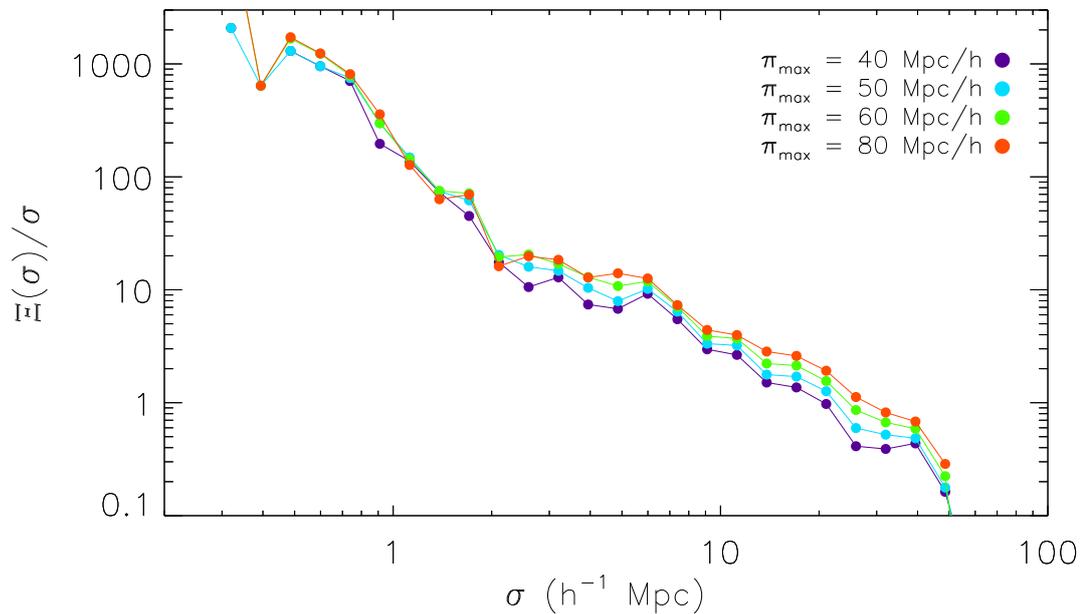
FIGURE 3.6: The effect of different values of π_{max} (the upper limit of the integral in equation 3.3.2) on the projected correlation function.

limit of $\pi_{\mathrm{max}} = 50 \ h^{-1}$Mpc in this work, avoiding considerable noise in $\xi(\sigma, \pi > 50)$. This function, free of redshift-space effects, can be related to the real-space correlation function (Davis & Peebles, 1983):

$$\frac{\Xi(\sigma)}{\sigma} = \frac{2}{\sigma} \int_{\sigma}^{\infty} \frac{r\xi(r)}{\sqrt{r^2 - \sigma^2}} \mathrm{d}r. \tag{3.13}$$

Assuming a power law form for $\xi(r)$ as in equation 3.11, we can fit $\Xi(\sigma)/\sigma$ to find the real-space correlation length, r_0, and power law slope, γ:

$$\frac{\Xi(\sigma)}{\sigma} = \left(\frac{r_0}{\sigma}\right)^{\gamma} H_{\gamma}, \tag{3.14}$$

where H_{γ} is related to the Gamma function, $H_{\gamma} = \Gamma(\frac{1}{2})\Gamma(\frac{\gamma-1}{2})/\Gamma(\frac{\gamma}{2})$.

Naturally, the requirement of having spectroscopic redshifts to accurately determine the galaxy pair separations limits the use of this method to those 1,635 radio sources with GAMA spectroscopy. Furthermore, by binning galaxy pairs in two dimensions and then integrating, we sacrifice signal-to-noise in any given σ bin, limiting the scope of the available data. We most likely underestimate the errors associated with $\Xi(\sigma)$, however, as the errors on the numerous π bins being integrated are treated as independent.
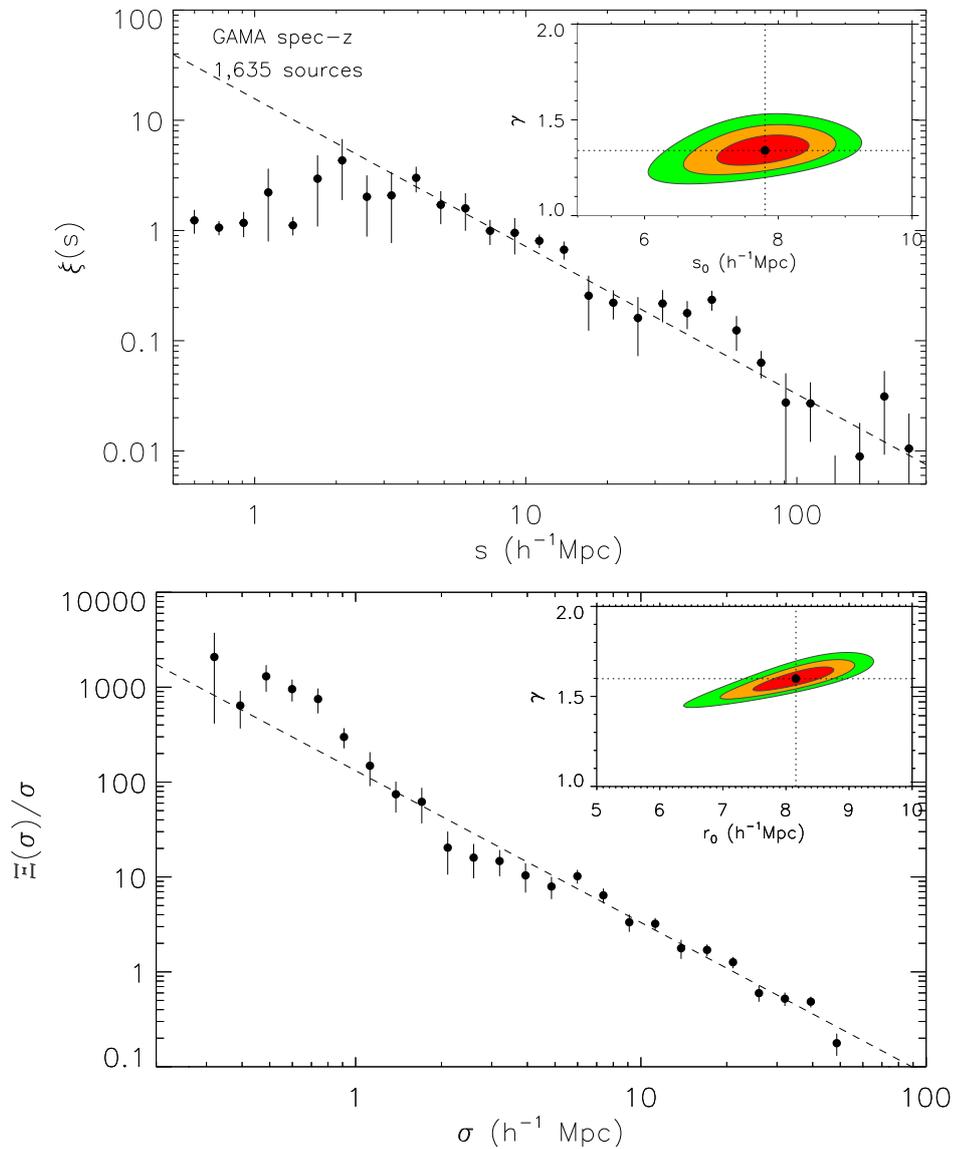
FIGURE 3.7: The redshift-space correlation function (top panel) and projected spatial correlation function (bottom panel) for the GAMA spectroscopically identified objects with bootstrap resampling errors. The dashed line shows the best fit power-law (over the ranges $2 < s < 300 \ h^{-1}$Mpc and $\sigma < 50 \ h^{-1}$Mpc, respectively), and the inset plot shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83. The flattening of $\xi(s)$ at small scales ($s \lesssim 3 \ h^{-1}$Mpc) highlights the effect of the fingers of God discussed in Section 1.2.1.
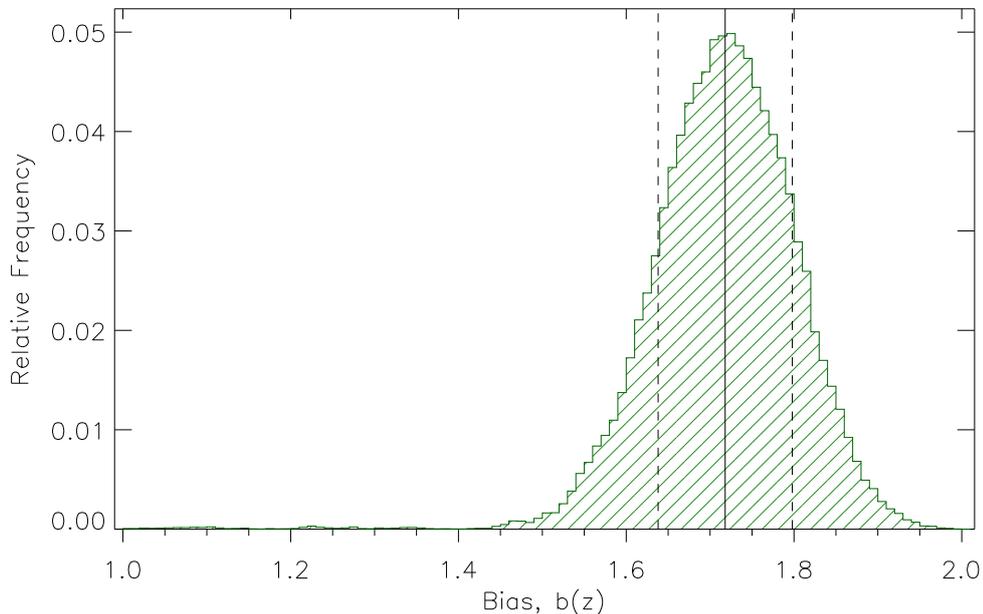
FIGURE 3.8: The posterior distribution of $b(z)$ as estimated from the MCMC power-law fits to $\xi(s)$. The solid vertical line shows the median value, with the dashed lines showing the $\pm 1\sigma$ deviations.

Given the advantages gained by having spectroscopic redshifts, we have measured the redshift-space and projected correlation functions of those radio sources with optical counterparts with spectroscopic redshifts in the GAMA survey itself (Table 3.1 and Figure 3.7). For this sample at $z \simeq 0.34$, we find a redshift-space correlation length $s_0 = 7.80^{+0.40}_{-0.53} \, h^{-1}$Mpc with $\gamma \simeq 1.34$, and correcting for redshift-space distortions, for the projected correlation function we find $r_0 = 8.16^{+0.40}_{-0.46} \, h^{-1}$Mpc with $\gamma \simeq 1.60$.

3.4 Mass Bias

We are able to infer the redshift-space correlation length s_0 from the redshift-space correlation function, or the real-space correlation length r_0 from the projected correlation function (Section 3.3.2). The bias parameter (as a function of redshift) may be defined as equation 1.29, with $r = 8 \, h^{-1}$Mpc. As described in equation 3.11 the numerator can be written

$$\xi_{\text{gal}}(8, z) = \left[\frac{r_0(z)}{8} \right]^{\gamma}. \tag{3.15}$$

The corresponding function for the denominator is given by Peebles (1980) as

$$\xi_{\text{DM}}(8, z) = \sigma_8^2(z)/J_2, \tag{3.16}$$

Function	Range (h^{-1}Mpc)	s_0/r_0 (h^{-1}Mpc)	γ	$b(z = 0.34)$
$\xi(s)$	$2 < s < 300$	$7.80^{+0.40}_{-0.53}$	$1.34^{+0.05}_{-0.05}$	$1.72^{+0.08}_{-0.08}$
$\Xi(\sigma)$	$\sigma < 50$	$8.16^{+0.40}_{-0.46}$	$1.60^{+0.04}_{-0.05}$	$1.92^{+0.10}_{-0.11}$

TABLE 3.1: Clustering parameters found from the redshift-space correlation function $\xi(s)$ and the projected correlation function $\Xi(\sigma)$ of the 1,635 source sample of radio galaxies with spectroscopic redshifts from GAMA.

where $J_2 = 72/[(3 - \gamma)(4 - \gamma)(6 - \gamma)2^\gamma]$ and the parameter σ_8^2 is the matter density variance in a comoving sphere of radius 8 h^{-1}Mpc. The combination of these equations gives the scale-independent evolution of bias with redshift, given only the correlation length and slope:

$$b(z) = \left[\frac{r_0(z)}{8}\right]^{\gamma/2} \frac{J_2^{1/2}}{\sigma_8 D(z)/D(0)}, \tag{3.17}$$

where $D(z)$ is the cosmological growth function. The redshift value is assumed to be the median of the distribution of objects.

3.5 Results and Discussion

Table 3.1 shows s_0 and r_0 derived from $\xi(s)$ and $\Xi(\sigma)$ respectively, with the former being subject to the redshift-space effects described in Section 1.2.1. These correlation lengths describe the clustering strength of the radio galaxies under investigation, with the inferred bias comparing this clustering with that of the underlying dark matter distribution. The fitted correlation lengths are consistent with one another, with the real-space correlation length, $r_0 = 8.16^{+0.40}_{-0.46}\ h^{-1}$Mpc, slightly exceeding the redshift-space equivalent, $s_0 = 7.80^{+0.40}_{-0.53}\ h^{-1}$Mpc. The redshift-space correlation function exhibits a flattening at $s \lesssim 3\ h^{-1}$Mpc, corresponding to a redshift-space distortion effect from peculiar velocities of galaxies within a cluster (see Section 1.2.1), and a shallower slope ($\gamma \simeq 1.34$) than found using $\Xi(\sigma)$ ($\gamma \simeq 1.6$). The MCMC chains for these fitted parameters are maintained in inferring a bias, with 1σ errors determined from the distribution of values (see Figure 3.8). We infer a bias of ~ 1.7 for the redshift-space calculation of the bias, with a real-space equivalent of ~ 1.9 at $z \simeq 0.34$. This shows that radio sources trace a massive (and therefore biased) galaxy population, somewhere between the clustering strength of ordinary local galaxies ($r_0 \sim 5\ h^{-1}$Mpc) and rich clusters ($r_0 \sim 20 - 25\ h^{-1}$Mpc).

The results of this spectroscopic sample lend themselves to comparison with the work of Brand et al. (2005) and Magliocchetti et al. (2004). Brand et al. (2005) directly measure the spatial correlation function (although ignoring redshift-space distortions) of 268 radio galaxies in 165 square degrees of

the Texas-Oxford NVSS Structure (TONS) survey with flux density $S_{1.4} > 3$ mJy and an optical limit of $R \lesssim 19.5$. The slightly higher radio flux limit removes some of the sources associated with star formation, but the brighter optical limits favour low-luminosity radio galaxies at lower redshift, with the median of the sample at $z \sim 0.3$. They find $r_0 = 6.1 \pm 1.1 \ h^{-1}$Mpc, assuming a linear ($\epsilon = \gamma - 1$) clustering model.

Magliocchetti et al. (2004) make joint use of FIRST and 2dFGRS to a radio flux density limit of 1 mJy and an optical limit of $b_J < 19.37$, working with a sample of 820 radio sources with redshifts $0.01 < z < 0.3$ over a larger area of ~ 375 square degrees. Their value of $r_0 \simeq 4.7 \pm 0.7 \ h^{-1}$Mpc is lower than for our sample, but perhaps by no more than expected given their lower redshifts and preferential selection of optically brighter sources. This effect is mitigated by their subsequent selection of only those sources whose spectra have signatures of AGN activity. This increases the redshift from $z \simeq 0.10$ to 0.13, and the correlation length to $r_0 \simeq 7.6 \pm 0.8 \ h^{-1}$Mpc, similar to that of our sample $z \simeq 0.34$, shown in Figure 4.5. We would expect to observe clustering between these two values, albeit boosted due to our sample having a higher median redshift. The close agreement with these past surveys also suggests that our analysis accounting for the noise variations in the GAMA fields is robust.

3.6   Conclusions

Using data from the FIRST survey, with its combination of extensive sky coverage, depth (1 mJy) and resolution ($5.4''$), there is potential for in-depth study of the clustering of radio sources across a range of redshifts. Cross-matching with GAMA, a spectroscopic survey carried out in the optical, provides spectroscopic redshifts for the optical hosts of these radio sources, adding precision to their distance estimates and allowing a direct measurement of the spatial correlation function. Multiple-component sources in groups within $72''$ are collapsed into a single object at the flux-weighted mean position of the grouped sources. Redshifts are attributed to the 16 per cent of radio sources which are within $3''$ of a GAMA source with high quality spectrum ($Q \geq 3$), resulting in a cross-matched sample of 1,635 at $z \simeq 0.34$.

The projected correlation function is used to correct for redshift-space effects in $\xi(s)$ and $\xi(\sigma, \pi)$, with a correlation length of $r_0 = 8.16^{+0.40}_{-0.46} \ h^{-1}$Mpc being found, corresponding to a bias estimated at $b(z = 0.34) = 1.92^{+0.10}_{-0.11}$. The redshift-space analysis finds $s_0 \simeq r_0$, within errors, but a smaller bias value of $b = 1.72 \pm 0.08$.

While these measurements, and improvements on the constraints on the parameters concerned, are

important, only a small minority of the available sources have the accurate redshifts required for a direct measurement of the spatial correlation function, with much of the advantage of these large catalogues being wasted. As a result, only relatively low redshifts are being studied, while it is at high redshifts that a clear understanding of radio galaxy clustering is lacking. Probing multiple epochs, and extending to higher redshifts, can be acheived in a number of ways, some of which will be described and carried out in Chapters 4 and 5.

4

Bias Evolution of FIRST Radio Sources to $z \sim 1.5$ in GAMA/SDSS/UKIDSS[†]

4.1 Introduction

Spectroscopic redshifts, as obtained through the GAMA survey in Chapter 3, add a crucial, accurately-determined third dimension to our observed radio galaxy samples. As a result, the projected spatial correlation function can be used to quantify the clustering strength of such a sample with good precision. This requirement of spectroscopic identification (most likely in the optical) can be very limiting, with spectroscopy (as is apparent from Chapter 2) being far less time efficient than photometric imaging observations for a given area and depth. By studying only those subsamples of radio surveys where an optical counterpart is found with a spectroscopic redshift, we waste the information encoded in the distribution of the majority of sources, which are unmatched, and are restricted, by the depth of the optical survey, to a relatively low-redshift sample. One of the advantages of using radio surveys is that radio emission from powerful AGN traverse vast distances with impunity, unabsorbed by intervening gas and dust, so it is important to make full use of these high-redshift sources, where available.

A further issue arising from the limited number of radio sources with spectroscopic redshifts is that

the sources cannot be divided into redshift-binned samples while preserving adequate signal-to-noise in the measured correlation function. The correlation length and bias at a single epoch are of limited use, and, as discussed in Section 3.5, similar results have been found by other authors at the redshift we reach with GAMA (albeit with smaller samples). The cosmological merit of these studies is best revealed through the observed evolution of galaxy clustering over cosmological time scales.

In this chapter, these discussed shortcomings of requiring spectroscopic redshifts from GAMA are overcome in two ways. Firstly, the angular correlation function is used to quantify the clustering of the radio sources, requiring only a redshift *distribution* to infer a correlation length, and therefore bias (Section 4.4). This relaxes the redshift requirements in the cross-matching process to include data from photometric surveys, deep enough to reach higher redshifts, and numerous enough to allow the binning of matched sources by redshift. Secondly, the SKADS simulated catalogues (Wilman et al., 2008) are used to provide a parent redshift distribution for the FIRST radio sources, such that spatial clustering parameters can still be estimated for those sources that do not have optical counterparts, further extending the redshift range over which the bias can be seen to evolve.

The outline of this chapter is as follows: Section 4.2 describes the surveys and catalogues used, while Sections 4.3 and 4.4 detail the angular correlation function measurements and their deprojection to infer spatial clustering properties. Section 4.5 discusses the linear bias inferred as a result, and Section 4.6 summarizes these results and presents conclusions.

4.2 Data

To infer the spatial clustering parameters of the radio galaxy population, redshifts of the sources (or their redshift distribution, at least) are required. We use spectroscopic redshifts from the GAMA survey, supplemented by photometric redshifts calculated using optical and near infrared photometry from SDSS and UKIDSS (described in Smith et al., 2011). These redshift catalogues are cross-matched with radio sources from the FIRST survey to assign optical counterparts and redshifts to each radio source.

4.2.1 Radio Surveys

The same FIRST catalogue is used as in Chapter 3 (Section 3.2.2), with potential extended sources (grouped within $72''$) collapsed and placed at their flux-weighted mean position. The SKADS simulated catalogue (Wilman et al., 2008) is also used more extensively in this chapter, providing redshift distributions and relative proportions of different radio populations (RQQs, FRIs, FRIIs, star-forming

Field	RA range (°)	Dec range (°)	Optical sources	$\overline{\sigma}$ (deg^{-2})
9h	[128.0, 142.0]	[-2.0, 3.0]	633,229	9,101
12h	[172.5, 186.5]	[-2.9, 2.1]	651,148	9,303
15h	[210.5, 224.5]	[-2.0, 3.0]	653,417	9,336

TABLE 4.1: Spatial boundaries and surface densities ($\overline{\sigma}$) for the three GAMA/SDSS/LAS fields used (comprising both spectroscopic and photometric redshifts). Due to incomplete coverage in one corner of the 9h field, the area is ~ 0.4 square degrees smaller in area than the full rectangular 12h and 15h fields. All subsequent analysis accounts for this.

galaxies, and starbursts).

The SKADS simulation accounts for the galaxy clustering of the radio sources by assigning dark matter halo masses to the various radio source populations, and the catalogue gives positions and observational values for the host galaxies rather than any of their components (found in the Components catalogue). As a result, no collapsing of multiple sources is required as for FIRST data, and the true clustering of the population can be found directly. However, in the interest of consistent treatment of the two radio catalogues, the same collapsing procedure is applied to the SKADS galaxies, reducing the number of sources by ~ 6 per cent representing approximately the proportion of separate, independent galaxies which will be spuriously collapsed.

4.2.2 Optical and Near-Infrared Surveys

The redshift catalogues and cross-matching procedure used in this chapter are the same as those used in Chapter 3 (GAMA spectroscopic redshifts, and SDSS/UKIDSS photometric redshifts). Figure 4.1 shows a comparison between spectroscopic and photometric redshifts, where we find an rms difference of 0.07, and their distributions. The number of redshifts in the combined GAMA/SDSS/UKIDSS catalogue, after star–galaxy separation, is shown in Table 4.1, for each of the three equatorial fields encompassing the GAMA fields.

4.2.3 Redshift Distributions

In order to study the spatial clustering properties of our galaxy samples, we require knowledge of their redshifts, either individually or at least their distribution. We have spectroscopic or photometric redshifts for those FIRST radio sources with optical identifications, but do not have direct distance information for the optically unidentified radio sources. To estimate the complete redshift distribution
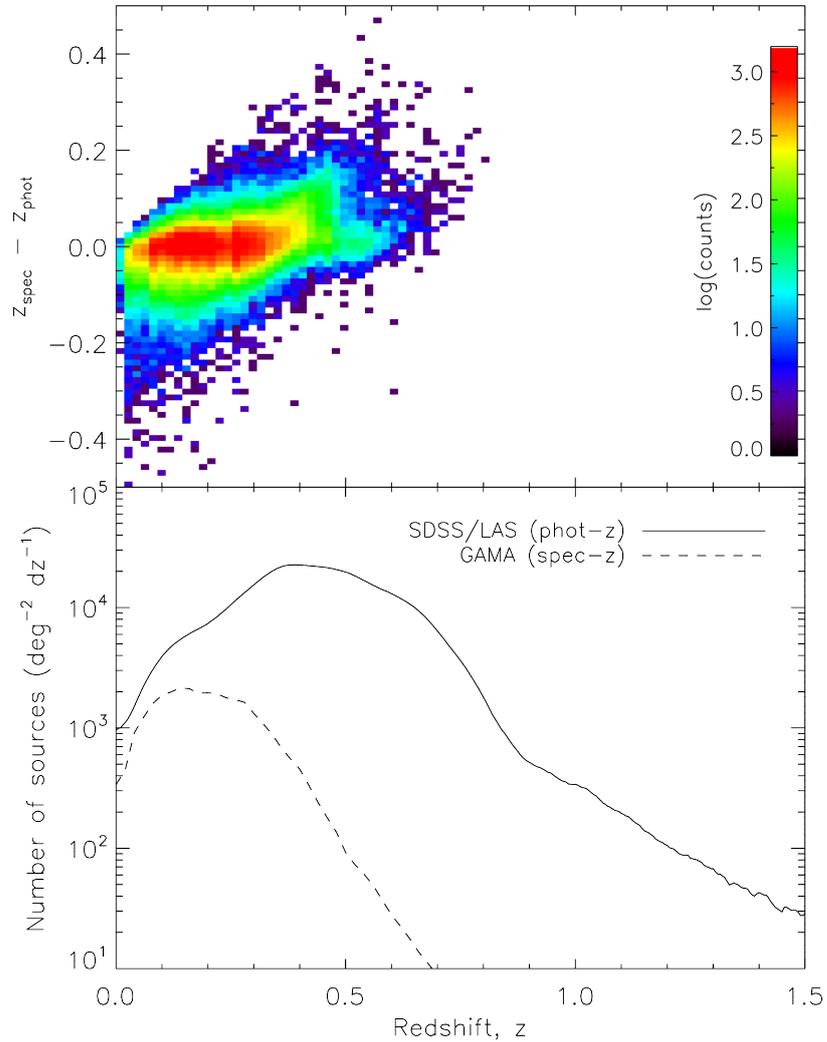
FIGURE 4.1: A comparison of SDSS/LAS photometric redshift estimates with GAMA spectroscopic ($Q \geq 3$) redshifts (top panel). The bottom panel shows the redshift distributions of the full photometric and spectroscopic catalogues (galaxies only).

Catalogue	No. of sources	Area (deg^2)	$\overline{\sigma}$ (deg^{-2})
FIRST	585,473	$\sim 10,000$	~ 58.5
GAMA	3,886	210	18.5
SKADS	32,061	398	80.6

TABLE 4.2: Areas and surface densities of $S_{1.4} > 1$ mJy radio sources in the FIRST catalogues, the GAMA counterparts to the radio sources and also for the SKADS simulated data set. Note that the low-density of sources in the whole (post-collapse) FIRST catalogue compared to the prediction from SKADS may be explained by the incompleteness at < 2 mJy (Sections 3.2.6 and 4.4.2) and the effects of resolution bias (Section 4.2.3).
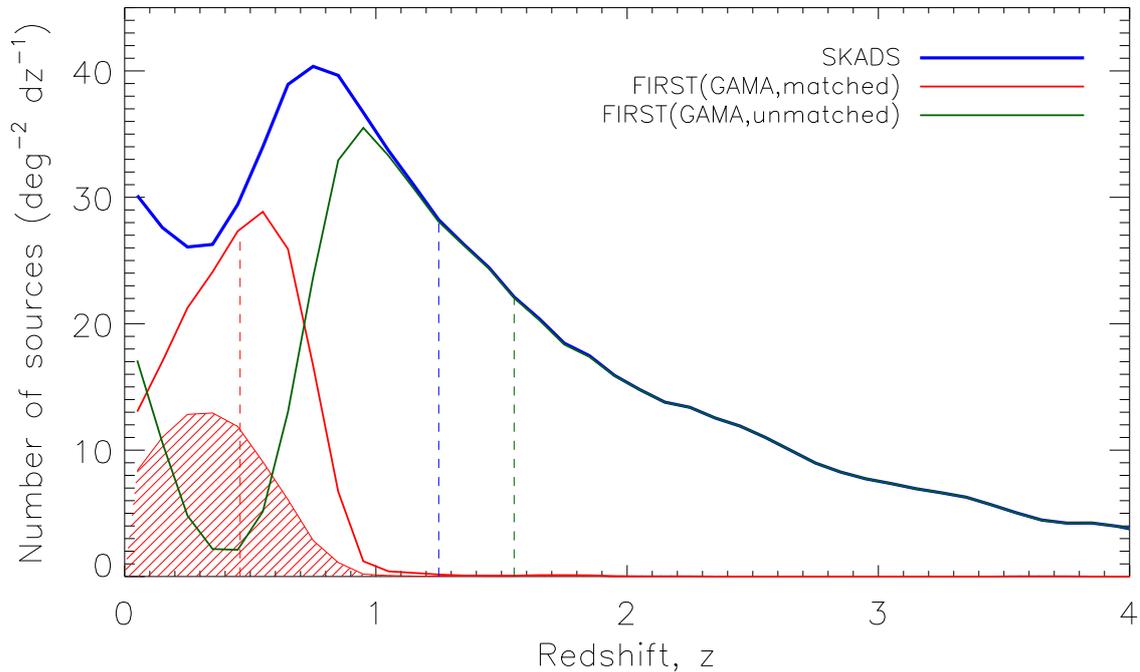
FIGURE 4.2: Redshift distributions of the 1 mJy SKADS (blue) and FIRST–GAMA–matched catalogues (red), where the SKADS $N(z)$ is assumed to be the redshift distribution of FIRST sources. The distribution of the radio sources not identified in GAMA (green) is inferred assuming a SKADS-like parent distribution and subtracting the GAMA distribution. Dotted lines mark the median redshifts for each set of objects, and the filled distribution describes the GAMA spectroscopic redshifts (i.e. removing SDSS/LAS photometric redshifts).

as a function of radio flux-density limit, we use the SKADS simulation, which is based on a range of observed luminosity functions (see Wilman et al., 2008, for full details).

Figure 4.2 shows the redshift distribution of the SKADS catalogue at 1 mJy and the cross-matched subset of FIRST. By making the assumption that FIRST sources (the wider catalogue as well as just those within the GAMA fields) should have a similar distribution to that from SKADS, we may compare directly the clustering of the real and simulated catalogues. Furthermore, as shown in Figure 4.2, a third distribution can be inferred by subtracting the cross-matched distribution from the SKADS "parent" distribution (following similar work by Passmoor et al. 2013). This describes the remaining unmatched radio sources with no direct redshift measurements.

It is clear that the distribution of unmatched sources shows an up-turn towards low-redshift ($z <$ 0.2), where there appears to be a large fraction of sources in the SKADS simulation that are either not detected in FIRST or are in FIRST but do not have a counterpart in GAMA. The latter explanation of this is unlikely, as we would expect that the majority of relatively bright radio sources at $z < 0.2$ to have a reasonably bright optical counterpart. We therefore suggest that these sources, which are dominated by star-forming galaxies in the SKADS simulation, and are predicted to be there based on

low-redshift far-infrared luminosity functions (Wilman et al., 2010), are likely to be resolved out by the VLA observations in the B-Array configuration used for the FIRST survey (Jarvis et al., 2010).

Simpson et al. (2012) also noted a similar deficit of low redshift sources compared with SKADS predictions in their radio survey of the Subaru/XMM-Newton Deep field, and attributed this to the effects of resolution bias. This resolution bias refers to the number of faint resolved sources missing from a peak-flux-density-limited survey because their extended emission is not detected by the radio interferometer. Bondi et al. (2003) performed detailed simulations to determine the effect of this bias on their deep VLA catalogue (used in Chapter 5), and their estimates suggest that around 25 per cent of sources could be missing due to this resolution bias, and that these would be concentrated at low redshift. Any low-redshift star-forming galaxies that have radio flux-density detectable in FIRST are unlikely to fall below the SDSS depth ($r < 22$), although low-surface brightness optical incompleteness may play a role if the radio data were deeper.

Thus we suggest that the discrepancy between the simulated redshift distribution and the observed redshift distribution at low redshift is consistent with being due to resolution bias. We also note that some of the sources in our cross-matched sample could be removed by our removal of sources with photometric redshifts $z < 0.002$.

4.2.4 Completeness

The dearth of low-redshift radio matches, which make up a minority of the sources at ~ 1 mJy, is unlikely to be caused by incompleteness in the FIRST survey, which is corrected for in Section 3.2.6. One would expect to preferentially lose the higher-redshift AGN at around the flux-density limit. For example, the SKADS simulation features ~ 5 low-redshift star-forming galaxies per square degree compared to ~ 30 AGN at higher redshifts, in the flux-density range of 1–2 mJy.

4.2.5 Sample Catalogues

From the 13,346 radio sources in the expanded GAMA fields used in this chapter, the cross-matching process described in Section 3.2.5 leaves 3,886 (29 per cent) of the original FIRST radio sources, of which 422 correspond to collapsed multi-component sources, an estimated 78 are spurious and 1,635 (42 per cent) have good quality GAMA spectroscopic redshifts.

Table 4.3 lists the various samples to be investigated, comprising one large FIRST sample and 5 subsamples of FIRST sources within the GAMA fields. The large FIRST sample is over a single $115° \times 64°$ patch ($127.5° < \alpha < 242.5°$; $0° < \delta < 64°$), while two smaller subsets contain all FIRST
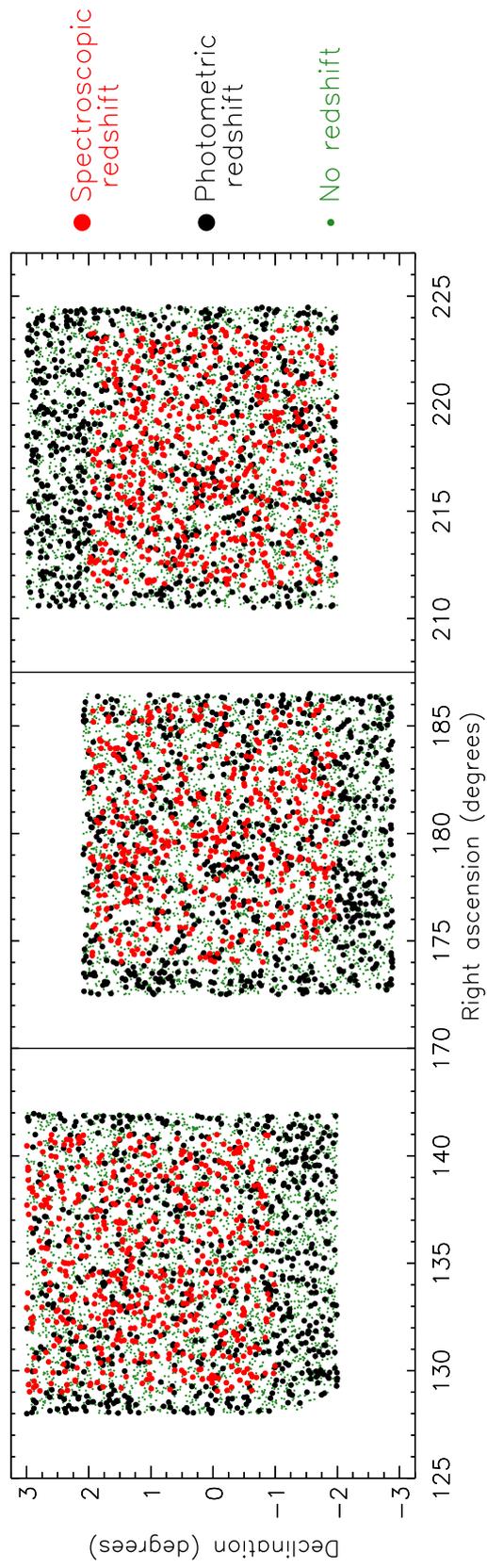
FIGURE 4.3: The angular distribution of the radio samples used. Optically identified FIRST sources are plotted in red (GAMA spectroscopic redshifts; see Chapter 3) or black (SDSS/UKIDSS photometric redshifts), and those without optical matches are plotted in green.

Sample	Area (deg^2)	No. of sources	Redshift	
			Limit	z_{med}
FIRST–all	5,922	342,615	-	1.21
FIRST–GAMA fields	210	13,346	-	1.21
FIRST–GAMA–unmatched	210	9,460	-	1.55
FIRST–GAMA–matched	210	3,886	-	0.48
FIRST–GAMA–matched	210	2,156	$z < 0.5$	0.30
FIRST–GAMA–matched	210	1,730	$z > 0.5$	0.65

TABLE 4.3: A summary of the samples for which the clustering properties will be measured: GAMA/SDSS cross-matched radio sources (further split into $z > 0.5$ and $z < 0.5$ samples), a large sample of FIRST sources, a smaller sample of all FIRST sources within the GAMA fields, and the remaining FIRST sources in these fields after removing GAMA cross-matches (all post-collapse of multiple sources). Redshift distributions are shown in Figure 4.2, and the enlarged GAMA area footprint is shown in Figure 4.3.

sources found within the enlarged GAMA footprint defined in Table 1 (FIRST–GAMA fields) and only those which are unmatched with GAMA sources (FIRST–GAMA–unmatched), shown in Figure 4.3. The matched GAMA sources are also subsequently split into low- and high-redshift subsets in order to attempt to demonstrate any redshift evolution. The optical cross-matching preferentially identifies low-redshift radio sources, allowing us to use the unmatched sources to probe higher redshifts ($z_{\mathrm{med}} = 1.55$) than the matched or full samples, further helping to observe any redshift-dependence of their clustering.

4.3 Angular Correlation Function

In the absence of sufficiently accurate distances to individual sources (from spectroscopic redshifts), the clustering of sources in an angular distribution can be described by the angular two-point correlation function, $w(\theta)$, defined in Section 1.2.1. The same estimators can be used as for the spatial correlation function (Section 3.3), using an unclustered distribution of sources for comparison, with the chosen estimator of Landy & Szalay (1993) being defined similarly as:

$$w(\theta) = \frac{n_r(n_r - 1)}{n_d(n_d - 1)} \frac{DD}{RR} - \frac{(n_r - 1)}{2n_d} \frac{DR}{RR} + 1, \tag{4.1}$$

where n_d and n_r are the number of real sources and random sources.

Errors on $w(\theta)$ are calculated analogously to those on $\xi(s)$ (described in Section 3.3), via the bootstrap resampling method. Likewise, the integral constraint due to limited area of the samples is
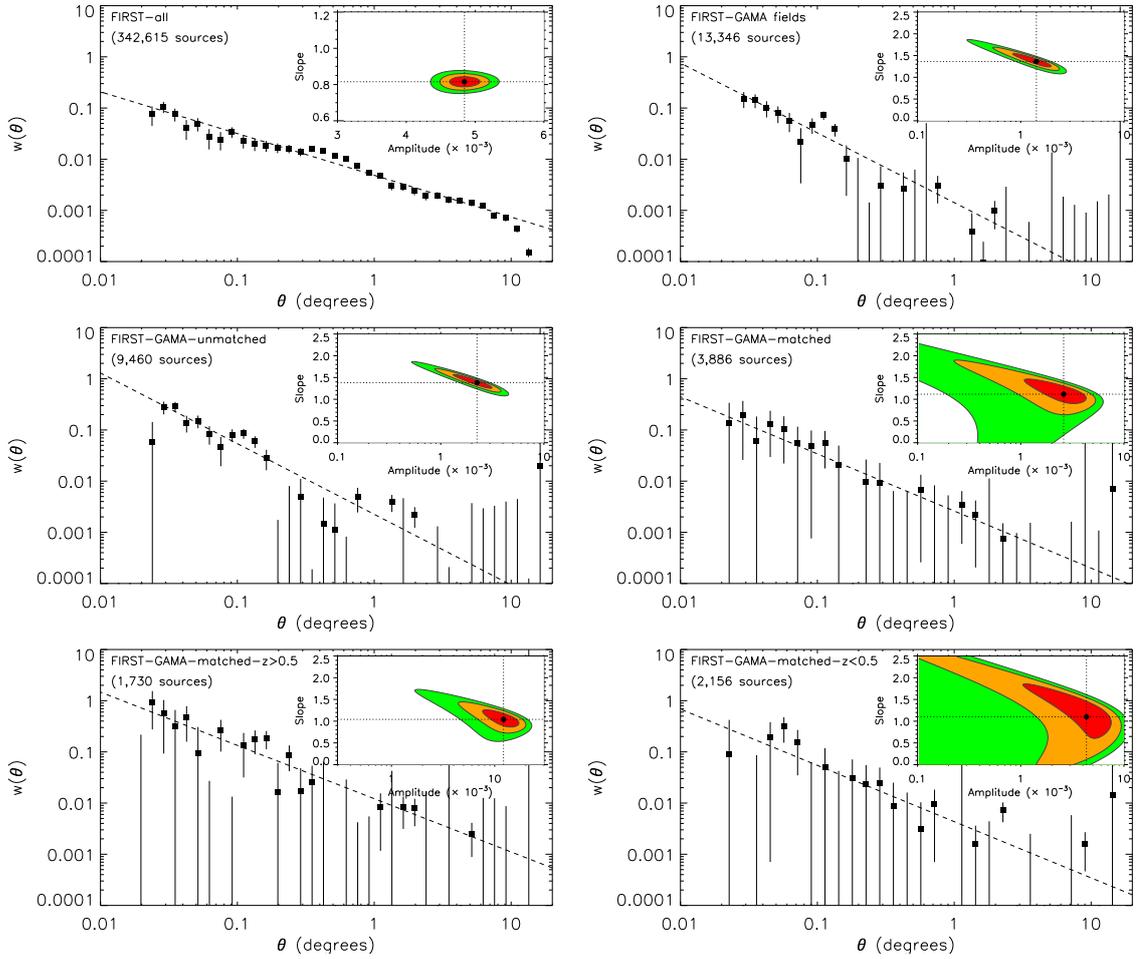
FIGURE 4.4: The angular correlation function for the 5 samples (sample name is denoted at the top left corner of each panel) within the extended GAMA fields (with bootstrap resampling errors) and the larger FIRST sample in the bottom right panel. The dashed lines show the best fit power-law (over the range $0.02° < \theta < 10°$) and the inset contour plots shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83.

described as follows:

$$w_{\mathrm{obs}}(\theta) = w(\theta) - \sigma^2, \tag{4.2}$$

where σ^2 represents the integral constraint, approximated by

$$\sigma^2 = \frac{\sum RR(\theta)w(\theta)}{\sum RR(\theta)}. \tag{4.3}$$

Traditionally, $w(\theta)$ has been fitted by a power law (e.g. Peebles, 1980) with a slope of ~ -0.8 commonly found for the clustering of objects of various masses (Bahcall & Soneira, 1983). While radio sources have been found to fit a distinct double power law (cf. Blake & Wall, 2002a; Overzier et al., 2003), since we have collapsed the multiple component sources, $w(\theta)$ should reduce to the canonical single power law form for our data.

We fit $w(\theta)$ with a single power law function of the form $w(\theta) = A\theta^{1-\gamma}$ (where γ is equivalent to that used in equations 1.14 and 3.11), fitted over the range $0.02° < \theta < 10°$. This is done using the Metropolis-Hastings algorithm to obtain a Markov Chain Monte Carlo (MCMC) array of 10^6 data points in γ–log(A) space. Parameter values quoted are minimum-χ^2 values and 1σ errors correspond to the region containing 68.3 per cent of the MCMC points. A Levenberg-Marquardt χ^2 minimization routine yields the same best fit values, and ordinary χ^2 contours coincide closely with those from the MCMC simulations.

Figure 4.4 shows the results of this angular correlation function method for the 5 samples of FIRST radio sources within the GAMA regions, and one larger FIRST sample, and we see clearly in the likelihood contours that the errors are smaller for the more numerous unmatched sources than the matched sources. The best fit parameters are shown in Table 4.4. The unmatched sample has a lower amplitude than the matched sample, but given these will come to represent complementary high- and low-redshift measurements, respectively, similar amplitudes do not imply similar clustering scales due to the increased angular diameter distance at high redshift.

4.4 Limber Inversion

If the redshift distribution of a set of objects is known, one may deproject the angular correlation function into the spatial correlation function (Section 1.2.2). This is the purpose of inverting the cosmological Limber equation (Limber, 1953; Peebles, 1980) to estimate the spatial correlation length, r_0 (discussed further in Section 3.3). This is often more useful than computing the spatial correlation function, $\xi(r)$, directly, as a complete set of individual redshifts is rarely available for a given survey, thus requiring the redshift distribution to be estimated via the luminosity function in order to deproject $w(\theta)$. Using the redshifts available for the objects in our catalogue, however, means we may apply a distribution directly from the data for the radio sources with optical counterparts.

The spatial correlation function is taken to be of the form described in equation 3.10, defined by the correlation length, slope, and a clustering index. The slope, γ, is the same as that used in the power law fit to the angular correlation function (where the magnitude of the slope is $\gamma - 1$; see Table 1.1), so we measure this parameter through the $w(\theta)$ function. The amplitude A of $w(\theta)$ has been expressed as a function of r_0 (in comoving coordinates) in the literature (Overzier et al., 2003; Kovač et al., 2007;

Kim et al., 2011; Elyiv et al., 2012), based on equation 1.20, as follows:

$$A = r_0^\gamma H_\gamma \left(\frac{H_0}{c}\right) \frac{\int_0^\infty N^2(z)(1+z)^{\gamma-(3+\epsilon)}\chi^{1-\gamma}(z)E(z)\mathrm{d}z}{\left[\int_0^\infty N(z)\,\mathrm{d}z\right]^2}, \tag{4.4}$$

where $N(z)$ is the redshift distribution, and $\chi(z)$ and H_γ are defined in Section 3.3, respectively. Equation 4.4 may simply be inverted to give the comoving correlation length, r_0 as a function of redshift distribution, correlation function slope, and angular clustering amplitude.

We employ redshift distributions as measured for the GAMA matched sources, and for the simulated SKADS sources, as detailed in Section 4.2.3. We also emphasize that using the SKADS $N(z)$ allows us to account for sensitivity variations across the FIRST survey as we are able to remove the correct number of sources of a given flux and redshift (as described in Section 3.2.6).

4.4.1   Results of Limber Inversion

We have measured $w(\theta)$ for the six samples described above, and fitted them with a power law (see Table 4.4 and Figure 4.4). Given the dependence on ϵ, the clustering index of choice, r_0 is shown in Figure 4.5 for the comoving ($\epsilon = \gamma - 3$) and linear ($\epsilon = \gamma - 1$) clustering models. The latter consistently gives lower values for r_0, with the unshown stable ($\epsilon = 0$) model values being between the two, but these differences tend to be comparable with the associated errors. The results of the comoving clustering model will be quoted hereafter, unless stated otherwise.

The large FIRST sample yields a $w(\theta)$ consistent with the literature, where we find a slope of $\gamma - 1 = 0.82 \pm 0.02$, the collapsing of close pairs successfully having removed evidence of a steeper, small-scale power law component. By assuming the same redshift distribution as a similar SKADS catalogue (adjusted in accordance with the noise variations in the FIRST survey; see Sec. 3.2.6), we find a correlation length of $8.20^{+0.41}_{-0.42}\,h^{-1}\mathrm{Mpc}$ at a median redshift of $z = 1.21$

We are able to infer the clustering properties from our real data within the GAMA fields at 3 different epochs utilizing the three subsamples described in Section 4.2.3 and Figure 4.2. These 3 subsamples have distributions of median $z = 0.48$, 1.21 and 1.55, for the matched, complete, and unmatched radio samples, respectively.

In each case, the $w(\theta)$ power law is found to be significantly steeper ($\gamma > 2$) than for the larger FIRST sample, albeit less well constrained owing to the smaller sample sizes. This steeper power law is borne out in the analysis where we find that the " FIRST–GAMA fields" sample has a larger clustering length than the larger FIRST sample with the same $N(z)$. The trend in these 3 subsamples, however, is for increasing r_0 with redshift. This is confirmed by analysing subsamples of the cross-matched objects.

Sample	N_{obj}	$A\ (\times10^{-3})$	γ	z_{med}	$r_0\ (h^{-1}\mathrm{Mpc})$		$b(z=z_{\mathrm{med}})$	
					$\epsilon=\gamma-1$	$\epsilon=\gamma-3$	$\epsilon=\gamma-1$	$\epsilon=\gamma-3$
FIRST–all	342,615	$4.87^{+0.12}_{-0.17}$	$1.82^{+0.02}_{-0.02}$	1.21	$11.82^{+0.43}_{-0.46}$	$8.20^{+0.41}_{-0.42}$	$4.39^{+0.19}_{-0.19}$	$3.14^{+0.16}_{-0.17}$
FIRST–GAMA fields	13,346	$1.31^{+0.35}_{-0.34}$	$2.35^{+0.11}_{-0.09}$	1.21	$11.27^{+0.82}_{-0.87}$	$10.53^{+0.64}_{-0.68}$	$6.40^{+0.42}_{-0.45}$	$5.95^{+0.48}_{-0.53}$
FIRST–GAMA–unmatched	9,460	$2.21^{+0.62}_{-0.57}$	$2.39^{+0.11}_{-0.10}$	1.55	$14.41^{+1.23}_{-1.46}$	$13.60^{+0.83}_{-1.07}$	$10.06^{+0.49}_{-0.50}$	$9.45^{+0.58}_{-0.67}$
FIRST–GAMA–matched	3,886	$2.76^{+0.94}_{-1.11}$	$2.15^{+0.12}_{-0.24}$	0.48	$8.24^{+1.75}_{-2.36}$	$6.72^{+1.81}_{-2.17}$	$2.65^{+0.98}_{-0.90}$	$2.13^{+0.90}_{-0.76}$
FIRST–GAMA–matched–z<0.5	2,156	$4.28^{+1.99}_{-1.89}$	$2.28^{+0.24}_{-0.51}$	0.30	$6.21^{+2.34}_{-4.34}$	$5.39^{+2.53}_{-3.92}$	$1.77^{+1.62}_{-1.23}$	$1.52^{+1.57}_{-1.07}$
FIRST–GAMA–matched–z>0.5	1,730	$12.26^{+2.90}_{-3.04}$	$2.04^{+0.12}_{-0.12}$	0.65	$17.07^{+4.50}_{-4.70}$	$10.67^{+3.22}_{-3.18}$	$5.74^{+2.56}_{-1.97}$	$3.56^{+1.59}_{-1.23}$

TABLE 4.4: Clustering parameters measured covering real and simulated radio sources ($S_{1.4} > 1$ mJy) with real (GAMA) and/or simulated redshift distributions (FIRST). Errors are 1σ from 10^6 (A, γ) and 10^5 (r_0, b) MCMC calculations.
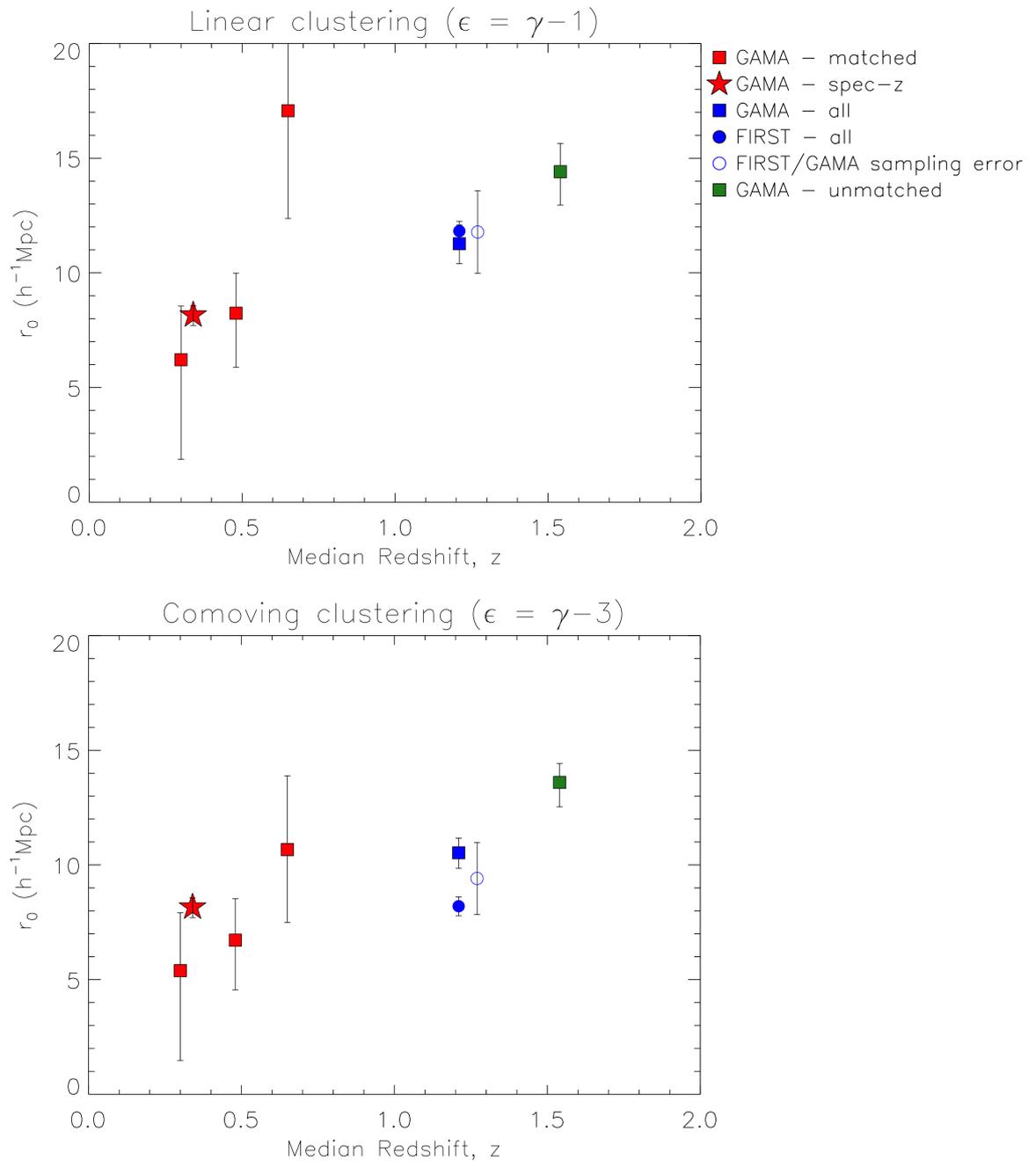
FIGURE 4.5: Correlation length for the observed radio samples described in Table 4.4 as calculated for two different clustering indices. Filled square symbols correspond to samples over the GAMA/SDSS/UKIDSS fields, while the circle symbols refer to the wider FIRST survey, and the star symbols correspond to the spectroscopic sample from Chapter 3. Red, green and blue colours refer to matched, unmatched, and total samples respectively (cf. Figure 4.2). The open circle beside the points at $z \sim 1.2$ shows the mean value and error associated with 20 independent patches of the FIRST survey, providing an estimate of the sample variance within a field size of 210 square degrees.

Small number counts limit us to taking a simple high- and low-redshift approach, but even within this redshift range of $z \simeq 0.3$ to $z \simeq 0.65$, r_0 is seen to increase.

In order to assess the effect of sample variance on our results over the GAMA fields, we also find $w(\theta)$ and r_0 for 20 separate 210-square-degree patches of the FIRST survey over a large range of declinations ($-5° < \delta < +55°$), assuming the SKADS $N(z)$ for each. We find a mean and standard deviation of correlation lengths from these subsamples to give $r_0 = 9.41 \pm 1.57\ h^{-1}$Mpc, lying between the two values found for our GAMA sample and for the larger FIRST sample at $z \simeq 1.21$, and placing both within this cosmic variance error.

4.4.2 Discussion

Using the angular clustering approach, without individual redshifts, Blake & Wall (2002a) find a clustering length for $S > 10$ mJy NVSS sources of $r_0 \sim 6\ h^{-1}$Mpc (independent of flux density limit) under the assumptions of $N(z)$ calculated from luminosity functions of Dunlop & Peacock (1990) and an Einstein-de Sitter (EdS) cosmological model. Likewise, Overzier et al. (2003) use NVSS and similarly-derived $N(z)$ from Dunlop & Peacock (1990), finding $r_0 \simeq 5 \pm 1\ h^{-1}$Mpc for flux density between 3 and 40 mJy. Both papers probe clustering at $z \sim 1$ and, while they differ slightly from one another, they are both considerably lower than our findings even with our lower flux limit diluting the clustering due to AGN. It is important to note, however, that the Dunlop & Peacock models used are poorly constrained at these low flux density limits due to a lack of volume coverage at low redshift in their combined surveys. Improvements in deep surveys have allowed Wilman et al. (2008) to better constrain low-power AGN and star-forming galaxies in the SKADS redshift distribution for use at the mJy level.

Any underestimation of r_0 could partly be a result of having fixed the angular clustering power law to a slope of -0.8, as is widely observed for normal galaxies, although even our larger FIRST sample with $\gamma - 1 = 0.77$ appears to exhibit stronger clustering. The common assumption of a slope of -0.8 reduces the power law fitting to a one-parameter problem, but a steeper slope (~ -1.0 to -1.4) may give a higher correlation length than found under the assumption of a universal slope of -0.8. It is likely, however, that these differing results are due mostly to our lower flux density limits and slightly higher median redshift.

Indeed, while our $w(\theta)$ fit for the larger FIRST sample does yield the expected $\gamma - 1 = 0.8$, Cress et al. (1996) find a higher value, similar to our smaller subsamples, of $\gamma - 1 \simeq 1.1$ yielding a larger correlation length at the same $z \sim 1$ of $r_0 \simeq 10\ h^{-1}$Mpc, more in keeping with our results.

Furthermore, this epoch, and corresponding correlation length, is in close agreement with observations of $z \sim 1$ extremely red objects (EROs) using the same $\epsilon = \gamma - 3$ clustering model (e.g. Daddi et al., 2001, 2002; McCarthy et al., 2001; Roche et al., 2002). This ties in with the suggestion by Willott, Rawlings & Blundell (2001) that EROs and high-redshift galaxies are different evolutionary stages of the same galaxy population, which in turn may be the progenitors of local bright ellipticals.

Two studies of spectroscopic samples of radio sources associated with luminous red galaxies (LRGs) by Wake et al. (2008) and Fine et al. (2011) investigate samples at a similar range of redshifts below $z = 0.8$ by different methods. Wake et al. (2008) fit a halo occupation distribution (HOD) model to the projected spatial correlation function of radio-detected ($L_{1.4} \gtrsim 10^{24.2}$ WHz^{-1}) LRGs at $z \simeq 0.55$ to give $r_0 \simeq 12 \ h^{-1}$Mpc, slightly greater than we find (Figure 4.5) due to their selection of the most optically luminous, more powerful AGN sources. Fine et al. (2011) show the evolution in this clustering by adding LRG samples from SDSS ($z \simeq 0.35$) and AAOmega ($z \simeq 0.68$). Over this very similar redshift range to our optically-identified radio sources, these authors find r_0 using angular auto- and cross-correlation functions which are consistent with no redshift evolution, falling in the range of ~ 10–$12 \ h^{-1}$Mpc. While we find a similar correlation length towards higher redshift, this result is contrary to the trend we observe with the GAMA/SDSS radio sources. Again, this could simply be a result of the different optical selection favouring considerably more powerful AGN at lower redshift, whereas our magnitude limit allows more of the low optical luminosity sources into our sample.

Looking at the AGN population from the other end of the electromagnetic spectrum, Elyiv et al. (2012) use a $N(z)$ derived from a range of luminosity functions for AGN detected in X-rays with the XMM–LSS, recovering the ubiquitous slope of -0.8 for the soft band sources, but -1.0 for the hard band. Their hard band sources exhibit a clustering length of $r_0 \simeq 10 \pm 1 \ h^{-1}$Mpc at $z \sim 1$ using a clustering index of $\epsilon = -1.2 \simeq \gamma - 3$, roughly corresponding to comoving clustering evolution. Accounting for the slightly lower redshift, this agrees well with a FIRST radio population expectedly dominated by AGN emission.

Surface Density Variations and Cosmic Variance

Figure 4.6 shows the surface density of FIRST varying with declination (more notably at lower flux-densities), highlighting a likely cause for some discrepancy between clustering measures over the wider FIRST survey and those same sources only within the GAMA fields. Plotting the surface density contrast ($\delta\sigma/\bar{\sigma}$) against declination highlights a feature at $\delta \lesssim 5°$ where the survey has about 10 per cent fewer sources than average with flux densities between 1 and 10 mJy. Above 10 mJy there is
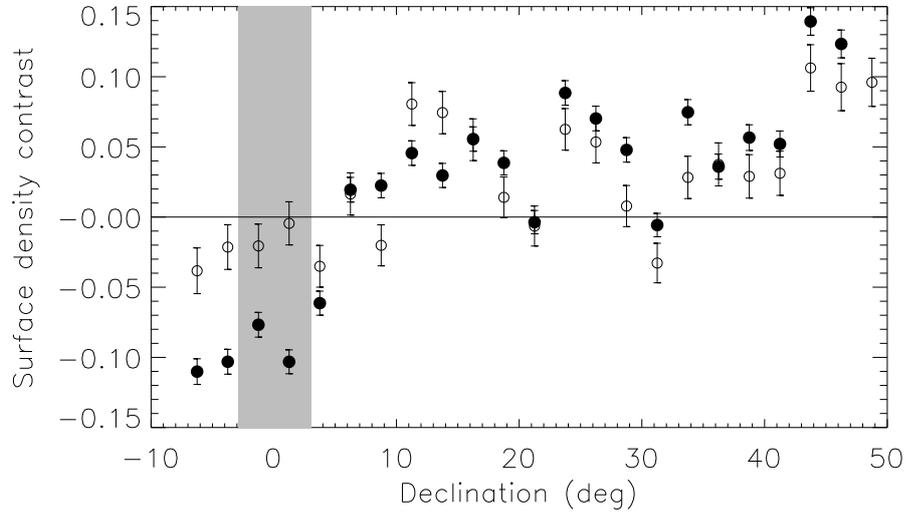
FIGURE 4.6: Surface density contrast of the collapsed FIRST catalogue against declination for sources with $1 < S_{1.4} < 10$ mJy (solid circles) and >10 mJy (open circles) with error bars of $\sqrt{N}$ for N sources in each bin. At declinations below $\sim 5°$ there is a dearth of sources at lower flux densities, affecting the GAMA fields (shaded).
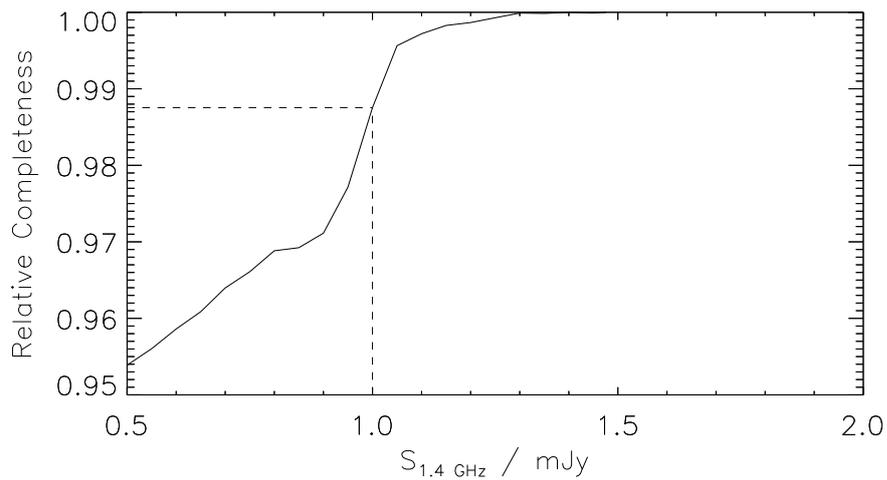
FIGURE 4.7: Estimated ratio of completeness curves for the FIRST survey over the three equatorial GAMA/S-DSS fields and a large area of the Northern sky. Higher rms noise values at lower declinations make the GAMA fields less complete at the lowest flux densities.

no such pronounced feature, meaning the equatorial regions containing the GAMA fields are missing a significant number of sub-10 mJy sources. Likewise, estimating the completeness of FIRST using noise values from the survey and model flux density distribution from SKADS, we find the noise in the GAMA regions to be such that roughly 1 per cent fewer sources are detected above 1 mJy than would typically be detected in the full survey (see Figure 4.7).

As a result, we might expect the clustering statistics to reflect a larger proportion of radio-loud AGN rather than low-redshift star-forming galaxies, which become more dominant at the mJy level. We demonstrate this effect by calculating r_0 and bias for 20 patches of FIRST, each 210 square degrees in area, over a declination range between $-6°$ and $+54°$. The open circles in Figure 4.5, slightly offset in z for clarity, represent the mean and standard error on these quantities. The GAMA fields yield results at the high end of what bias we might find based on the surface density fluctuation, while the larger FIRST sample at $\delta > 0$ is towards the lower end, however we note that both are consistent with the expected variance with patches of 210 square degrees.

4.5 Mass Bias

We are able to infer the real-space correlation length r_0 from the projected correlation function, as was the purpose of Section 3.3.2, or from the angular correlation function (Section 4.4). The bias parameter (as a function of redshift) may be defined as equation 1.29, and related to r_0 as in Section 3.4:

$$b(z) = \left[\frac{r_0(z)}{8}\right]^{\gamma/2} \frac{J_2^{1/2}}{\sigma_8 D(z)/D(0)},\tag{4.5}$$

where, once again, the redshift value is assumed to be the median of the distribution of objects in a given sample.

4.5.1 Results

The bias inferred from the angular correlation function is shown in Table 4.4 for both ϵ values corresponding to linear and comoving clustering evolution. These results are compared in Figure 4.8 with the bias model of Mo & White (1996), used by Wilman et al. (2008) in populating dark matter haloes in the SKADS simulations. They assign a particular halo mass to each source type in the simulation: FRI, FRII and radio-quiet AGN, normal star-forming galaxies, and starbursts, with assumed halo masses of 10^{13}, 10^{14}, 3×10^{12}, 10^{11} and $5 \times 10^{13} \ h^{-1} M_\odot$, respectively, and impose a plateau for each model at high redshifts, above which the assumption of a fixed halo mass breaks down. The model $b(z)$ for
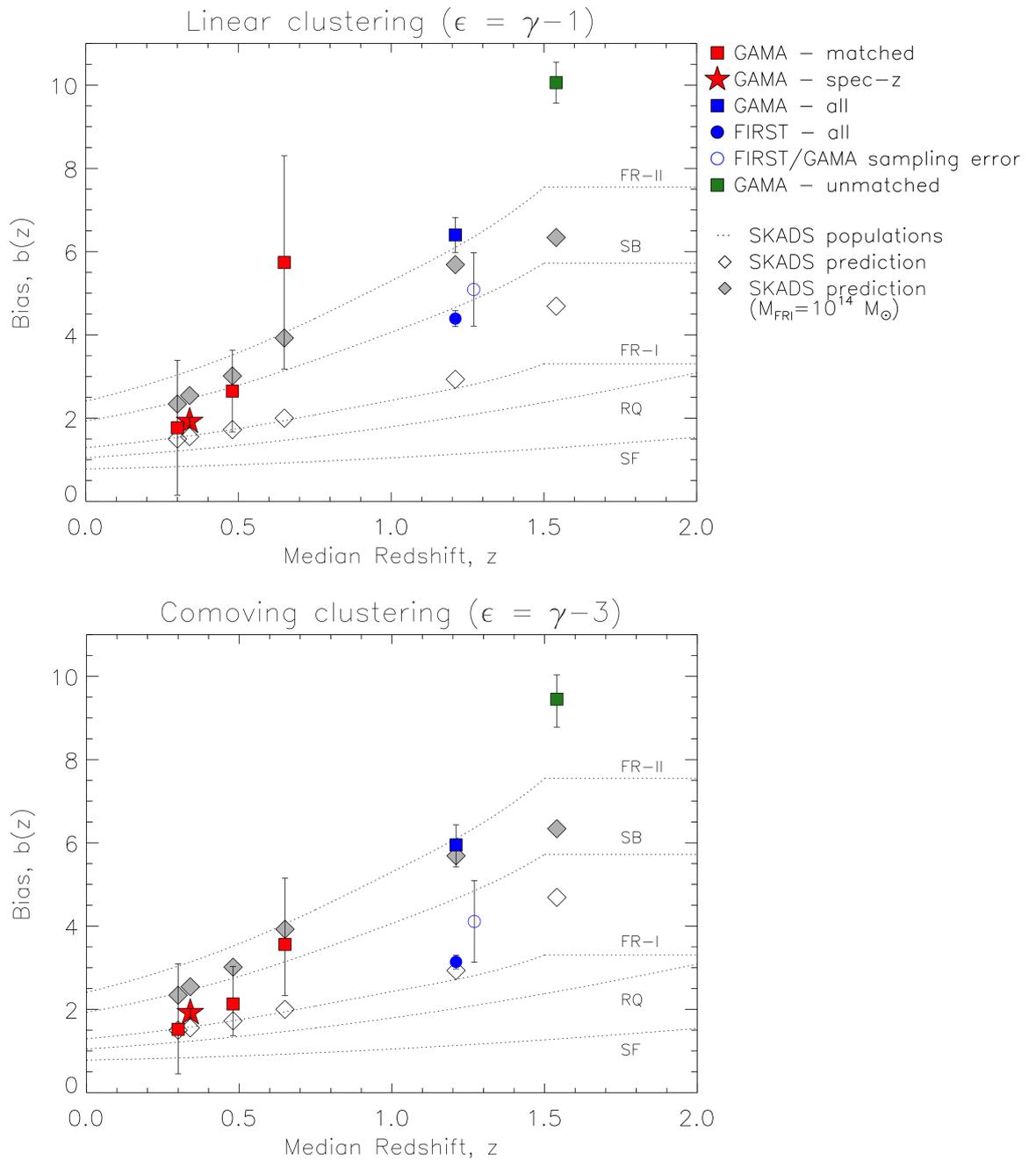
FIGURE 4.8: Linear bias parameter for the observed radio samples described in Table 4.4 as calculated for two different clustering indices, with symbols as defined in Figure 4.5. Overplotted are lines showing the model bias evolution based on the halo masses assigned in the SKADS simulation of individual source populations: FRI, FRII and radio-quiet (RQ) AGN, starbursts (SB), and normal star-forming (SF) galaxies. Open diamonds corresponding to each point mark the aggregate bias expected of the samples, assuming relative population abundances and masses used in redshift-matched SKADS samples. The grey filled diamonds show this same model bias if we increase the assumed halo mass of FRI sources to equal that of the FRII sources.

each of these source types is denoted by a black dotted line, showing the stronger bias for those objects residing in more massive haloes.

By weighting the SKADS sources by the ratio between the sample and SKADS redshift distributions, we have estimated the relative proportions of the population masses. Calculating the model bias of the sample as the weighted mean of the biases of each individual population, we show a predicted bias at the 5 different redshifts probed in Figure 4.8 (*open diamonds*), accounting for the change in galaxy halo mass being observed. These follow the same trend as our measurements, with the bias rising more steeply with redshift than any individual model population owing to the increasing AGN fraction towards higher z. Our measurements over the GAMA fields agree qualitatively with the model, but with those associated only with the GAMA fields exceeding the predictions to varying degrees. For those sources with measured redshifts, the difference is within the stated errors, but at higher redshifts our measurements exceed the SKADS bias values by $\sim 3\sigma$ (we return to this point in Section 4.5.2).

4.5.2   Discussion

Within our matched sample and incorporating simulated data, we observe the increasing of $b(z)$ in line with models corresponding to fixed masses (Mo & White, 1996; Matarrese et al., 1997), although we find a relatively high value for the high-redshift sources, particularly when using the linear clustering parameter. Whilst the GAMA data points do not place especially tight constraints on this evolution, and the well-constrained point from $\Xi(\sigma)$ is limited to one epoch (Chapter 3), the assumption of the redshift distributions for the full and unmatched subsamples probe redshifts up to $z \simeq 1.55$ and give more shape to the evolution. The agreement is good at lower redshifts, the GAMA-matched results (*black*) lying $< 1\sigma$ from the model prediction as well as agreeing with the 2SLAQ radio-detected LRGs examined by Wake et al. (2008) ($b \simeq 3.0$). The higher redshift points, however, significantly exceed the values employed in the SKADS simulation. The wider FIRST sample is an exception, falling significantly below the trend displayed by the narrower samples, but closely matching the Mo & White bias prescription from SKADS. Taken at face value, this apparent excess clustering suggests a greater proportion of more highly biased AGN observed than was assumed in the simulations, or perhaps a considerable underestimate of the typical halo masses of any or all of the galaxy types considered.

Our unmatched FIRST sources, especially, in the GAMA field are more strongly clustered than would be expected by simply subtracting the cross-matched $N(z)$ from the SKADS $N(z)$ and assuming the same population as a matching SKADS sample. This means either that the assumed redshift distribution is skewed towards low redshift, or the fraction of more massive galaxies is higher than

expected. Inspection of Figure 4.2 shows that our assumed $N(z)$ for this point contains a significant fraction of $z < 0.5$ sources where one would expect to observe a greater proportion of less massive, star-forming galaxies rather than the more massive AGN. Given that if these low-redshift objects were detected in FIRST, we would expect them to have optical counterparts in SDSS/GAMA (cf. almost complete cross-identification at $z \simeq 0.6$ in Figure 4.2), it is perhaps unrealistic to assume these missing low-z sources remain in our unmatched sample. The cause of this large fraction of unmatched sources is discussed in Section 4.2.3, however, to determine the effect that these sources have on our measured clustering length and bias, we implement cuts on the SKADS redshift distribution, where we remove all sources at $z < 0.05$ and $z < 0.1$ from the simulations. This is a coarse way of simulating the combined effect of the resolution bias and the removal of very low-redshift sources with photometric redshifts with $z < 0.002$. We then recalculate the correlation length and bias using these new redshift distributions and find that the correlation length and bias increases significantly (~ 25–50 per cent increase in bias). Although subject to many uncertainties, this may point to a higher proportion of highly biased objects in the radio source population at high redshift.

One of the main uncertainties in the SKADS simulation is indeed the evolution and bias of the FRI-type objects, which dominate the source counts at the flux-density limit of the FIRST survey at $z > 0.5$. In the SKADS simulation the FRIs are less biased than the FRIIs, however it is becoming clear that there is a large overlap in the stellar mass distributions between the generally less radio luminous FRIs and their FRII counterparts at high redshifts (e.g. McLure et al., 2004; Herbert et al., 2011), thus it is possible that the bias prescription for the FRI sources in the SKADS simulation is underestimated. If we assign a similar halo mass to the FRI as for the FRIIs, then we find that the expected average bias of radio sources in our sample ($z \simeq 1.21$) to be in the region of $b = 5.7$; this is much more closely aligned with our measured value of $b = 5.95^{+0.48}_{-0.53}$.

Deeper surveys over smaller areas will be able to address this issue better than the relatively shallow FIRST data. Indeed, deep multi-wavelength surveys are now beginning to tackle the question of redshift evolution (e.g. Smolčić et al., 2009; McAlpine & Jarvis, 2011; Simpson et al., 2012; McAlpine, Jarvis & Bonfield, 2013), and with slightly more area could measure the clustering length. This is the goal of Chapter 5, to follow.

4.6 Conclusions

Using radio observations from the FIRST survey and optical/infrared data (SDSS/UKIDSS LAS) over the GAMA survey field area, we have obtained a cross-matched sample of radio galaxies with optical host galaxy redshifts (~ 42 per cent spectroscopic). We have measured the angular correlation function of radio sources over the extended GAMA survey area. Assuming parent redshift distributions from SKADS (Wilman et al., 2008), we have inferred the spatial correlation length r_0 and the mass bias $b(z)$ for the matched radio sources and the radio sources without optical identifications, extending our redshift range up to $z \simeq 1.55$. The results can be summarized as follows:

1. The angular correlation function follows a power law, but we find a steeper slope of $1 - \gamma \simeq -1.2$ compared with that found for other classes of galaxy and often assumed in the literature.

2. Our cross-matched sample yields a spatial correlation length of $r_0 \simeq 8.5\ h^{-1}$Mpc to $r_0 \simeq 10.7$ h^{-1}Mpc at $0.3 < z < 0.65$ (assuming stable clustering evolution). Adding the assumption of the SKADS parent $N(z)$, we measure the clustering length up to median redshift $z \simeq 1.55$, where we find $r_0 \simeq 12.2\ h^{-1}$Mpc.

3. We measure the bias as a function of redshift across the subsamples finding it to increase from $b(z = 0.30) \simeq 2.8$ to $b(z = 1.55) \simeq 9.2$. The trend in the bias found is consistent with that inferred from the projected correlation function of spectroscopic cross-matches in Chapter 3.

4. These values were compared with predicted values from a model assuming population fractions and masses from SKADS, and were in qualitative agreement with, but exceeded, the prescribed values at high redshift. This is most probably due to a combination of surface density fluctuation in the FIRST survey, leading to a shortfall of sources at ~ 1 mJy which in turn biases our sample towards a higher fraction of strongly clustered AGN at high redshifts and/or potential inaccuracies in the halo masses of particular radio subpopulations used in SKADS.

5. If we assign a similar halo mass to the FRI sources as assumed for the FRII sources in our radio survey then we find that we can reproduce our bias value at high redshift. A more highly biased FRI population may in turn lead to a highly biased tracer of the high-redshift Universe for cosmological applications with radio surveys (e.g. Raccanelli et al., 2012; Camera et al., 2012).

While we place modest constraints on the clustering evolution of mJy radio sources and suffer from limited cross-identification with the optical surveys, our use of a well-constrained redshift distribution

from SKADS allows us to extend our redshift range to $z > 1$. This highlights some potentially important discrepancies with a simple model bias, implying stronger clustering at higher redshifts than is expected from the fixed halo masses assumed in the SKADS simulations through some combination of a greater proportion of massive AGN than previously thought and/or an increased typical halo mass being observed. This idea will be investigated further, in Chapter 5, using smaller, deeper fields to reach higher redshifts more directly, without the need for a model $N(z)$.

<div align="right">

5

</div>

Evolution in the Bias of Faint Radio Sources to

$z \sim 2.2$ in VIDEO †

5.1 Introduction

Chapters 3 and 4 use a combination of spectroscopic and photometric redshifts to investigate the clustering of FIRST radio sources to $z \simeq 1.5$, and provide a comparison between spatial and angular correlation function results (with and without the precision of spectroscopic redshifts, respectively). However, redshift measurements are lacking at high redshift, where the clustering is stronger but poorly constrained over such large areas. While sky coverage alone, with surveys such as NVSS, provides the statistical power required to measure the strength of the clustering of radio sources over large scales to depths of a few mJy (see Section 1.3.5), the depths of similarly large-area optical surveys, spectroscopic or otherwise, do not allow for optical identification of the radio sources with any significant completeness. Additionally, small-area surveys have been carried out at the μJy level, which also have potential uses for cosmology and large-scale structure measurements. At this level, the radio population becomes less dominated by FRI and FRII type active galactic nuclei (AGN), and we observe a greater fraction of star-forming galaxies. However, the lower flux-density limit also extends the range at which

we can detect AGN, reaching beyond $z \sim 1$ where the bias of radio sources is poorly understood. It is important to measure the bias of the radio sources to these high redshifts, and to know how it evolves, in order to inform cosmological experiments dependent on disentangling the observed galaxy clustering from other effects, such as cosmic magnification (e.g. Scranton et al. 2005; Wang et al. 2011) and the integrated Sachs-Wolfe effect (ISW; McEwen et al. 2007; Giannantonio et al. 2008; Raccanelli et al. 2008). In particular the large volume of the Universe that will be opened up by the SKA and its precursors may provide important information on the very largest scales (e.g. Raccanelli et al., 2012; Camera et al., 2012).

The aim of this chapter is to investigate the bias of a sample of faint radio sources, extending to $z > 2$ where observational measurements have been lacking. We use 1.4 GHz radio data from the VLA–VIRMOS Deep Field (Bondi et al., 2003), and the VLA–COSMOS Deep survey (Schinnerer et al., 2010), each covering an area of 1 square degree to $S \gtrsim 90\,\mu$Jy. The former overlaps with optical photometry from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) Deep-1 field (D1) and near-infrared photometry from the VISTA Deep Extragalactic Observations (VIDEO; Jarvis et al. 2013) survey with a depth of $K_\mathrm{s} < 23.5$, with similar overlap between COSMOS and the UltraVISTA survey. We overcome the signal-to-noise issue of having a far smaller sample than a wider, shallower NVSS-like survey by inferring the properties of the radio sources by the combined use of the angular correlation function of K_s-selected VIDEO sources (with sufficiently large numbers to keep uncertainties small) and the angular cross-correlation of these VIDEO sources with the radio sample (see e.g. Guo et al., 2011; Hartley et al., 2013, for similar use of this technique). With reliable photometric redshifts out to $z \sim 4$ for all of the galaxies used, we have a good knowledge of the redshift distributions of our samples as well as estimates of their radio luminosity. Even when coarsely binning by redshift, this gives us valuable constraints on the bias of these radio sources in bins up to a median redshift of $z = 2.15$, and an insight into the clustering specifically of typical radio AGN at high redshift.

This chapter is organized as follows: Section 5.2 describes the multi-wavelength surveys from which we construct our galaxy samples. Section 5.3 details the correlation function methods used to calculate the galaxy bias, and Sections 5.4 and 5.5, respectively, show our results and present our discussion of them. The results are summarized in Section 5.6.

5.2 Data

For the twin requirements of this chapter (deep radio data and corresponding deep near-infrared photometry, with redshifts), we find two similar deep fields, observed independently. Both fields are observed with the Visible and Infrared Survey Telescope for Astronomy (VISTA) facility in Chile. This is a 4.1m wide-field survey telescope with a 1.65-degree field of view and a 67-megapixel near-infrared camera. Both are 1–2 square degrees in size with similar depths in the K_s band, and radio coverage with with Very Large Array (VLA) at 1.4 GHz to flux densities below 0.1 mJy.

5.2.1 VVDS/VIDEO

Near-Infrared observations

The VISTA Deep Extragalactic Observations (VIDEO; Jarvis et al. 2013) survey covers ~ 12 square degrees over three fields in five near-infrared bands, tracing the evolution of galaxies and large-scale structure from the present out to $z = 4$, and higher for AGN and the most massive galaxies. The survey photometry reaches 5σ AB-magnitude depths of 25.7, 24.5, 24.4, 24.1 and 23.8 (in $2''$ apertures) in Z, Y, J, H and K_s bands, respectively.

The VIDEO-XMM3 tile detailed in Jarvis et al. (2013) overlaps with the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS; Ilbert et al. 2006) Deep-1 field (D1) optical data (u*, g', r', i', z' bands) over a $1° \times 1°$ area. The combination of these optical and near-infrared data allow for improved photometric redshifts with 3.1 per cent catastrophic outliers at $z < 1$, and further improvements at $1 < z < 4$ expected due to VIDEO's sensitivity to the Balmer and 4000 Å breaks at these redshifts (see Jarvis et al. 2013 for further details).

Due to particularly bright sources in the VIDEO images adding to the noise in surrounding areas and obscuring faint nearby sources, we apply a mask to our parent near-infrared catalogue. A circular area of $0.01°$ in radius is excised, centred on any sources brighter than 13.5 mag.

Any remaining stellar contaminants are removed from the VIDEO catalogue by the same means used in McAlpine et al. (2012) following the method of Baldry et al. (2010). We use the $J - K$ and $g - i$ colours to define a stellar locus, shown in Figure 5.1. All sources more than 0.12 mag redward in $J - K$ of this stellar locus are considered galaxies and remain in our VIDEO sample.

Finally, we impose a magnitude limit of $K_\mathrm{s} < 23.5$ on the near-infrared catalogue, corresponding roughly to the 22.6 (Petrosian magnitude) cut-off applied in the radio cross-matching by McAlpine et al. (2012).

Radio observations

The radio data used in this analysis come from Very Large Array (VLA) observations at 1.4 GHz by Bondi et al. (2003) in the VLA–VIRMOS Deep Field, corresponding to the same 1 square degree area covered by both VIDEO and CFHTLS-D1. This survey used the VLA in B-configuration, giving approximately $6''$ resolution, producing a final mosaic image of nearly uniform noise at $\sim 17\,\mu$Jy.

McAlpine et al. (2012) used a likelihood ratio (LR) method (Sutherland & Saunders, 1992) to identify infrared counterparts to radio sources in the VLA–VIRMOS Deep Field. The LR is given by

$$\mathrm{LR} = \frac{q(m)f(r)}{n(m)},\tag{5.1}$$

where $f(r)$ is the radial probability distribution function of the offsets between radio and infrared positions, $q(m)$ is the expected distribution of the true infrared counterparts as a function of K_s-band magnitude, and $n(m)$ is the magnitude distribution of the full catalogue of K_s-band detected objects. Related to $q(m)$ is the quantity Q_0, which estimates the fraction of radio sources with infrared counterparts above the magnitude limit of the VIDEO survey as follows:

$$Q_0 = \frac{N_\mathrm{matches} - \left(\sum_m n(m) \times \pi r^2 \times N_\mathrm{radio}\right)}{N_\mathrm{radio}},\tag{5.2}$$

where N_matches is the number of possible counterparts within r_max of the radio positions of N_radio radio sources in the catalogue. The reliability of a potential counterpart j to a radio source i is given by

$$\mathrm{Rel}_i = \frac{\mathrm{LR}_i}{\sum_j \mathrm{LR}_j + (1 - Q_0)}.\tag{5.3}$$

Of the 1054 5σ detections, 915 were found to have reliable ($\mathrm{Rel}_i > 0.8$) counterparts in VIDEO. These are a combination of radio-loud AGN, quasars and star-forming galaxies (see e.g. McAlpine, Jarvis & Bonfield, 2013).

We impose a flux density limit of $S > 90\,\mu$Jy on the radio sources. While the nominal detection limit of Bondi et al. (2003) is $80\,\mu$Jy, with some 41 sources detected at lower flux densities down to ~ 60 μJy, there is still appreciable incompleteness at this limit which merits the slightly more conservative cut used in our analysis. This results in a final sample of 766 objects in our radio sample, and 95,826 in the corresponding infrared sample.
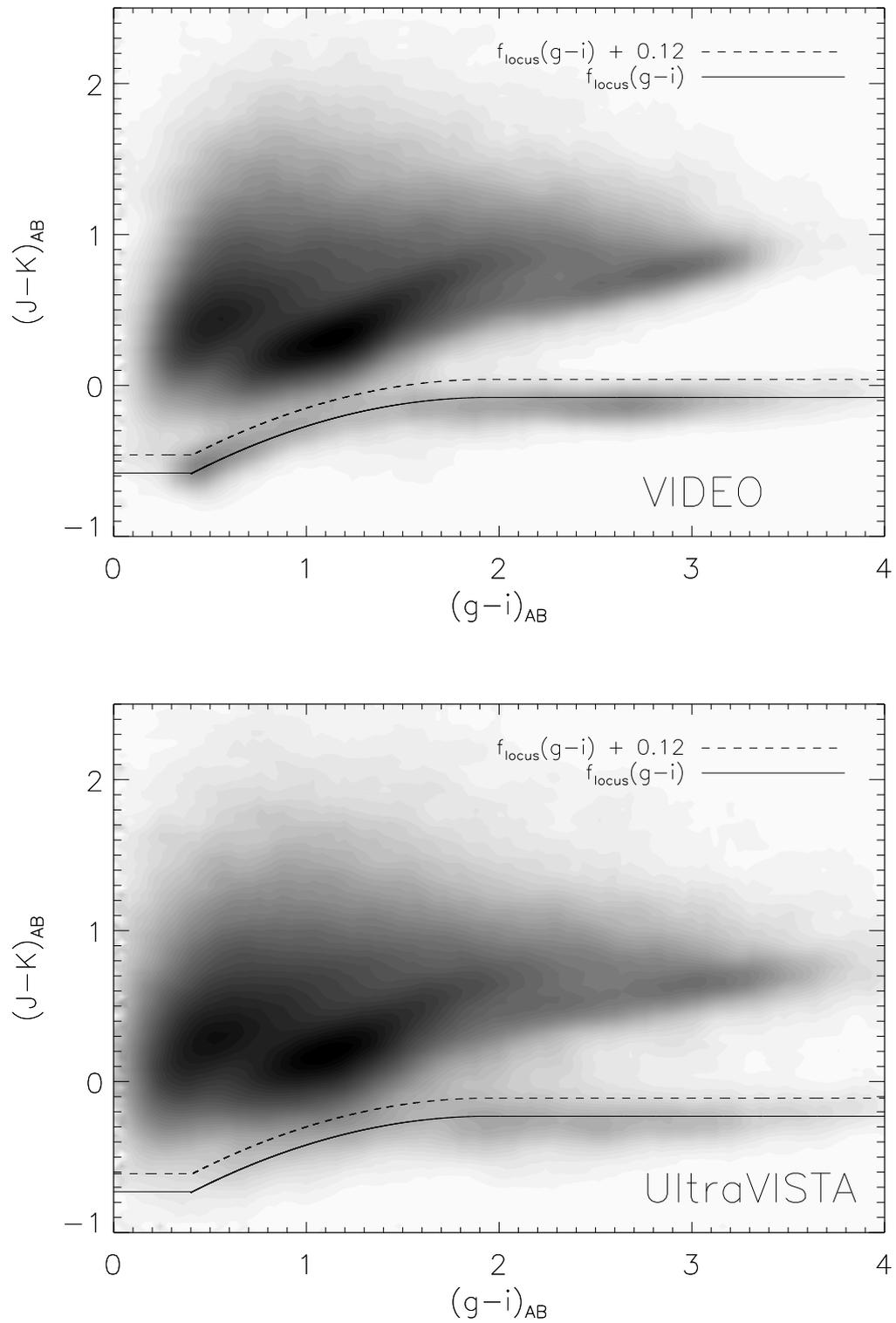
FIGURE 5.1: *J–K* vs. *g–i* colours of the VIDEO sources ($K_s < 23.5$; top panel) and UltraVISTA sources ($K_s < 23.65$; bottom panel), with the stellar locus (solid line) and imposed galaxy cut-off (dashed line) as used by McAlpine et al. (2012). The stellar locus fit for the UltraVISTA data is shifted 0.15 mag lower to account for a discrepancy in the K_s photometry.

5.2.2   COSMOS/UltraVISTA

Near-Infrared observations

UltraVISTA (McCracken et al., 2012) is another VISTA survey, observing the COSMOS field in $YJHK_s$ and narrow-band filters, reaching 5σ AB-magnitude depths of ~ 25 in Y and ~ 24 in JHK_s (in $2''$ apertures). Muzzin et al. (2013) produced a public K_s-selected catalogue covering all 1.62 square degrees of the COSMOS/UltraVISTA field with PSF-matched photometry in 30 bands between 0.15 and 24 μm from *GALEX*, Subaru, CFHT, VISTA and *Spitzer*. The included photometric redshifts have a catastrophic outlier fraction of only 1.6 per cent.

Radio observations

Radio observations in this field formed the VLA–COSMOS Deep project (Schinnerer et al., 2010) which used the VLA in A-configuration to achieve ~ 12 μJy noise in the central $50' \times 50'$ at $2.5''$ resolution (in addition to the Large project at $1.5''$ resolution). We use the central $1° \times 1°$ of this field to match the size and geometry of the VIDEO field, and to keep the noise level low and relatively uniform across our sampled area.

We repeat the star-galaxy separation and LR matching procedure for these data, as in McAlpine et al. (2012), with two small differences. Muzzin et al. (2013) show an offset in the UltraVISTA K_s photometry when compared with other surveys, UltraVISTA being systematically fainter. While this offset is said to be < 0.1 mag, when performing the star-galaxy separation we find a shift of 0.15 mag of the stellar locus in $J - K$ best matches the photometry (see Figure 5.1). Secondly, the expected positional error is adjusted for the claimed calibration error of $0.25''$ in Schinnerer et al. (2010), rather than $0.1''$ in the Bondi et al. (2003) radio data. Of the 1,733 radio sources in our sample area, we find 1,306 with reliable UltraVISTA counterparts, before imposing radio flux density and K_s magnitude limits.

5.2.3   Final Samples

The final galaxy samples in both VIDEO and COSMOS fields are geometrically similar $1° \times 1°$ areas and we define identical infrared and radio limits to both fields in order for them to remain directly comparable as much as possible. We impose a $K_s < 23.5$ magnitude cut to VIDEO and UltraVISTA (after 0.15 mag correction) sources and radio flux density limit $S_{1.4} > 90$ μJy for the VLA sources.

Field	RA (°)	Dec (°)	IR sources	$z_{\text{IR,med}}$	Radio sources	$z_{\text{radio,med}}$
VIDEO	36.5	-4.5	95,826	1.09	766 (575)	1.05
COSMOS/UltraVISTA	150.0	2.25	122,495	0.92	794 (664)	1.01

TABLE 5.1: Central RA and declination, galaxy counts ($z < 4$) and median redshifts for the two 1 square degree fields under investigation. Radio source numbers in brackets are high-power ($L_{1.4} > 10^{23}$ W Hz^{-1}), assuming a spectral index $\alpha = 0.8$.

The samples are summarized in Table 5.1.

5.3 Methods

5.3.1 Angular Correlation Function

The angular autocorrelation function (see Section 1.2.1) is calculated for the densely populated near-infrared samples in the same way as described in Chapter 4. This will provide a high signal-to-noise ratio with which we can infer the clustering properties of the background near-infrared galaxy population. In Figure 5.3 we show $w(\theta)$ of our K_{s}-selected VIDEO and UltraVISTA galaxies in order to understand their inherent large-scale clustering properties as a stepping stone towards investigating the relative clustering between these galaxies and the subset of radio sources in the same VIDEO field.

5.3.2 Cross-Correlation Function

Closely related to the two-point correlation function is the two-point cross-correlation function, which compares two different data sets D_1 and D_2. The same procedure is followed as for the auto-correlation of one data set, but we modify equation 4.1 as follows:

$$w_{\text{cross}}(\theta) = \frac{D_1 D_2 - D_1 R - D_2 R}{RR} + 1, \tag{5.4}$$

where $D_1 D_2$ and $D_i R$ are analogous to DD and DR above, and each quantity is normalized.

The cross-correlation function is also fitted with a power law in the same manner as the auto-correlation function and describes the relative cross-clustering of two populations with one another. Here, we use the cross-correlation function of VIDEO K_{s}-selected galaxies with radio sources in order to provide a higher signal-to-noise measurement of the clustering than would be possible with the radio auto-correlation alone. Figure 5.4 shows the angular cross-correlation function for the radio sources
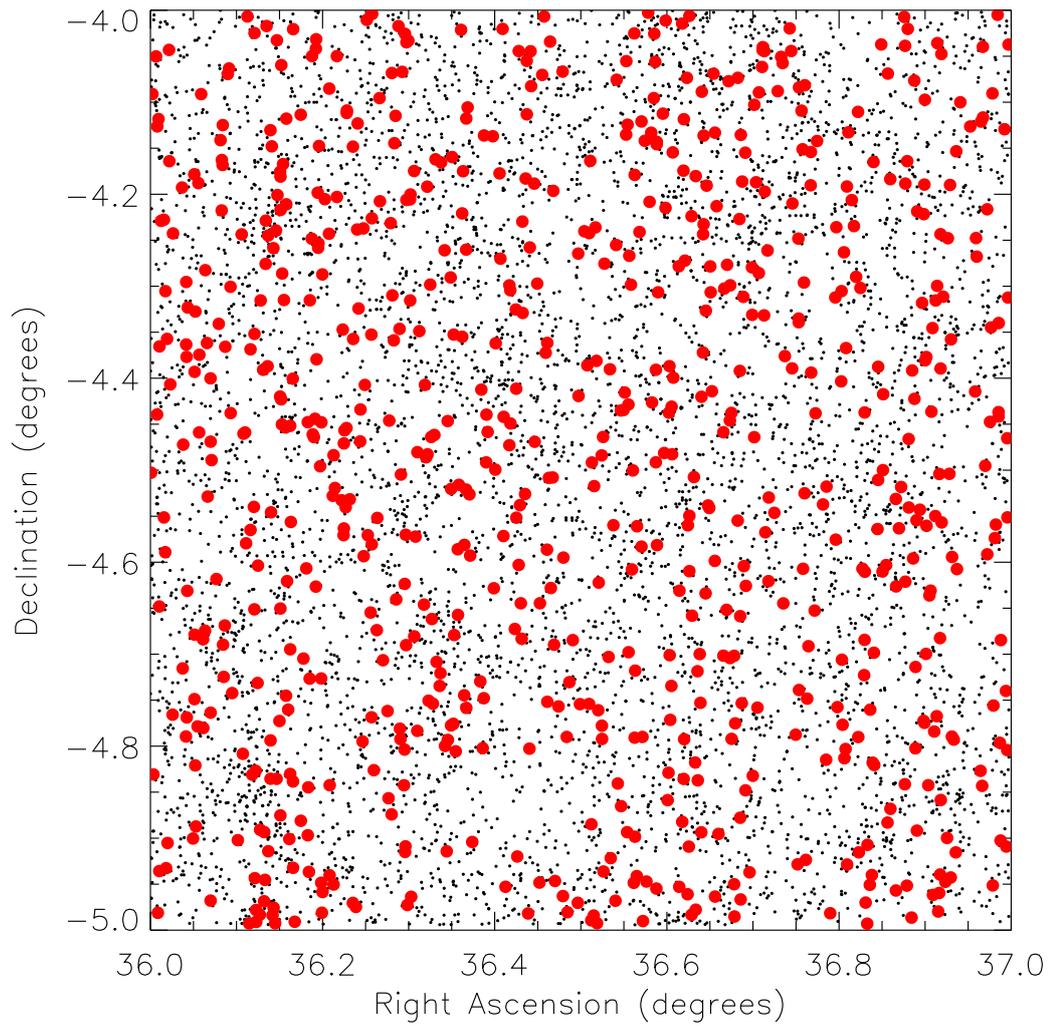
FIGURE 5.2: VIDEO IR sources shown in black (limited to $K_s < 20$ for clarity) with their radio counterparts ($K_s < 23.5$; $S_{1.4} > 90$ μJy) shown in red. Regions of radius $0.01°$ around the brightest VIDEO sources have been masked out of both catalogues.
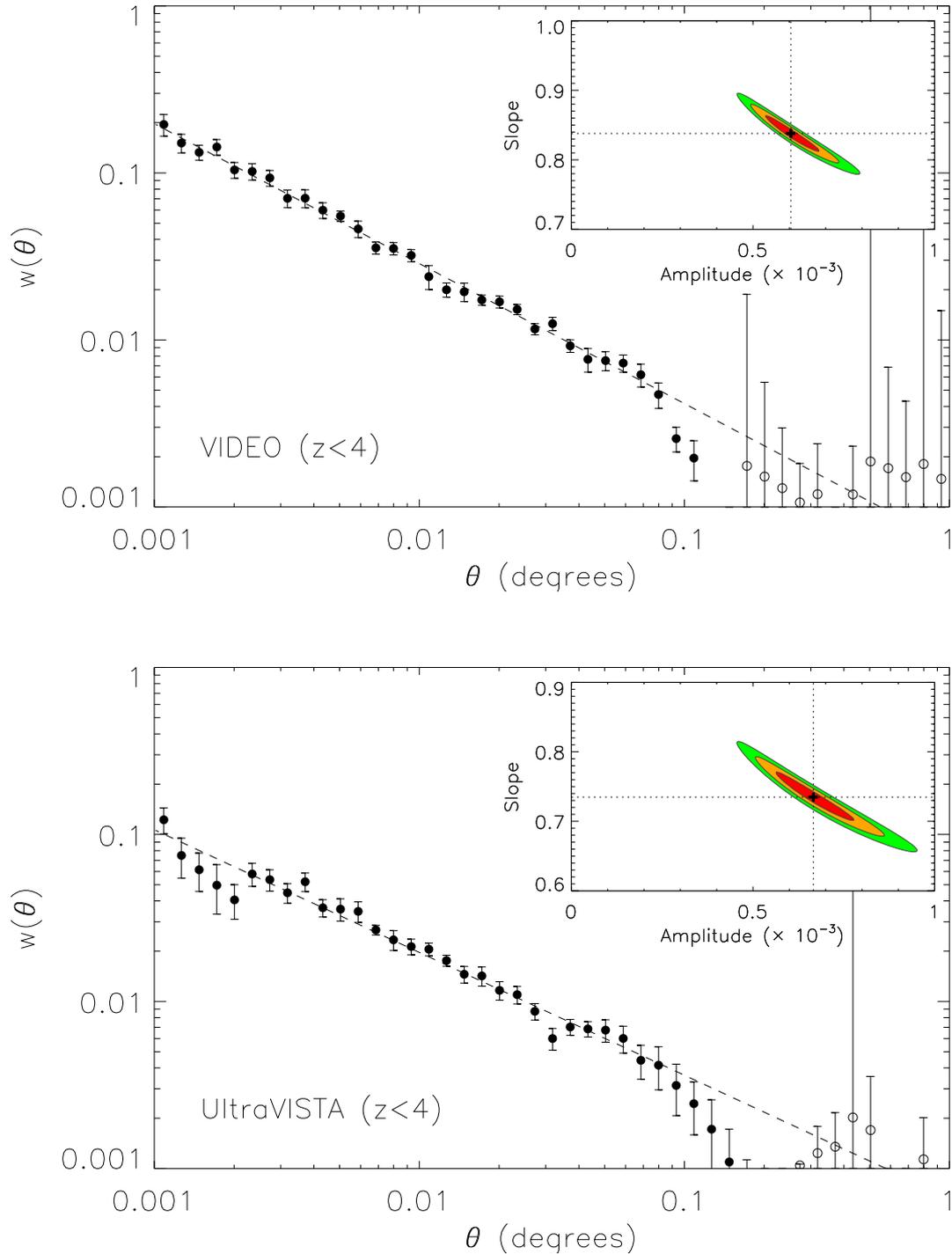
FIGURE 5.3: The angular correlation functions for $K_s < 23.5$ VIDEO/UltraVISTA infrared sources (top/bottom panel, respectively), with bootstrap resampling errors. The dashed line shows the best fit power-law (for $\theta < 0.08°$) and the inset plot shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83. Points plotted in grey are the absolute values where $w(\theta) < 0$.
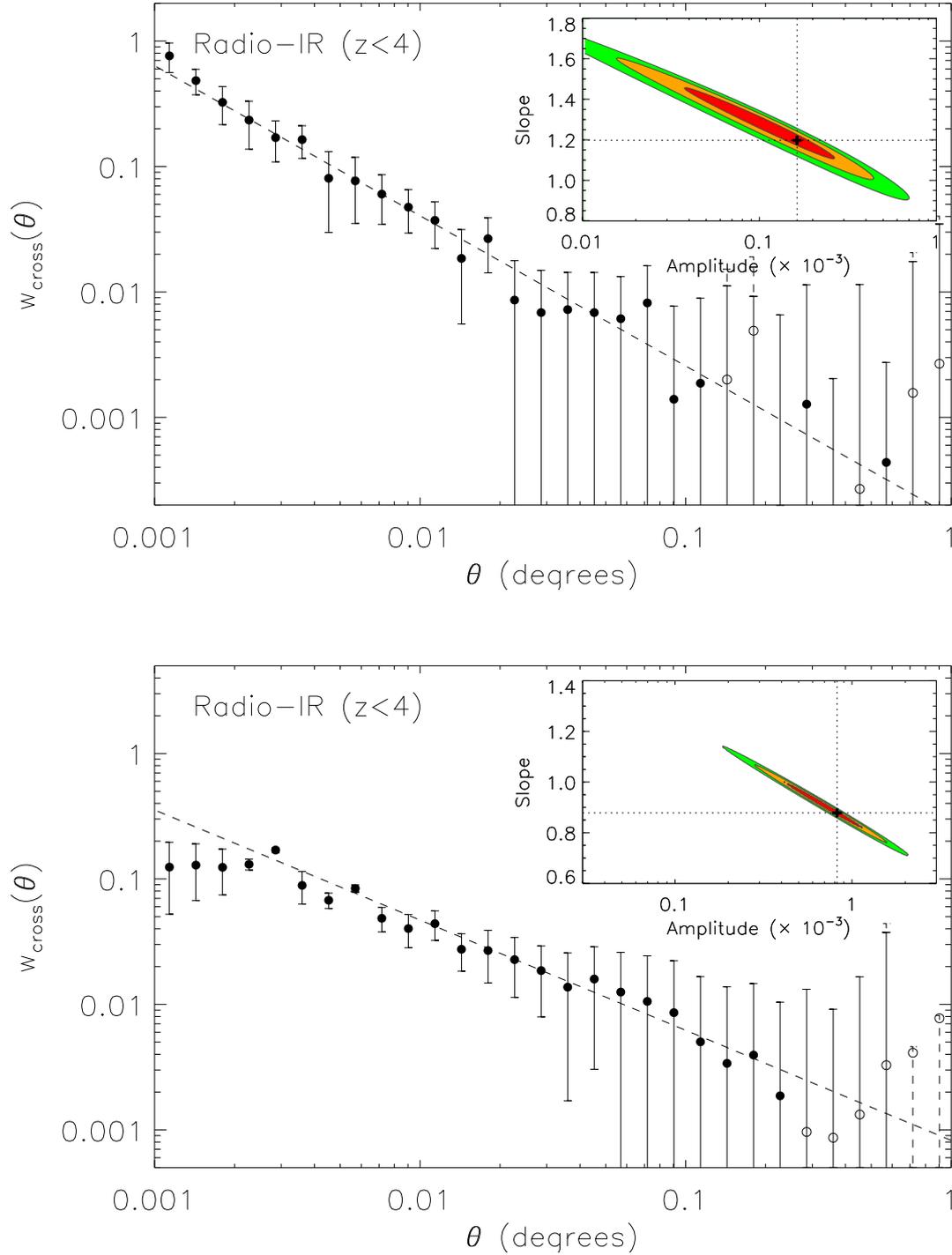
FIGURE 5.4: The angular cross-correlation function of the radio counterparts with the VIDEO/UltraVISTA infrared sources (top/bottom panel, respectively), with bootstrap resampling errors. The dashed line shows the best fit power-law and the inset plot shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83. Points plotted in grey are the absolute values where $w_{\mathrm{cross}}(\theta) < 0$.

with the near-infrared selected sources, which, in the following sections, we use to infer the clustering properties of the radio galaxies themselves.

5.3.3   Spatial Clustering and Limber Inversion

If the redshift distribution of a set of objects is known, one may deproject the angular correlation function into the spatial correlation function (see Section 4.4). Determining $\xi(r)$ directly is difficult, as a complete set of individual redshifts is rarely available for a given survey, thus requiring the redshift distribution to be estimated in order to deproject $w(\theta)$. Using the photometric redshifts available for the galaxies in the VIDEO survey, however, means we may apply a redshift distribution directly from the data. The photometric redshift distributions for both the K_s-selected galaxies and the radio galaxies are calculated by adding the normalized probability distributions of the individual photometric redshifts of each object. This ensures that we fully incorporate the uncertainties associated with the photometric redshifts when determining the clustering of the various populations. These redshifts distributions are shown in Figure 5.5.

An epoch-dependent form of the spatial correlation function is assumed (see e.g. de Zotti et al., 1990; Overzier et al., 2003, and references therein):

$$\xi(r, z) = \left(\frac{r_0}{r}\right)^{-\gamma} \times (1 + z)^{\gamma - (3 + \epsilon)}, \tag{5.5}$$

where ϵ parameterizes the clustering model being assumed. Following Chapter 4, we assume the co-moving model with $\epsilon = \gamma - 3$. The spatial correlation function slope, γ, is the same as that used in the power law fit to the angular correlation function (where the magnitude of the slope is $\gamma - 1$), so we measure this parameter through the $w(\theta)$ function. The correlation length r_0 is calculated in the same manner as described in Chapter 4 (equation 4.4).

5.3.4   Mass Bias

The bias of the near-infrared sources is calculated with the same method used in previous chapters. The analogous equation to equation 1.29 for the cross-correlation is

$$b_{K\mathrm{R}}^2(z) = \frac{\xi_{K\mathrm{R}}(r, z)}{\xi_{\mathrm{DM}}(r, z)} = b_K(z) b_\mathrm{R}(z), \tag{5.6}$$

where K and R subscripts denote K_s-selected galaxies, and radio galaxies respectively. Given that we are able to calculate the bias of the near-infrared sources, $b_K(z)$, from their auto-correlation function,
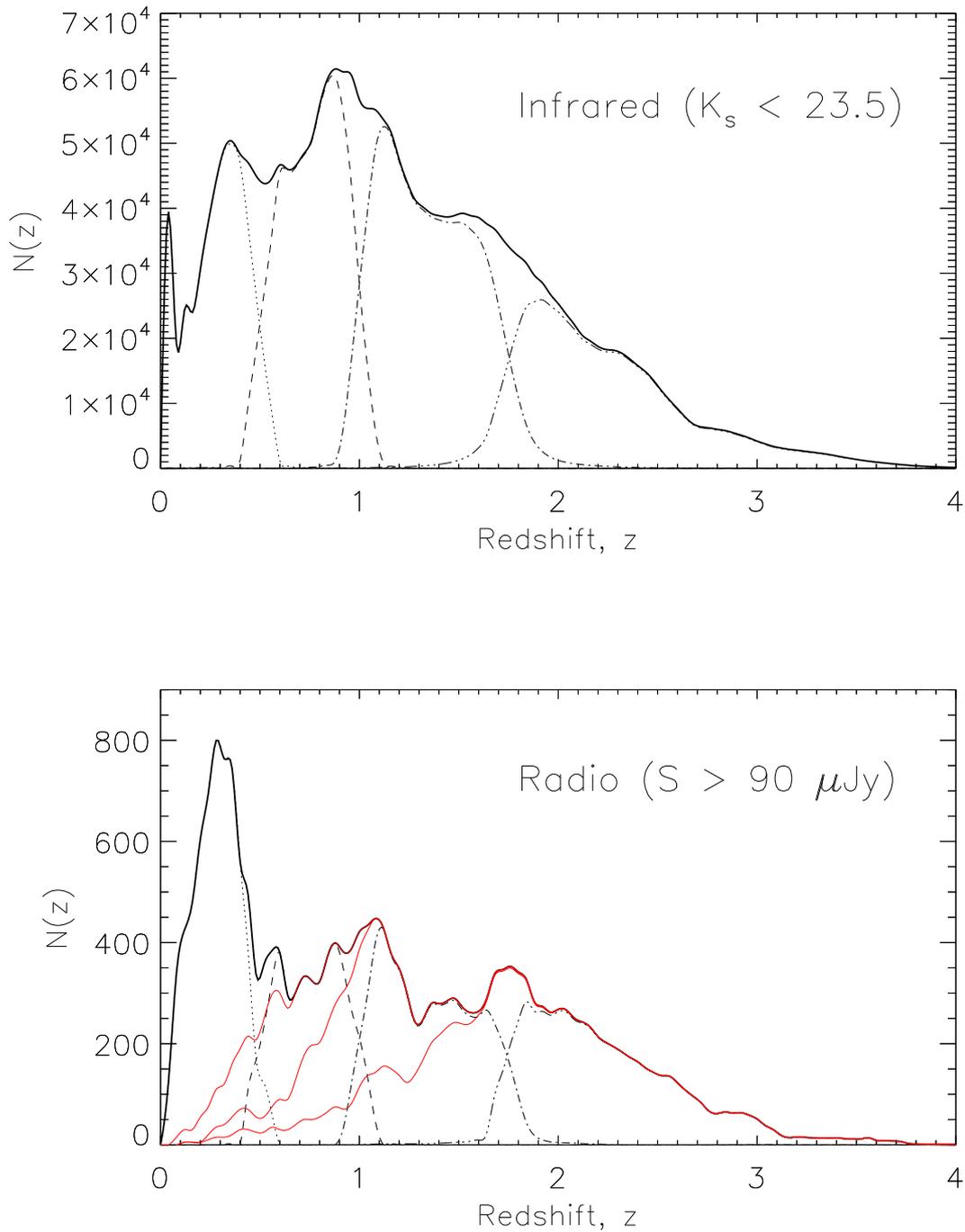
FIGURE 5.5: Redshift distributions of infrared-selected sources (top panel) and radio-selected sources (bottom panel). The four broken lines show the disributions of sources in our four redshift bins. The three red lines on the right panel show the distributions of radio sources with $L > 10^{23}$, $10^{23.5}$ and 10^{24} W Hz^{-1} in increasing order of median redshift.

Redshift range	IR–IR			Radio–IR		
	$z_{\mathrm{med,IR}}$	$A(\times 10^{-3})$	γ	$z_{\mathrm{med,radio}}$	$A(\times 10^{-3})$	γ
$z < 4$	0.92	$0.68^{+0.08}_{-0.07}$	$1.73^{+0.02}_{-0.02}$	1.01	$1.09^{+0.33}_{-0.28}$	$1.82^{+0.05}_{-0.05}$
$z < 0.5$	0.30	$2.37^{+0.31}_{-0.28}$	$1.67^{+0.03}_{-0.03}$	0.31	$0.85^{+0.29}_{-0.23}$	$2.12^{+0.06}_{-0.06}$
$0.5 < z < 1$	0.78	$2.70^{+0.23}_{-0.21}$	$1.76^{+0.02}_{-0.02}$	0.77	$0.81^{+0.18}_{-0.15}$	$2.14^{+0.04}_{-0.04}$
$1 < z < 2$	1.34	$0.42^{+0.06}_{-0.05}$	$2.02^{+0.03}_{-0.03}$	1.40	$0.18^{+0.11}_{-0.10}$	$2.22^{+0.14}_{-0.09}$
$2 < z < 4$	2.44	$1.40^{+0.36}_{-0.27}$	$1.92^{+0.05}_{-0.05}$	2.37	$0.30^{+0.17}_{-0.11}$	$2.27^{+0.11}_{-1.11}$

TABLE 5.2: Fitted parameters from the auto-correlation function (IR–IR) $w(\theta)$, of K_{s}-selected UltraVISTA sources and cross-correlation function (Radio–IR) $w_{\mathrm{cross}}(\theta)$ of these IR sources with radio sources from COSMOS.

the radio bias may be given by

$$b_{\mathrm{R}}(z) = \frac{b^2_{K\mathrm{R}}(z)}{b_K(z)}. \tag{5.7}$$

To account for the fact that we are measuring these quantities for discrete samples with different redshift distributions where z for corresponding bins is subtly different between the NIR and radio sources, we multiply the right hand side of equation 5.7 by $D(z_K)/D(z_{\mathrm{R}})$, where $D(z)$ is the cosmological growth function, which is a relatively small correction. This gives a final quantity, $b_{\mathrm{R}}(z_{\mathrm{R}})$, which describes the bias of the radio sources at the redshift of those radio sources:

$$b_{\mathrm{R}}(z_{\mathrm{R}}) = \frac{b^2_{K\mathrm{R}}(z_{\mathrm{R}})}{b_K(z_K)} \frac{D(z_K)}{D(z_{\mathrm{R}})}. \tag{5.8}$$

5.4 Results

5.4.1 Problems with COSMOS/UltraVISTA

There are several issues with the data over the COSMOS/UltraVISTA field that bring any subsequent analysis into doubt. While the near-infrared catalogue includes measurements in 30 photometric bands, making the corresponding redshift estimates quite reliable, there are some issues with K_{s} band and radio data with which we select our samples. Muzzin et al. (2013) demonstrate that the UltraVISTA K_{s} band zero point is too large by about 0.08 mag, after comparing with Two Micron All-Sky Survey (2MASS; Skrutskie et al., 2006), the NOAO Extremely Wide-Field Infrared Imager (NEWFIRM) Medium-Band Survey (NMBS; Whitaker et al., 2011) and CFHT/WIRCAM imaging from McCracken et al. (2010).

As the UltraVISTA K_s band is used as the anchor filter when calculating the zero-point offsets of the remaining filters, this should not have any significant effect on the source colours, however. Note that this is not the case for the VIDEO data, which are calibrated using magnitudes and colours from 2MASS. Nevertheless, the colours shown in Figure 5.1 appear to be 0.1–0.2 mag bluer in $J - K_s$ than the VIDEO sources, with a less well defined stellar locus, and an apparent dearth of sources at the blue end of the locus. As also mentioned in Section 5.2, the positional error of the radio source positions from Schinnerer et al. (2010) are 2.5 times as large as the corresponding Bondi et al. (2003) data used for the VIDEO field, despite having superior angular resolution. This has a negative impact on the cross-matching of these sources with potential near-infrared counterparts found nearby (in projection).

The combination of these issues appears to have an effect on both the auto- and, in particular, the cross-correlation functions. Figure 5.3 shows $w(\theta)$ for the two fields considered, with UltraVISTA yielding larger error bars despite having a numerical advantage in sample size over the VIDEO sample, and has a significantly shallower slope (although still consistent with -0.8). In contrast to the VIDEO correlation function, UltraVISTA also demonstrates notable deviations from a smooth power-law shape, at $\theta < 0.002°$ (7.2″) and $\theta \sim 0.03°$ (1.8′). In the cross-correlation functions shown in Figure 5.4, the UltraVISTA and VIDEO results differ significantly, the former giving a much flatter power law and a pronounced break at $\theta \lesssim 0.003°$ (11″). Given that the less densely populated VIDEO field exhibits the increased overdensity expected towards smaller separations, this feature of the UltraVISTA cross-correlation function could hint at a problem with the cross-matching procedure (affecting scales smaller than the errors in radio source position). Combined with concerns about the UltraVISTA photometry, this leads us to proceed only with the analysis based on the VIDEO field hereafter. The fitted correlation function parameters are shown in Table 5.2, but take no further part in our discussion.

5.4.2 The evolution of the bias of near-infrared selected galaxies

In order to calculate the bias of the radio sources, we first calculate the angular autocorrelation function of K_s-selected VIDEO galaxies, finding the data to be fit very well with a single power law (Figure 5.3). This fit describes the data very well over almost two decades in angular scale, from $\theta = 0.001°$ (3.6″) to $0.08°$ (4.8′) . The correlation length and bias of these sources is then calculated using their redshift distribution (Figure 5.5). We measure the evolution of the clustering of these sources in four bins of median redshift ranging from $z = 0.3$ to $z = 2.15$. These quantities are listed for the binned and full $z < 4$ samples in Table 5.3 and shown in Figure 5.6.

The power law slope of $w(\theta)$ is $\gamma - 1 = 0.85$ for all but the highest redshift bin at $1.75 < z < 4$

Redshift range	IR sources	z_{med}	$A \, (\times 10^{-3})$	γ	$r_0 \, (h^{-1}\text{Mpc})$	$b_K(z)$
$z < 4$	95,826	1.09	$0.61^{+0.05}_{-0.05}$	$1.84^{+0.02}_{-0.02}$	$2.79^{+0.08}_{-0.09}$	$1.12^{+0.03}_{-0.03}$
$z < 0.5$	17,603	0.33	$2.55^{+0.32}_{-0.30}$	$1.85^{+0.03}_{-0.03}$	$2.08^{+0.07}_{-0.07}$	$0.59^{+0.02}_{-0.01}$
$0.5 < z < 1$	25,461	0.79	$2.06^{+0.18}_{-0.17}$	$1.85^{+0.02}_{-0.02}$	$2.59^{+0.08}_{-0.09}$	$0.91^{+0.02}_{-0.03}$
$1 < z < 1.75$	31,507	1.33	$1.45^{+0.12}_{-0.11}$	$1.85^{+0.02}_{-0.02}$	$2.73^{+0.10}_{-0.10}$	$1.21^{+0.04}_{-0.04}$
$1.75 < z < 4$	21,255	2.16	$0.69^{+0.12}_{-0.10}$	$1.98^{+0.03}_{-0.03}$	$3.80^{+0.18}_{-0.19}$	$2.23^{+0.12}_{-0.12}$

TABLE 5.3: Angular clustering parameters from the autocorrelation function $w(\theta)$, and the inferred correlation length and bias for our various K_{s}-selected galaxy samples from the VIDEO survey.

FIGURE 5.6: Linear bias of VIDEO K_{s}-selected sources as a function of median redshift. Open circles correspond to the four independent redshift bins used while the filled circle is the bias for the full sample of 95,826 sources with $K_{\text{s}} < 23.5$.

Redshift range	Luminosity (WHz^{-1})	Radio sources	z_{med}	$A\ (\times 10^{-3})$	γ	$r_0\ (h^{-1}\mathrm{Mpc})$	$b_{\mathrm{KR}}(z)$	$b_{\mathrm{R}}(z)$
$z < 4$	All	766	1.02	$0.17^{+0.14}_{-0.08}$	$2.19^{+0.11}_{-0.11}$	$3.83^{+0.21}_{-0.24}$	$1.57^{+0.12}_{-0.15}$	2.13 ± 0.27
$z < 0.5$	All	234	0.29	$0.99^{+0.47}_{-0.33}$	$2.15^{+0.08}_{-0.08}$	$2.57^{+0.16}_{-0.18}$	$0.70^{+0.05}_{-0.05}$	0.57 ± 0.06
$0.5 < z < 1$	"	139	0.75	$1.28^{+0.74}_{-0.54}$	$2.07^{+0.11}_{-0.10}$	$3.94^{+0.41}_{-0.48}$	$1.38^{+0.20}_{-0.19}$	1.80 ± 0.36
$1 < z < 1.75$	"	194	1.35	$0.70^{+0.59}_{-0.41}$	$2.09^{+0.16}_{-0.13}$	$4.62^{+0.68}_{-0.75}$	$2.11^{+0.48}_{-0.39}$	4.09 ± 1.20
$1.75 < z < 4$	"	199	2.15	$0.35^{+0.29}_{-0.20}$	$2.20^{+0.16}_{-0.13}$	$5.61^{+0.83}_{-1.00}$	$3.60^{+0.99}_{-0.84}$	8.55 ± 3.11
$z < 4$	$\log(L) > 23.0$	575	1.37	$0.16^{+0.11}_{-0.07}$	$2.24^{+0.10}_{-0.10}$	$5.57^{+0.33}_{-0.36}$	$2.76^{+0.34}_{-0.30}$	7.62 ± 1.27
"	$\log(L) > 23.5$	499	1.55	$0.11^{+0.09}_{-0.06}$	$2.28^{+0.13}_{-0.12}$	$5.57^{+0.50}_{-0.52}$	$3.04^{+0.61}_{-0.46}$	9.91 ± 2.48
"	$\log(L) > 24.0$	372	1.77	$0.11^{+0.12}_{-0.06}$	$2.28^{+0.15}_{-0.14}$	$5.30^{+0.47}_{-0.61}$	$3.10^{+0.64}_{-0.54}$	11.14 ± 3.01

TABLE 5.4: Angular clustering parameters from the cross-correlation function $w_{\mathrm{cross}}(\theta)$, and the inferred correlation length and bias for radio-IR cross-clustering. The inferred radio bias $b_R(z)$ is shown in the final column.
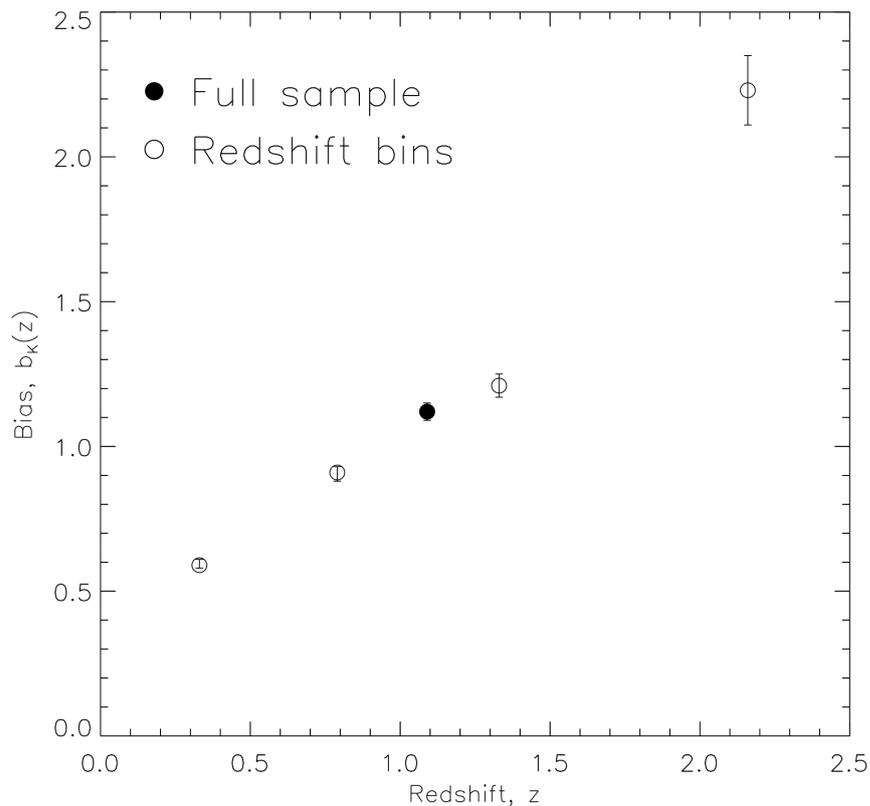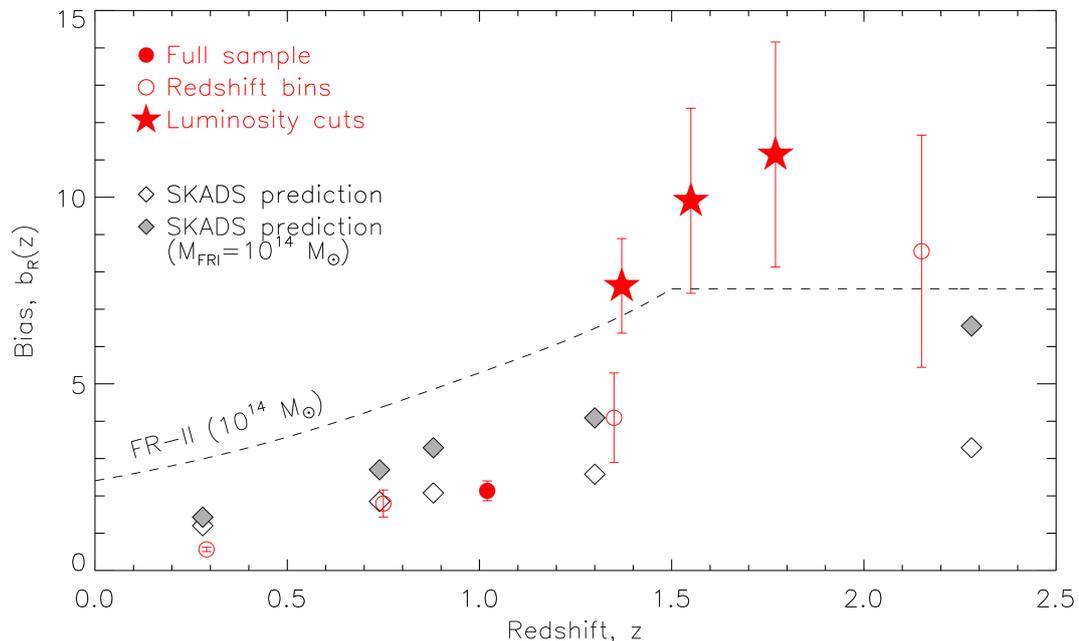
FIGURE 5.7: Linear bias of VIDEO-identified radio sources as a function of median redshift. Open circles correspond to the four independent redshift bins used while the filled circle is the bias for the full sample of 766 radio sources with $K_s < 23.5$ and $S_{1.4} > 90$ μJy. Star symbols correspond to lower luminosity limits of 10^{23}, $10^{23.5}$ and 10^{24} W Hz^{-1} from low to high redshift. The dashed line shows the FRII bias adopted by Wilman et al. (2008) in the SKADS simulations, and the diamond symbols show the expected bias based on the SKADS prescriptions (open) and with the FRI halo mass increased to 10^{14} $M_\odot$ (filled), matching the FRIIs.

where it rises to 0.98. While the angular clustering amplitude decreases with increasing redshift, Limber inversion gives a gently increasing correlation length, r_0 with increasing redshift, from 2.08 ± 0.07 h^{-1}Mpc at $z \simeq 0.33$ to $3.80^{+0.18}_{-0.19}$ h^{-1}Mpc at $z \simeq 2.16$. This corresponds to a bias increase from $0.59^{+0.02}_{-0.01}$ to 2.23 ± 0.12 and reflects the mix of galaxy populations present in a K_s-selected survey at these depths (i.e. higher redshift sources have higher mass in a flux-density limited K_s-band selected sample). The clustering of K_s-band selected galaxies as a function of galaxy mass and redshift will be studied in much greater detail in a future paper (see Section 6.3).

5.4.3 The clustering and bias of faint radio sources

In this work, we are interested in the evolution in the bias of the faint radio source populations. We measure this using the cross-correlation function of the radio and K_s-selected galaxies. Figure 5.4 shows the cross-correlation function of radio and K_s-selected galaxies for the full sample of galaxies. We also determine the cross-correlation in the four redshift bins as defined in the previous section (see Figure 5.9). Furthermore, we impose lower limits on the radio luminosity of 10^{23}, $10^{23.5}$ and 10^{24}

W Hz^{-1}, as each of these provides a slightly different sample of AGN with the fraction of normal star-forming galaxies gradually diminishing with radio luminosity (see Figures 5.5 and 5.8). Showing how this affects the measured clustering is important to gain fresh insight into the bias of AGN in an as yet uninvestigated luminosity and redshift range.

The cross-correlation length and radio-infrared relative bias ($b_{K\mathrm{R}}$), are calculated using the radio source redshift distribution and we then infer the radio bias using equation 5.8. Figure 5.7 shows the bias as a function of redshift for the radio sources. These quantities are also tabulated in Table 5.4 for the various redshift bins and for the different luminosity limits.

We find a steeper power-law slope for the cross-correlation function than for the corresponding infrared autocorrelation function, with $\gamma - 1 > 1$ in all cases. Likewise, the radio-infrared correlation length is greater than the infrared galaxy correlation, giving a relative bias, $b_{K\mathrm{R}}$, increasing even more strongly with redshift from $0.70^{+0.05}_{-0.05}$ to $3.60^{+0.99}_{-0.84}$ between the lowest and highest redshift bins at $z_{\mathrm{med}} = 0.29$ and 2.15 respectively. Accounting for the results of the infrared galaxy correlation function, this corresponds to a radio source bias of 0.57 ± 0.06 to 8.55 ± 3.11 over the same redshift range.

Imposing a minimum radio luminosity criterion to our radio source sample, we find r_0 to be greater than that for the full sample, meaning that the higher-luminosity AGN are more strongly clustered than the general radio source population. The three different luminosity cuts result in similar correlation lengths ($r_0 \simeq 5.5\,h^{-1}\mathrm{Mpc}$), with a slight decrease at the high luminosity end, but still within errors. The radio bias, however, increases as the radio luminosity increases, from $b_{\mathrm{R}} = 7.62 \pm 1.27$ for $L > 10^{23}$ W Hz^{-1} (approximately double the bias for the full radio sample at $1 < z < 1.75$ with a similar median redshift) to $b_{\mathrm{R}} = 11.14 \pm 3.01$ for $L > 10^{24}$ W Hz^{-1}. This increase is likely due purely to the increasing median redshift of higher luminosity sources, rather than a significantly more massive or more clustered sample.

5.5 Discussion

5.5.1 The clustering of near-infrared selected galaxies

As the measurement of the clustering of the radio source population is dependent on our measurement of the clustering of the near-infrared galaxies through the cross-correlation function, it is informative to compare our results for the near-infrared galaxy clustering to other results in the literature.

Our results for the K_{s}-selected VIDEO sources, based on tightly constrained angular correlation functions (see Figures 5.3 and 5.9), show an increasing clustering strength with redshift, as would be
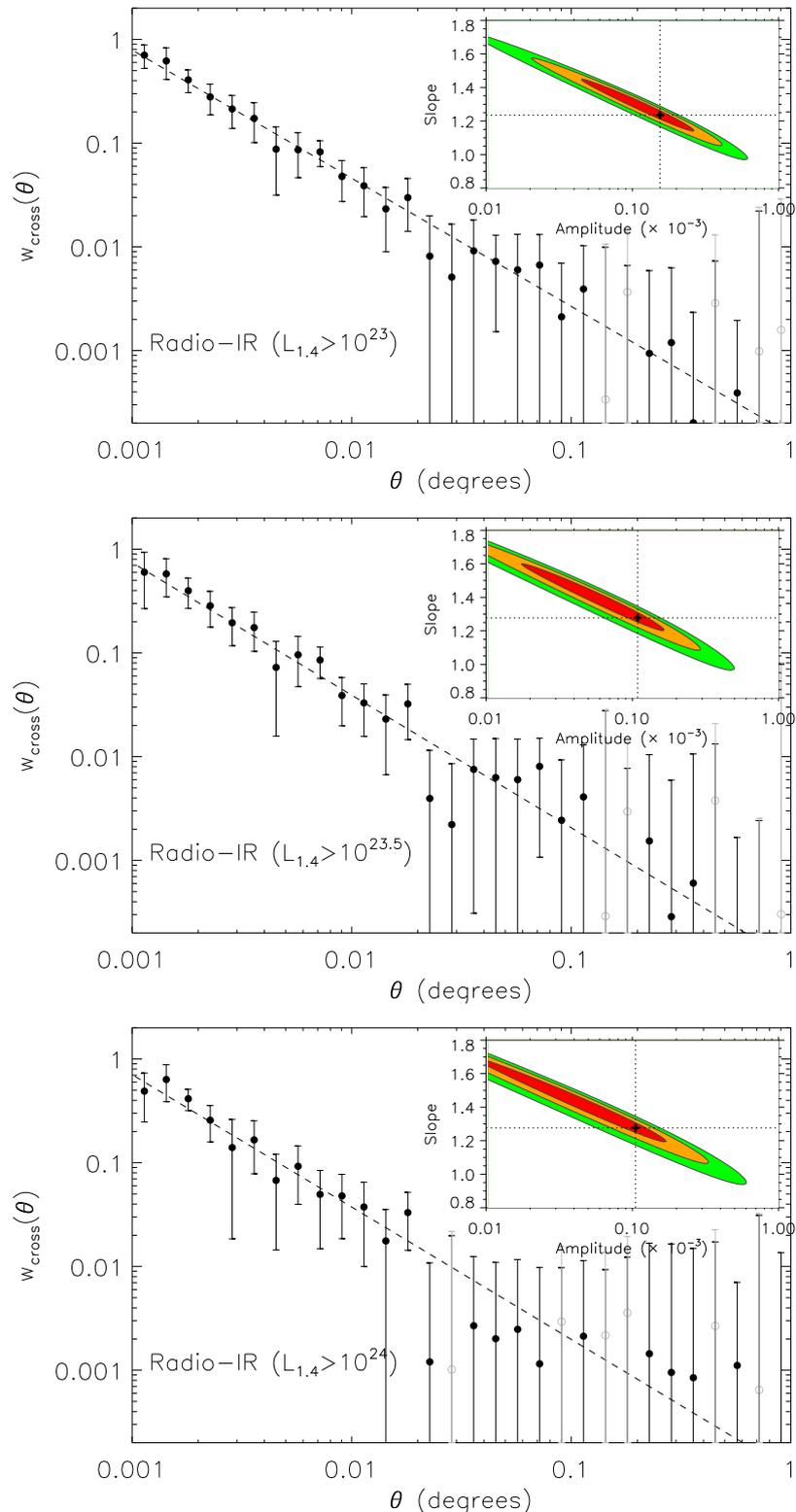
FIGURE 5.8: The angular cross-correlation function of the $L_{1.4} > 10^{23}$, $10^{23.5}$ and 10^{24} WHz^{-1} radio counterparts with the VIDEO infrared sources (with bootstrap resampling errors). The dashed line shows the best fit power-law and the inset plot shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83. Points plotted in grey are the absolute values where $w_{\text{cross}}(\theta) < 0$.
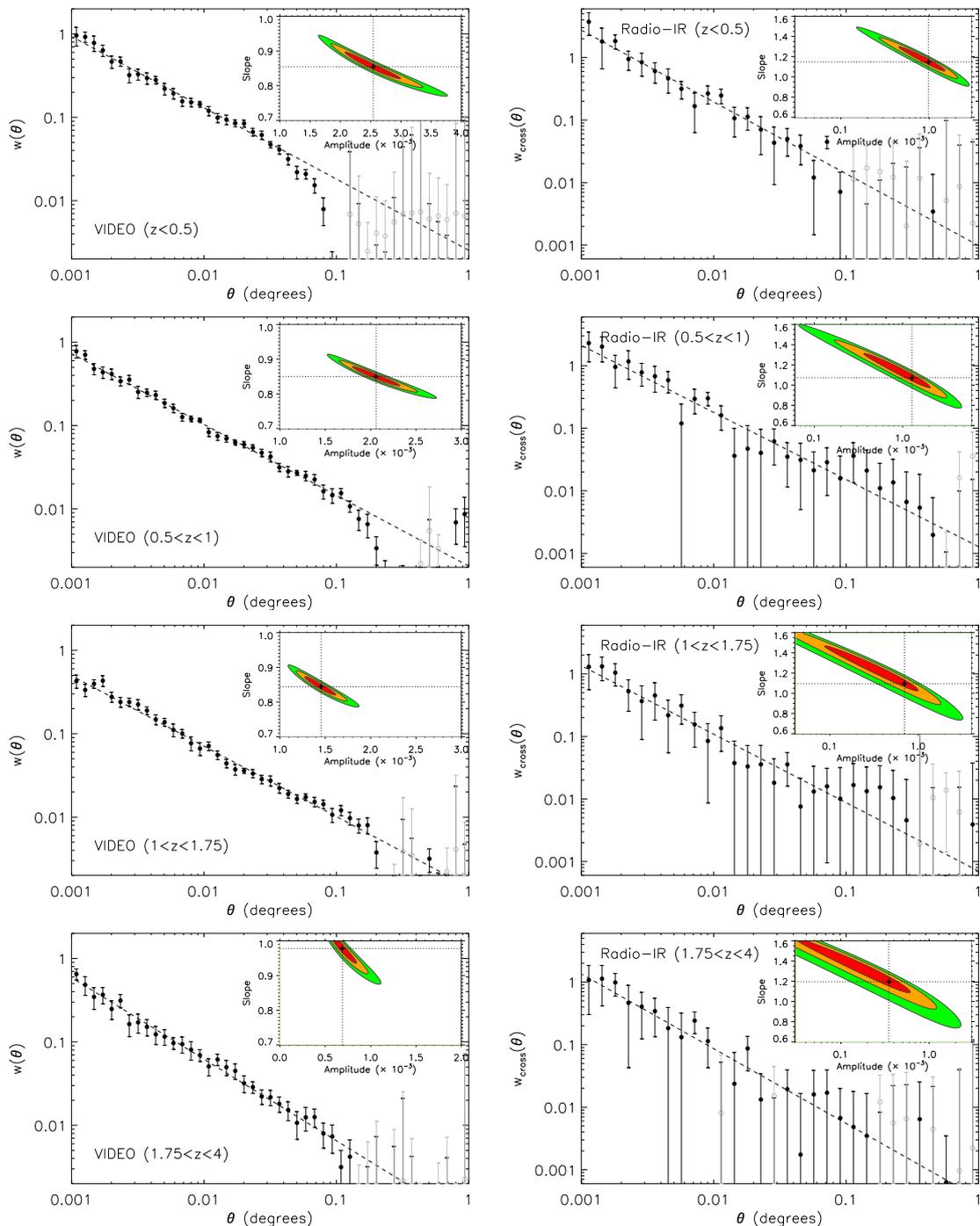
FIGURE 5.9: The angular autocorrelation function (left) of the $K_s < 23.5$ VIDEO sources, and cross-correlation function (right) with the radio counterparts (with bootstrap resampling errors) in 4 redshift bins. The dashed lines shows the best fit power-law and the inset plot shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83. Points plotted in grey are the absolute values where $w_{\text{cross}}(\theta) < 0$.

expected for a population formed near peaks in the mass distribution, with the dark matter structures collapsing over time, catching up to the galaxy clustering. However, comparison with similar studies of near-infrared galaxies suggests that we underestimate this clustering.

Furusawa et al. (2011) investigated the mass-dependent clustering of a similarly derived sample of more than 50,000 $K < 23.5$ galaxies from 0.63 square degrees of the Subaru/XMM-*Newton* Deep Survey and UKIRT Infrared Deep Sky Survey/Ultra Deep Survey (SXDS/UDS). While they do not show results for a combined sample across all redshifts, or indeed all masses, we find our redshift-binned r_0 to be lower than even their least clustered (lowest mass) galaxies at similar redshifts. Similarly, Bielby et al. (2013) find r_0 as a function of mass and redshift through the angular correlation function of a $z \lesssim 2$ galaxy sample from 2.4 square degrees of the WIRCam Deep Survey (WIRDS) in J, H and K_s bands combined with optical data from the CFHTLS Deep fields. By deprojecting $w(\theta)$ analytically rather than using the Limber method, they find the correlation length to increase from $r_0 \simeq 4.5\,h^{-1}$Mpc at $z \simeq 0.5$ to $r_0 \simeq 6.3\,h^{-1}$Mpc at $z \simeq 1.75$.

Without reaching the same depth as ours and other authors' work, Quadri et al. (2007) use the Limber method to establish the correlation length and bias of a sample of $K < 21$ (Vega magnitude) galaxies at $2 < z < 3.5$ using four $10' \times 10'$ fields of the MUSYC survey. Their bias, at $z \simeq 2.6$, of 3.3 ± 0.5 is slightly greater than we would expect extrapolating beyond our $z \sim 2.16$ figure of $b = 2.23 \pm 0.12$. We observe a slightly fainter population of sources, but Quadri et al. also show that there is little to no significant effect of limiting K-magnitude on clustering at these magnitudes. They do note the possible limitation, however, of fixing the slope parameter $\beta = \gamma - 1 = 0.8$, as is done by Furusawa et al. (2011) (but not Bielby et al. 2013). The inset in Figure 5.3 shows that fixing the slope to 0.8 rather than fitting for it, we would artificially boost the amplitude parameter, A, and therefore r_0 and bias along with it. This partly addresses the difference between our highest-redshift result and that of Quadri et al. (2007) or even Furusawa et al. (2011), but not Bielby et al. (2013).

Ichikawa et al. (2007) provide severely limited constraints (due to a 24.4-square-arcmin area) on the clustering of K-selected sources at fainter levels ($K < 25$) using the MOIRCS Deep Survey in the GOODS-North region. Contrary to the shallower sample used by Quadri et al. (2007), they find a significant decline in r_0 with K-band magnitude at $1 < z < 2$ and weaker evidence of such a decline at $2 < z < 4$. Even so, there is a roughly 1–2σ disagreement in our results at $K < 23.5$, more consistent with their results at $K < 25$.

We have explored possible causes for discrepancy in our results compared with the aforementioned authors. Changing the range of our fits to $w(\theta)$ to match other work and fixing the slope to -0.8
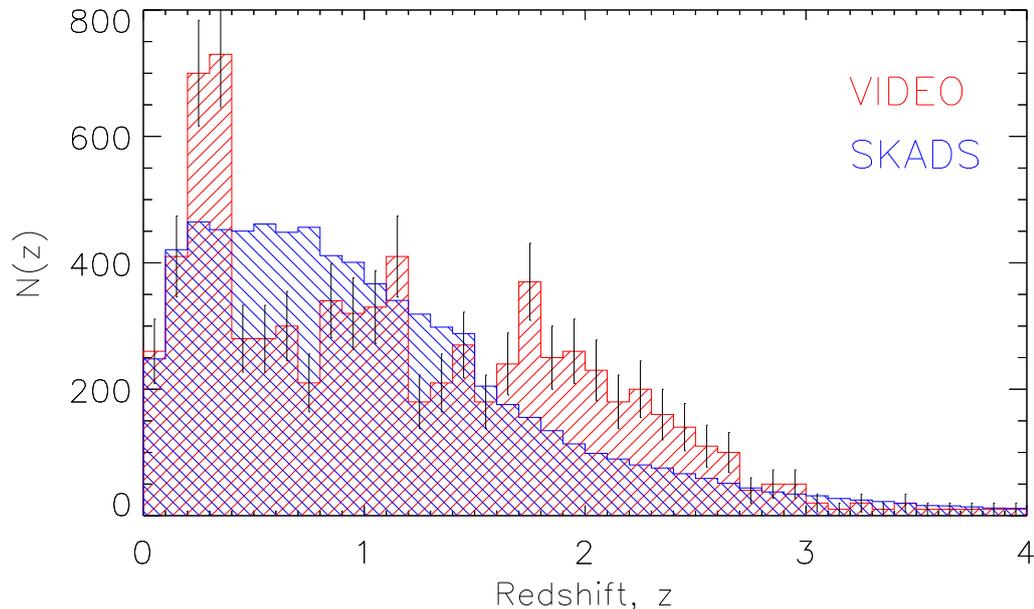
FIGURE 5.10: Redshift distributions of the observed radio sources in the VIDEO field (red), with Poisson error bars, and in the SKADS simulations (blue) using the same selection criterion at 1.4 GHz (all such sources also satisfied the $K < 23.5$ of the real data).

(as is common practice where data are sparse), has little combined effect on the resulting correlation lengths and biases. The difference then, may lie in the redshift distributions used, which do impact the spatial clustering measures, rather than the angular clustering amplitudes, which are similar in this work. In order to verify that our redshift distribution is robust, we have calculated r_0 for our data, but assuming the distribution of photometric redshifts from a similar sample taken from the deeper UltraVISTA survey (McCracken et al., 2012) over the COSMOS field, finding again that our results are not significantly altered. The redshift catalogue from Muzzin et al. (2013) includes photometry in 30 bands over the 1.62-square-degree COSMOS/UltraVISTA field, giving redshifts with a catastrophic outlier fraction of just 1.6 per cent and whilst their distribution is not identical to that of the VIDEO sources, the consistent results suggest that we may be confident in our results with these redshifts.

5.5.2   Clustering of faint radio sources

Our results show a strong evolution in the clustering and bias of faint radio sources with redshift. We also find that the more luminous subsamples are strongly biased, as would be expected if the dominant population were radio-loud AGN predominantly hosted by massive elliptical galaxies (e.g. Jarvis et al., 2001; Dunlop et al., 2003; Herbert et al., 2011).

While Wilman et al. (2003) investigated the clustering of a radio sample down to 0.2 mJy, direct measurements of the clustering of radio sources below the mJy level are generally lacking in the present literature. However, some predictions for the bias of our sample may be taken indirectly from the SKA Design Study (SKADS; Wilman et al. 2008, 2010) semi-empirical simulations of extragalactic radio continuum sources. These simulations provide a catalogue of radio sources over 400 square degrees to a depth of 10 nJy in 15 radio and mid-infrared bands, as well as providing K-band magnitudes and classifications of 5 different radio source types: normal star-forming galaxies, starbursts, radio-quiet AGN, FRI and FRII type AGN. Each of these source types are attributed a fixed halo mass and a bias based on these masses and the Mo & White (1996) bias model.

By imposing radio luminosity cuts to our sample at $L_{1.4} > 10^{23}$ W Hz^{-1} and higher, we effectively bias it towards AGN (Condon, Cotton & Broderick, 2002). The increasing minimum luminosity raises the median luminosity of these subsamples, and therefore the median redshift as the highest redshift sources must be extremely luminous to reach the radio detection limit. Within errors, and ignoring the imposed flattening of the SKADS bias prescription beyond $z = 1.5$ (see Wilman et al., 2008, for more details), these 'AGN-only' results are consistent with the SKADS FRII bias with a halo mass of $10^{14}\,M_\odot$. This suggests, in agreement with results of Chapter 4, that the typical halo mass of the more numerous FRIs (we expect ~ 5 FRIIs in the 1 square degree that we use here, compared with ~ 340 FRIIs) must be in this region as well. The increasing bias with luminosity (and therefore redshift), additionally, provides observational evidence that the truncation of the AGN bias models at $z = 1.5$ in Wilman et al. (2008) is not justified. While the reason for this plateau is to prevent an unphysical rise in the model bias beyond redshifts where it can be constrained in the literature, we show that the bias continues to evolve to $z > 2$.

To investigate this, we use a SKADS sample catalogue of $S_{1.4} > 90\ \mu$Jy radio sources, giving a redshift distribution (see Figure 5.10) and proportions for each source population, comprising FRI and FRII radio AGN, radio-quiet quasars, star-forming galaxies and starbursts. Weighting the five different bias models associated with these different populations by the relative proportions, we determine the expected bias for each of our data points shown in Figure 5.7. The difference in redshift distributions shown in Figure 5.10, however, demonstrates the potential effects of sample variance when using a small area such as this VIDEO field. Furthermore, we show these predictions for the case where we assume the same (higher) halo mass for the FRI sources as for the FRII sources, as suggested in Chapter 4. This boosts the bias as expected, particularly at the higher redshifts, consistent with our upper two redshift bins, but also exacerbates an apparent underestimate at low redshift.

Deviations from the SKADS predictions could be the result of miscalculating the proportions of AGN and normal galaxies – a greater than expected fraction of AGN at high-redshift would appear to increase the observed bias – but the luminosity functions used by Wilman et al. (2008) should be well-constrained so as to alleviate this potential issue. The halo masses of the different types of galaxy, and how they evolve with redshift however, are less well known. Assuming robust luminosity functions, our results would appear to suggest a less biased (i.e. less massive) population at low redshift and a more biased one at higher redshift, contrary to the single fixed mass model of SKADS. Although we note that the cosmological volume sampled at low redshift is relatively small and as such we expect the clustering signal to be dominated by sample variance, which we do not consider here as we are most interested in the high-redshift measurements, whereas at low-redshift large-area surveys are more suitable.

5.6 Conclusions

We have made use of complementary optical and infrared data from CFHTLS and VIDEO to infer the spatial clustering of radio sources at the > 90 μJy level in a 1 square degree field (VIDEO-XMM3) through the angular cross-correlation function. With a $K_\mathrm{s} < 23.5$ sample of $\sim 100,000$ galaxies with reliable photometric redshifts out to $z \sim 4$, we find the auto-correlation function, spatial correlation length, and linear bias of the full sample and four independent redshift bins. Furthermore, we use the cross-correlation of these sources with our deep radio sample of 766 galaxies to infer their bias. The results can be summarized as follows:

1. The slope of $w(\theta)$ for infrared-selected galaxies is generally consistent (~ -0.85) but slightly larger than the canonical value of -0.8 often assumed in the literature. This does not significantly affect our results, and we find an increasing bias and correlation length in qualitative agreement with similar infrared studies, but with significantly weaker clustering than found elsewhere due to a better constrained redshift distribution.

2. By cross-correlating with this background of infrared galaxies, we are able to find a robust cross-clustering signal for the radio galaxies across 4 redshift bins up to a median redshift of $z = 2.15$.

3. Combining the auto- and cross-correlation function results, we are able to disentangle the clustering scales of the two galaxy populations to find a radio source bias which increases from 0.57 ± 0.06 at $z \simeq 0.29$ to 8.55 ± 3.11 at $z \simeq 2.15$.

4. Placing lower limits on radio luminosity at 10^{23}, $10^{23.5}$ and 10^{24} W Hz^{-1} effectively reduces the sample to an AGN-dominated subsample, giving a high-redshift measurement of the AGN bias of $b(z = 1.37) = 7.62 \pm 1.27$, $b(z = 1.55) = 9.91 \pm 2.48$ and $b(z = 1.77) = 11.14 \pm 3.01$, respectively.

5. The radio bias found at higher redshifts is greater than that expected by assuming the models used in the SKADS radio simulations. However, assigning a similar halo mass to the FRI sources as assumed for the FRII sources in our radio sample largely addresses this discrepancy. Indeed, results using the high-luminosity subsample appear to confirm that low-radio-luminosity AGN have a typical halo mass similar to that assigned to FRII radio galaxies (10^{14} $M_\odot$).

6

Conclusions

"The possession of knowledge does not kill the sense of wonder and mystery. There is always more mystery."

Anaïs Nin

This chapter summarizes the results of Chapters 2–5 before going on to discuss how the clustering of radio galaxies relates to the broader study of galaxy evolution. I then describe several ongoing projects related to the work presented in this thesis: 2D spatial cross-correlation of low-redshift radio sources with the GAMA spectroscopic survey; halo occupation distribution analysis of deep VIDEO data; and cross-correlation of radio emission with the CMB.

6.1 Summary

In this thesis, I have used radio sources to trace large-scale structure in the Universe, with the implicit goal of following the evolution with cosmic time. In the first instance, understanding the distributions of galaxies in space (rather than simply viewed in projection), requires a measure of distance to an observed source, in addition to the angular positions on the sky (right ascension and declination). This preliminary obstacle is addressed in Chapter 2, where I describe the spectroscopic observations of a radio-selected sample of galaxies at $S_{1.4} > 100$ mJy, in the GAMA/H-ATLAS fields, and subsequently calculate redshifts from the emission lines found in 30 out of 79 targets observed in 5 nights on the

NTT. The low success rate of 38 per cent is indicative of the nature of these powerful radio sources; the synchrotron radio emission is able to traverse vast cosmological distances, unabsorbed by intervening gas and dust, while the optical emission is far weaker, as well as being subject to absorption en route to the observer. Additionally, this work made clear the importance of accounting for the unique problem (to radio observations) of a single host AGN being the driver of multiple discrete radio detections, separated by distances approaching 1 Mpc (or $\sim 2'$ on the sky, at $z \sim 1$). The detection rate for the host AGN of such multiple-component radio sources in the NTT observations was at just 25 per cent, compared to 47 per cent for the compact sources.

An alternative approach was taken in Chapter 3, looking towards a more automated approach for building larger samples of radio sources with accurate redshifts. This was done by taking a 144-square-degree redshift survey (GAMA) and finding as many matches as possible from the FIRST radio survey, down to a flux-density limit of 1 mJy. The issue of multiple-component radio sources was accounted for by collapsing groups of sources (within $72''$) and cross-matching to the flux-weighted mean position of these groups. The spatial correlation function was used, for a sample of 1,635 matched radio sources in GAMA, to find a correlation length of $r_0 \simeq 8.2 \, h^{-1}$Mpc at $z \simeq 0.34$, and infer a bias of $\simeq 1.9$.

The matching success rate of radio sources in the GAMA spectra (limited to $r < 19.8$) was so low as to preclude any attempted redshift binning to observe a change with redshift of the clustering properties of these sources. In any case, it is predominantly the low-redshift, optically-brighter galaxies being observed in this sample, rather than the more massive, more distant AGN, which are the more highly biased tracers of mass in the Universe. Using 9-band photometric redshifts from SDSS/UKIDSS (and an increased area of 210 square degrees), a matched sample of 3,886 photometric and/or spectroscopic redshifts was produced, at a higher median redshift of $z = 0.48$. The spatial clustering properties were derived from the angular correlation function of this sample, along with a high- and low-redshift division, giving some shape to the evolution of the bias of these radio sources. Using redshift distributions from the SKADS simulated radio catalogues, however, we have made full use of the remaining unmatched radio sources, calculating their $N(z)$, with which r_0 and the bias can be inferred, as for the matched sample. This extended the redshift range we could investigate, the unmatched sources being estimated to have a median redshift of $z = 1.55$. The corresponding correlation length $r_0 \simeq 14$ h^{-1}Mpc, and bias $b \simeq 9.5$, while being measured indirectly through their projected angular clustering, and the use of a model redshift distribution, reach a high redshift which has been hitherto poorly studied.

The SKADS simulated data were again used to estimate the relative proportions of different radio

galaxy populations (FRIs and FRIIs, radio-quiet quasars, starburst galaxies, and normal star-forming galaxies) for each radio sample studied, by weighting SKADS sources to match the observed redshift distributions. These were used, with the bias models of Mo & White (1996), and the halo masses assumed by Wilman et al. (2008) in producing the SKADS simulations, to compare our observations with these models. At low redshift (the matched sources), the results agree with the model bias, but our observations find the radio population to be much more strongly biased than the model at higher redshift. We show, however, that increasing the assumed halo mass of the more numerous FRI galaxies ($M = 10^{13} \, M_\odot$) to match that of the FRIIs ($M = 10^{14} \, M_\odot$) brings the model towards closer agreement with the results. The conclusion drawn is that the halo masses of these AGN are being underestimated and/or there are simply a greater proportion of massive AGN than previously thought, and estimated from the luminosity functions employed by Wilman et al. (2008), although these results could also be explained in part by e.g. bias in the redshift estimation of the radio sources, differing proportions of FRII to FRI sources to those defined in SKADS simulations, or the bias model assumed in SKADS underestimating the true bias at high redshift.

Chapter 5 does not rely on simulated data or assumed redshift distributions, instead maximizing efficiency by cross-matching deeper data, in both the radio sample and optical/IR photometric redshift catalogues, over a smaller field of 1 square degree. Matching a deep radio sample from VVDS ($S_{1.4} > 90 \, \mu$Jy) with $\sim 100{,}000$ reliable photometric redshifts out to $z \sim 4$ with VIDEO/CFHTLS ($K_\mathrm{s} < 23.5$), the matching success rate is far higher. Despite the smaller radio sample size of just 766, the angular cross-correlation between these sources and the full K_s-selected VIDEO sample is used to find a significant clustering signal, and a measure of the bias in 4 redshift bins between $z = 0.29$ and $z = 2.15$ ($b = 0.57 \pm 0.06$, increasing to $b = 8.55 \pm 3.11$). A model bias was estimated similarly to those of Chapter 4, with our results in reasonable agreement at low redshift, but showing a more highly-biased population at high-redshift than predicted by the model. Again, increasing the assumed mass of the FRI galaxies brings the model into good agreement with our results, confirming the conclusions of Chapter 4. Finally, further weight was added to this argument by applying radio luminosity cuts of $L > 10^{23} \, \mathrm{W \, Hz}^{-1}$ and higher, investigating the bias of the AGN-fuelled radio sources. These are found to be consistent with a halo mass of $10^{13} \, M_\odot$ (assumed for FRIIs in SKADS), and with a continuing rise in clustering strength beyond $z = 1.5$.

These measurements of the radio bias will become crucial in the application of future radio surveys to constrain cosmological models. Cosmological probes such as BAO, the ISW effect, and cosmic magnification (described in Section 1.2.5), will use these results to disentangle the radio clustering

properties from the underlying cosmological information revealed by these phenomena (Raccanelli et al., 2012; Camera et al., 2012; Ferramacho et al., 2014).

6.2 Links to Galaxy Evolution

The results presented in this thesis, and particularly those at higher redshift, have shown that radio galaxies cluster on scales between that of normal galaxies and clusters. While at low redshift, these may include normal galaxies where the radio emission is a tracer of star formation, the sources with higher radio luminosities (such as the majority of those in FIRST, and the subset of those in the VVDS-VLA survey) represent synchrotron emission arising from jets formed in AGN. Radiative and mechanical feedback from AGN is invoked to curtail black hole and galaxy growth, with star formation being suppressed (e.g. Croton et al., 2006; Bower et al., 2006; Dubois et al., 2010; Vogelsberger et al., 2014; Taylor & Kobayashi, 2015), but these black hole accretion and feedback processes are not fully understood, and there appears to be a necessary distinction between the radiative feedback and mechanical feedback from the jets in radio-loud AGN.

There is no correlation between the radio and emission-line luminosity of low radio-luminosity AGN, and the probability of radio loudness, at a given galaxy mass, is independent of whether the galaxy hosts an optically-identified AGN (Best et al., 2005). This suggests that optical AGN and low radio-luminosity AGN (mostly FRIs) are independent, and triggered by different physical mechanisms. Often these are described in terms of cold (or 'quasar') mode or hot (or 'radio') mode accretion, respectively (Croton et al., 2006). In high radio-luminosity AGN (mostly FRIIs), however, radio luminosity is strongly correlated with emission-line luminosity, showing the further distinction required within the subset of radio-loud AGN. The feedback introduced by AGN accretion processes can be broadly separated into radiative and mechanical contributions, adding thermal and kinetic energy to the surrounding medium, respectively. Mocz, Fabian & Blundell (2013) model these methods of feedback, distinguishing between FRIIs with high accretion rates and FRIs with low accretion rates both contributing kinetically due to their powerful radio jets, with quasars showing the same high accretion rate as the FRIIs but adding only radiative feedback to their surroundings. The FRII feedback dominates above $z \sim 1.5$ with FRIs dominating below, and radiative quasar feedback being subdominant at all epochs, but little work has been done on the mechanical feedback from powerful radio sources which dominate at high redshift.

The results presented in this thesis are in line with those of Mandelbaum et al. (2009) who found that

radio-loud AGN are hosted by more massive galaxies, and more massive haloes, than those of radio-quiet AGN or other radio galaxies. Hatch et al. (2014) also finds that radio-loud AGN reside in denser environments than radio-quiet analogues, and are commonly found in protoclusters – the progenitors of $z = 0$ clusters – at high-redshift. Given that mechanical feedback arises from the radio-loud AGN population, it therefore becomes very difficult to disentangle the role of the environment and that of AGN activity in determining the key physical mechanisms governing massive galaxy evolution. More subtle insight into this complex problem requires understanding of the separate contributions of the FRIs and FRIIs, which have been assumed to reside in haloes of different mass, but which may not be the case. High-excitation radio galaxies (HERGs) and low-excitation radio galaxies (LERGs) share a similar dichotomy in accretion modes to the FRIIs and FRIs, respectively, with it being suggested that it is the LERGs with a higher mass, more likely to be hosted by central galaxies of groups and clusters (Best & Heckman, 2012). An understanding of the clustering of these AGN and relating it to the dark matter haloes is therefore imperative to comparing with simulations (often founded in some way upon a halo occupation distribution model; see Section 6.3.2), and constraining more sophisticated feedback models. Of particular relevance to this would be to carry out a proper study of the clustering of powerful AGN, which are assumed to provide mechanical feedback, in the simulations and compare with the results of this thesis.

6.3 Ongoing and Future Work

Following on from the work described in this thesis, there are several projects for which there has been insufficient time to present meaningful results in an additional chapter. Some of this work is ongoing, and some will be the subject of future study by myself and collaborators. In this section, three such avenues of investigation will be discussed.

6.3.1 Spatial Cross-Correlation in GAMA

Chapter 3 shows the bias we can infer from the spatial clustering of a sample of radio sources with accurately known redshifts, but these results are limited by this requirement of spectroscopic redshifts. In Chapter 5, a significant clustering signal was found for far smaller radio samples by employing the cross-correlation with a more densely populated infrared-selected sample. The results using these VIDEO data were, however, limited by being derived from clustering viewed in angular projection, and relying on the assumptions concerned with the Limber equation. The natural extension to this work is to
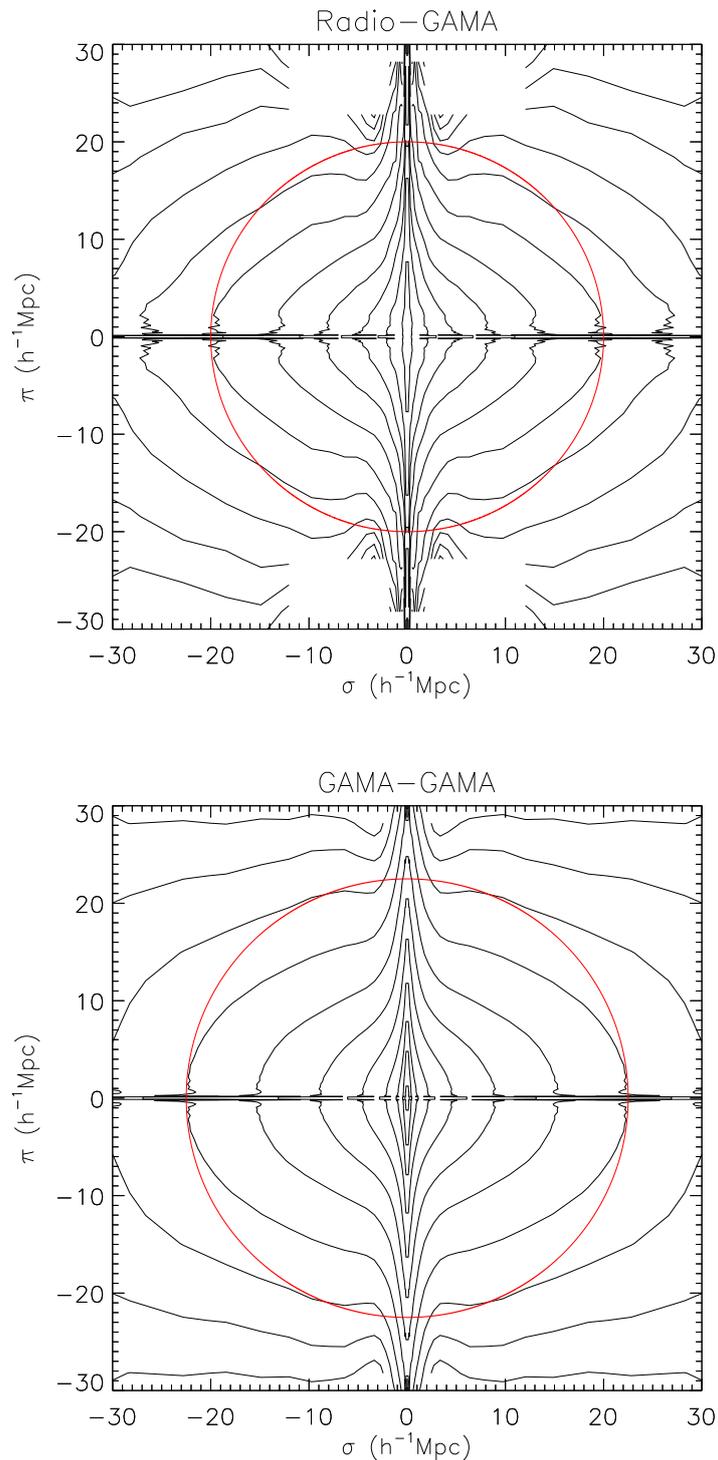
FIGURE 6.1: Contour plots of the 2D spatial correlation function $\xi(\sigma, \pi)$ (cf. Figure 1.5). The top panel represents the cross-correlation function of GAMA Phase I galaxies and the subset of those galaxies identified with FIRST radio sources (see Section 3.2), and the bottom panel represents the auto-correlation function of the GAMA sample. Overplotted in red is a circle, showing the expected contour shape in absence of redshift-space effects.
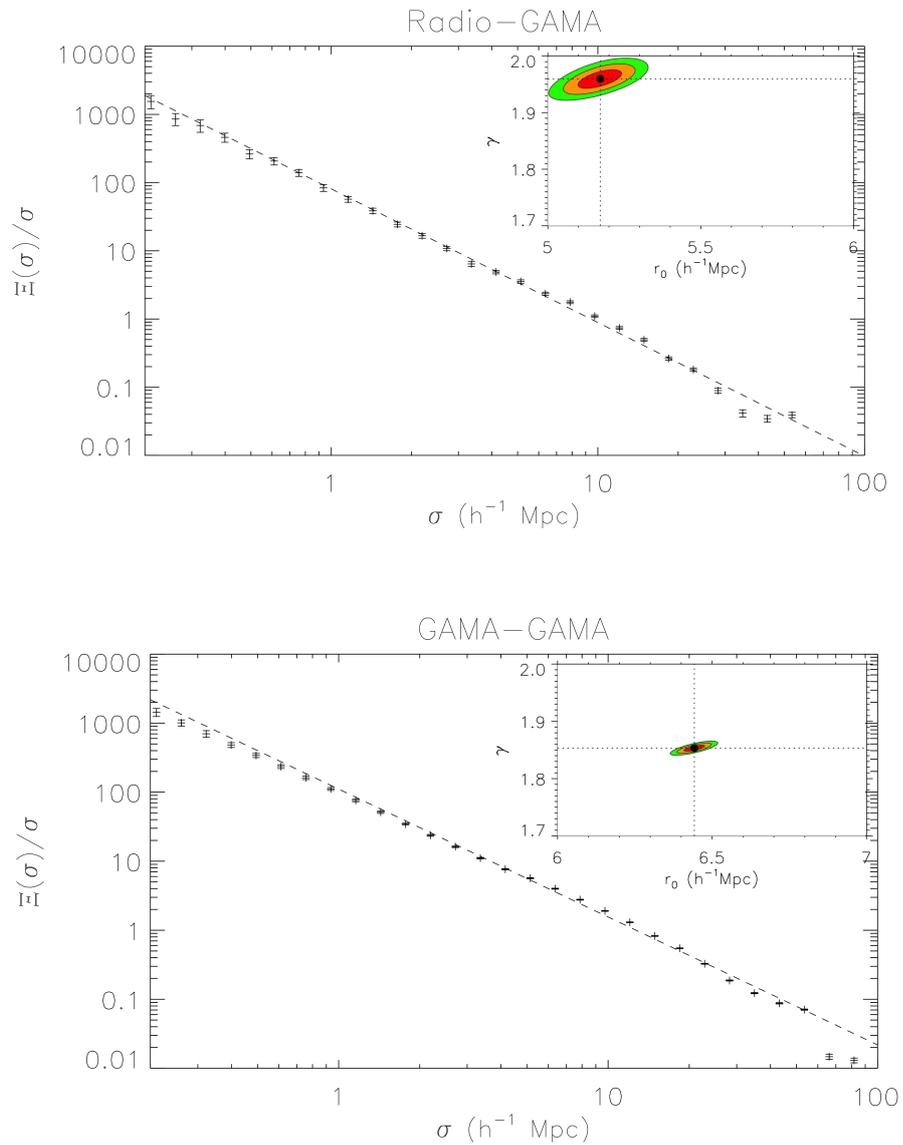
FIGURE 6.2: The projected cross-correlation function of GAMA Phase I galaxies and the subset of those galaxies identified with FIRST radio sources (top panel), and the projected auto-correlation function of the GAMA sample (cf. Figure 3.7, bottom panel). The dashed lines shows the best fit power-laws (over the range $\sigma < 60\ h^{-1}\mathrm{Mpc}$), and the inset plot shows parameter fits at $\Delta\chi^2 = 2.30$, 6.18 and 11.83.

cross-correlate the sample of 1,635 radio sources at $S_{1.4} > 1$ mJy, with GAMA spectroscopic redshifts, with the full GAMA catalogue (127,803 sources). Combined with the auto-correlation function of the larger GAMA catalogue, in a similar manner to the method of Chapter 5, the bias of both radio and non-radio samples can be calculated over a range of redshifts (albeit at lesser distances than reached by the VIDEO/CFHTLS photometric redshifts; $z \lesssim 1$). Figures 6.1 and 6.2 show some preliminary results to this end.

The quality of the GAMA data is highlighted in Figure 6.1, where the 2D spatial correlation function, $\xi(\sigma, \pi)$, is shown in the form of contours, symmetric about $\pi = 0$ and $\sigma = 0$, illustrating redshift-space distortions in the observed galaxy distribution. A circle is plotted in red, representing the concentric circles that would be expected from such contours, with real-space distances used rather than redshift-space distances. Both the cross-correlation and auto-correlation functions clearly show the flattening of large-scale structures at separations of tens of Mpc (the Kaiser effect, due to infall), and the apparent elongation of structures along the line-of-sight direction, π, at separations less than a few Mpc ("fingers of God", due to peculiar velocities). Such detail could not be seen from the auto-correlation function of the small sample of radio sources used in Chapter 3, but the clarity of these features suggests that integrating along the π direction to correct for these redshift-space distortions will produce a healthy clustering signal in the projected correlation function, $\Xi(\sigma)$.

Figure 6.2 shows the results of this integration, and subsequent fitting for the real-space correlation length, r_0, and slope parameter, γ. Incorporating the full redshift range of the catalogues, the errors on $\Xi(\sigma)$ are far smaller than those found in Chapter 3, with suitably strong constraints on r_0. Using a model dark matter power spectrum to estimate the matter correlation function, a direct measure of the bias can also be made (e.g. Ross et al., 2009). Binning $\xi(\sigma, \pi)$, additionally, by redshift, the signal-to-noise from these functions should remain high, precisely tracing the clustering evolution of both populations (radio and non-radio) at $z < 1$, improving significantly on the low-redshift results of Chapter 4.

6.3.2   Halo Occupation Distribution Modelling with Deep Fields

An improvement that could be made to the work of Chapters 4 and 5 is the use of a more sophisticated treatment of $w(\theta)$ and the way in which we infer the bias from it. Halo occupation distribution (HOD) models (see e.g. Cooray & Sheth, 2002, for further details) describe the probability distribution for the number of galaxies N hosted by a dark matter halo as a function of its mass M. This allows galaxy clustering to be tackled from the perspective of an intuitive relationship between the mass of dark matter haloes and the number of galaxies found in a given halo (e.g. Blake, Collister & Lahav, 2008; Ross,
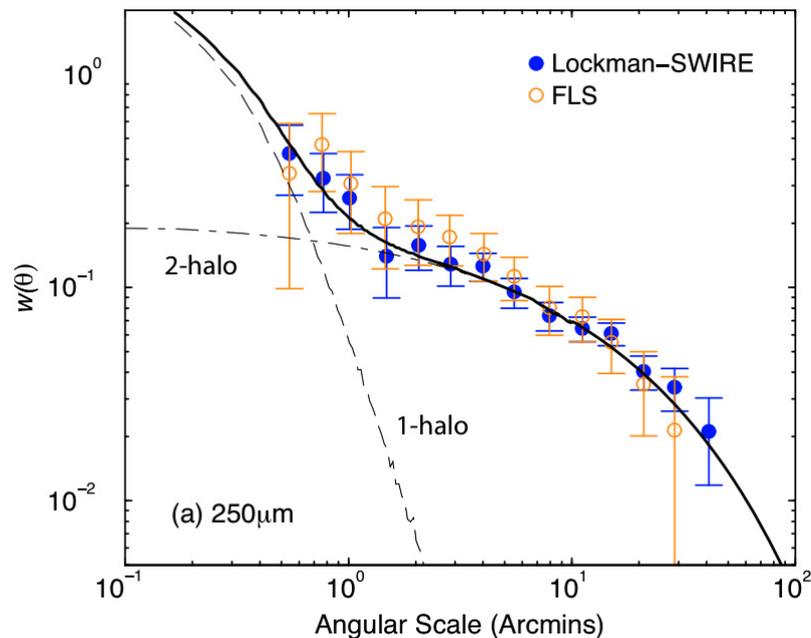
FIGURE 6.3: The angular correlation function of 250-μm-selected sources in two wide fields with HerMES (Cooray et al., 2010). The lines overplotted are illustrative halo models, showing the contribution from the 1-halo term (multiple sources in the same halo) and the 2-halo term (linear clustering).

Percival & Brunner, 2010; Cooray et al., 2010; Geach et al., 2012). Separate mean occupations are assumed for core galaxies found in the centre of a dark matter halo, and satellite galaxies, described by:

$$\langle N|M\rangle = \langle N_{\mathrm{c}}|M\rangle \left[1 + \langle N_{\mathrm{s}}|M\rangle\right], \tag{6.1}$$

where the satellite galaxies are distributed independently of whether a central galaxy also exists (although a central galaxy is not always necessarily assumed to be a prerequisite for satellites). With the application of a chosen model, the contributions to the galaxy power spectrum can be calculated and used to reproduce a correlation function to be fitted to observed data. Figure 6.3 shows an example, taken from Cooray et al. (2010), with a fitted $w(\theta)$ based on such an HOD model for sub-mm sources in HerMES. The fit to the data is demonstrably better than a simple power law fit, and a galaxy bias can be derived from the same HOD fit, when provided with a model for the bias as a function of halo mass (e.g. Sheth, Mo & Tormen, 2001).

Application to Results of Chapter 5

It is clear from Chapter 4 that the GAMA matched sources are too sparse to differentiate a power-law form of $w(\theta)$ to any more sophisticated form. The VIDEO data in Chapter 5 better constrain the shape, but still with a quite noisy cross-correlation function, and the narrow dimensions of the field limit the
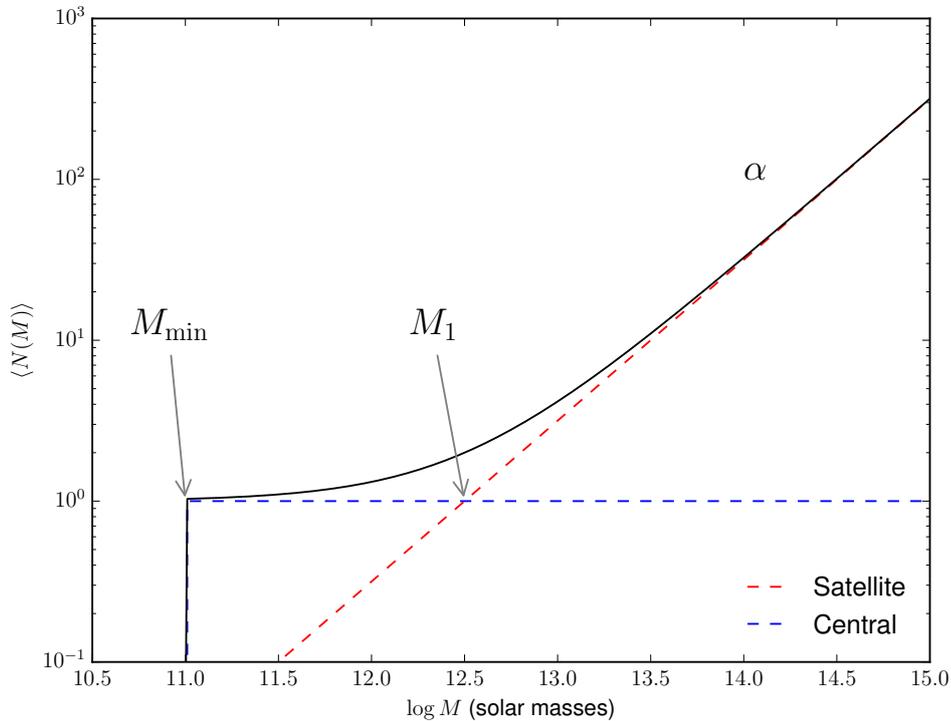
FIGURE 6.4: Example halo occupation distribution function, demonstrating the significance of the three model parameters in a simple HOD model. M_{min} represents the halo mass above which all dark matter haloes are assumed to have a galaxy at their centre. M_1 represents the halo mass above which a halo is expected, on average to have one satellite galaxy, with α describing how steeply the number of satellite galaxies per halo increases with halo mass as a power law.

usable range to $\theta \lesssim 0.1°$. The results of some preliminary work to interpret these data within the framework of an HOD model are presented in this section. This was performed with an adaptation to the `halomod` Python package, written by Steven Murray, which allows the use of several HOD models, dark matter halo profiles and halo mass functions (Murray, Power & Robotham, 2013). I have chosen to use the simplest model from Zehavi et al. (2005), described as follows:

$$\langle N_{\mathrm{c}}|M\rangle = \begin{cases} 0 & \text{if } M < M_{\mathrm{min}} \\ 1 & \text{if } M \geq M_{\mathrm{min}} \end{cases},$$

$$\langle N_{\mathrm{s}}|M\rangle = \left(\frac{M}{M_1}\right)^{\alpha}, \tag{6.2}$$

where M_{min} is the minimum dark matter halo mass which can host a central galaxy, M_1 corresponds to the mass of haloes that contain, on average, one additional satellite galaxy ($\langle N_{\mathrm{s}}(M_1)\rangle = 1$) and α sets the rate at which haloes accumulate satellites when growing in mass (see Figure 6.4). The transition at M_{min} is modelled as a step function, but, in reality, this is likely to have some scatter $\sigma_{\log M}$ (e.g. Ross

et al., 2009; Cooray et al., 2010; Geach et al., 2012):

$$\langle N_c|M\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log\left(M/M_c\right)}{\sigma_{\log M}}\right)\right].$$ (6.3)

As shown in Figure 6.3, contributions to the clustering amplitude can be separated into those arising from pairs of galaxies within the same halo (one-halo term) and those in different haloes (two-halo term). Expressed in these terms, the spatial correlation function is given by

$$\xi(r) = [1 + \xi_{1h}(r)] + \xi_{2h}(r).$$ (6.4)

The one-halo term is further separated into contributions from central–satellite and satellite–satellite galaxy pairs. The former is given by

$$\xi_{1h}^{c-s} = \frac{2}{n_g^2}\int_{M_{vir}(r)}^{\infty} dM \frac{dn(M)}{dM} N_c(M) N_s(M) \frac{\rho(r,M)}{M},$$ (6.5)

where n_g is the galaxy number density, $n(M)$ is the halo number density, $\rho(r,M)$ is the halo density profile, taken from Navarro, Frenk & White (1997), and $M_{vir}(r)$ is the virial mass, as a function of scale. The satellite–satellite contribution is a convolution of the halo mass function with the halo density profile, given in k-space as

$$P_{1h}^{s-s}(k) = \frac{2}{n_g^2}\int_0^{\infty} dM \frac{dn(M)}{dM} N_c(M)[N_s^2(M)u(k,M)]^2,$$ (6.6)

where $u(k,M)$ is the Fourier transform of $\rho(r,M)$. The corresponding correlation function contribution is

$$\xi_{1h}^{s-s}(r) = \frac{1}{2\pi^2}\int_0^{\infty} dk P_{1h}^{s-s}(k) k^2 \frac{\sin(kr)}{kr},$$ (6.7)

and we obtain the one-halo term:

$$\xi_{1h}(r) = \xi_{1h}^{c-s}(r) + \xi_{1h}^{s-s}(r) - 1.$$ (6.8)

The two-halo term can be calculated in Fourier space from the dark matter correlation function and a bias model, taken from Tinker et al. (2005):

$$P_{2h}(k,r) = P_{matter}(k)\left[\int_0^{M_{lim}(r)} dM \frac{dn(M)}{dM} b(M,r) \frac{N(M)}{n_g'(r)} u(k,M)\right]^2,$$ (6.9)

where $P_{matter}(k)$ is the matter power spectrum from CAMB (Lewis & Challinor, 2011) using the fitting formulae of Smith et al. (2003) and $b(M,r)$ is the scale-dependent bias of haloes, expressed (Tinker et al., 2005) as

$$b^2(M,r) = B^2(M)\frac{[1 + 1.17\xi_{matter}(r)]^{1.49}}{[1 + 0.69\xi_{matter}(r)]^{2.09}},$$ (6.10)

where $B(M)$ is the bias model of Sheth, Mo & Tormen (2001) and $\xi_{\text{matter}}(r)$ is the non-linear dark matter correlation function. $M_{\text{lim}}(r)$ is the mass limit due to halo exclusion, determined by the methods of Tinker et al. (2005) and Blake, Collister & Lahav (2008). The restricted number density, n'_g, is that of galaxies in haloes up to the halo exclusion limit.

Given $\xi(r)$, the angular correlation function for our HOD model is given by the Limber equation (Section 1.2.2):

$$w(\theta) = 2 \int_0^\infty \mathrm{d}x \, f(x)^2 \int_0^\infty \mathrm{d}u \, \xi\left(\sqrt{u^2 + x^2\theta^2}\right), \tag{6.11}$$

where $f(x)$ describes the radial distribution of sources:

$$f(x) = N(z) \left[\frac{\mathrm{d}x(z)}{\mathrm{d}z}\right]^{-1}, \tag{6.12}$$

$N(z)$ and $x(z)$ being the redshift distribution of the galaxy population and the comoving radial coordinate (equation 3.7), respectively.

The data used are as presented in Chapter 5, with the exception that for the purposes of the HOD fitting, the cross-correlation function is not used and the radio auto-correlation function is estimated instead as follows:

$$w_{\text{radio}}(\theta) \simeq \frac{w_{\text{cross}}^2(\theta)}{w_{\text{IR}}(\theta)}. \tag{6.13}$$

This neglects the correction applied to equation 5.8 for slightly differing redshift distributions, but this is a small effect. The HOD model is then fitted to the data by a MCMC method, with flat priors chosen arbitrarily for the three parameters[†]:

$$10^5 \leq M_{\text{min}} < 10^{15},$$
$$M_{\text{min}} \leq M_1 < 10^{20}, \tag{6.14}$$
$$0 \leq \alpha < 2.5.$$

Given that `halomod` does not yet have an integrated fitting procedure for angular correlation functions, the new HOD parameters were propagated through all of the above calculations for each point in the MCMC chains, very inefficiently. As such, the results shown are limited to 6 chains of 2,000 points each per fit.

Figure 6.5 shows the $z < 0.5$ infrared- and radio-selected angular correlation functions fitted with a the HOD model described above. Two lines are overplotted in red and green, defined by the median parameters and the best fit (minimum χ^2) parameters, respectively. Figure 6.6 shows the distributions of these three parameters in the MCMC chains of 12,000 points, showing that α is particularly poorly

---

[†] M_{min} and M_1 given in units of solar masses.
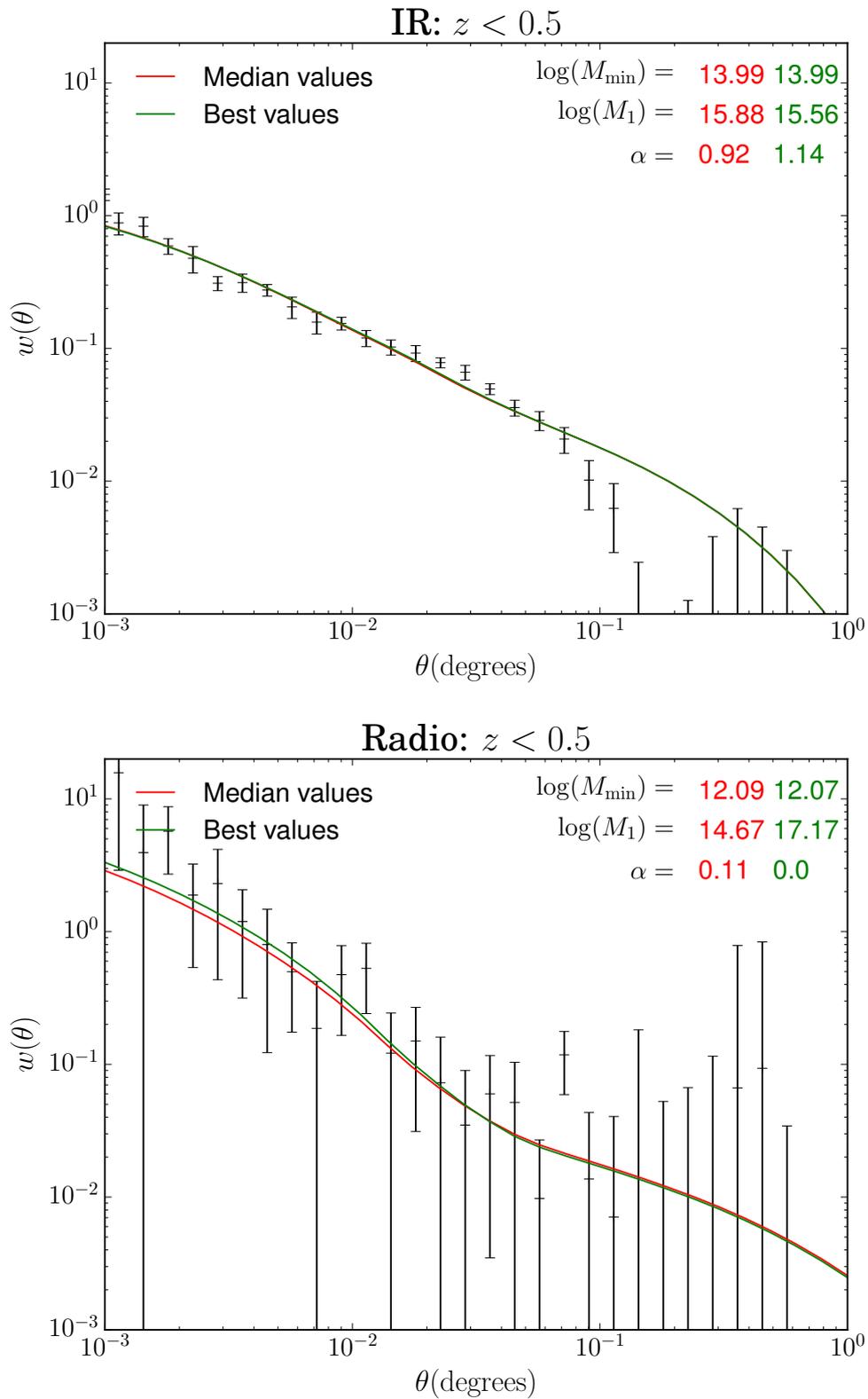
FIGURE 6.5: Example HOD model fits to the IR (top panel) and radio (bottom panel) correlation functions at $z < 0.5$. Values shown for the M_{min} and M_1 parameters are in solar masses.

FIGURE 6.6: Example HOD model parameter distributions from the combined MCMC chains for the IR (top panel) and radio (bottom panel) samples at $z < 0.5$. The dotted lines on the histograms show the Values shown for the M_{min} and M_1 parameters are in solar masses.

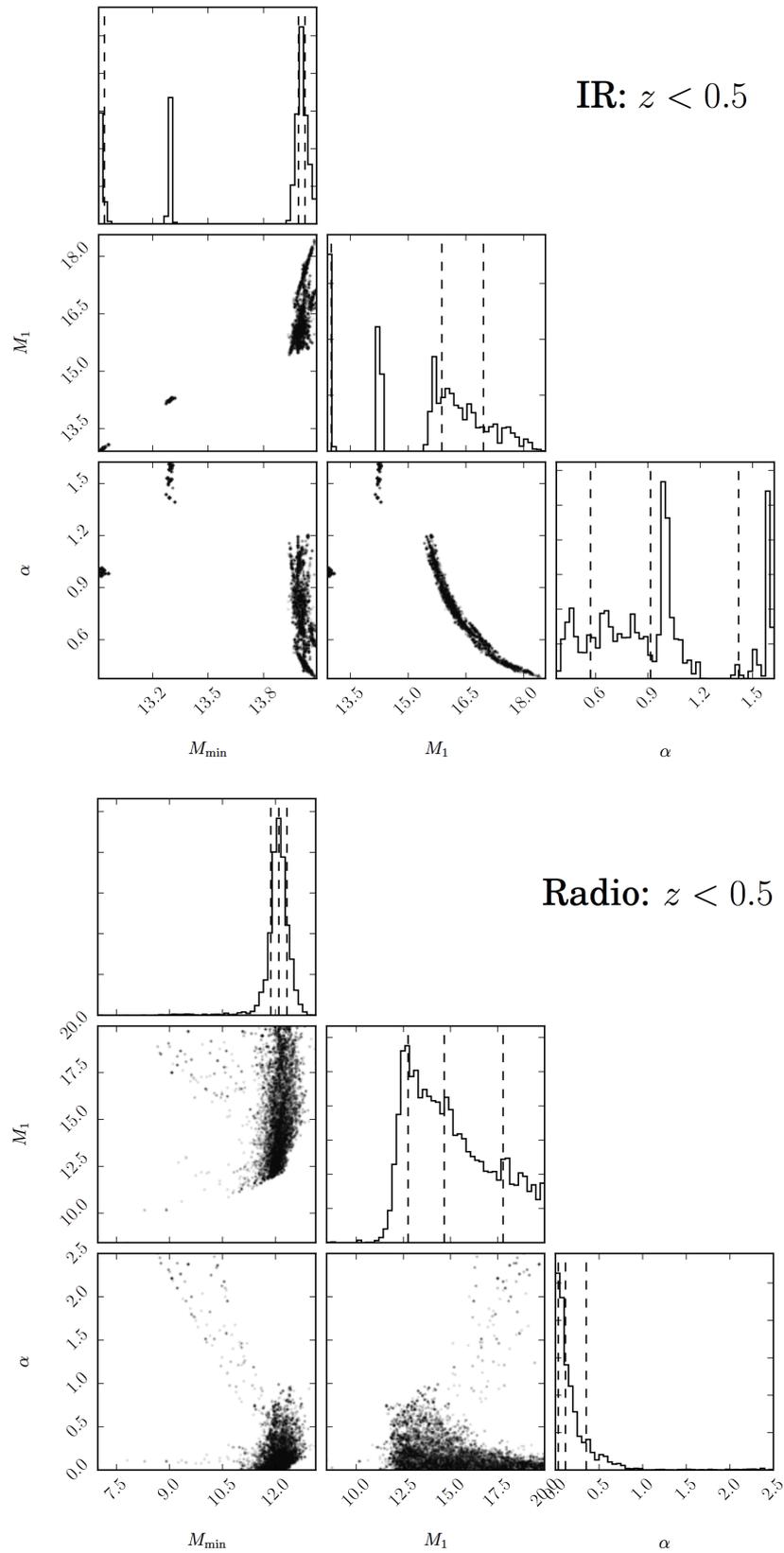constrained for the IR case, and being forced to zero (the minimum allowed value in the fit) in the radio case, suggesting no increase in halo occupation number with mass. For comparison, results from simulations (e.g. Zheng et al., 2005) and observations (e.g. Blake, Collister & Lahav, 2008; Ross et al., 2009; Ross, Percival & Brunner, 2010; Sawangwit et al., 2011; Coupon et al., 2012; Kim et al., 2014) typically find $\alpha \approx 1$, or greater, across a range of luminosities, redshifts and populations. Due to the noisy parameter distributions, the median and best-fit parameters shown here can be significantly different. Figures 6.7 and 6.8 show these same plots in the remaining three redshift bins, for the infrared- and radio-selected correlation functions, respectively.

In the case of the fits to the IR $w(\theta)$, we see that the MCMC fitting establishes a well-defined region in the parameter space, fitting the data reasonably well, albeit with some degeneracies apparent between M_1 and α ($0.5 < z < 1$), and between M_1 and $M_{\min}$ ($1.75 < z < 4$). In the $1 < z < 1.75$ case, however, α has been forced towards zero (from initial values of $\alpha = 1.0 \pm 0.5$), resulting in very poor fits to the data. It is worth noting that one may not expect a good fit to the near-infrared data as they describe a sample with a large range of masses as varying as a function of redshift, but this is being investigated in more detail in Hatfield, Lindsay, Jarvis et al. (in prep).

The fits to the radio correlation functions appear to suffer simply due to the poor signal-to-noise of $w(\theta)$, particularly as redshift increases. In all three redshift bins at $z > 0.5$, the parameter space is explored in almost its entirety (as allowed by the assumed priors), with no discernible preference, rendering the median parameter fits meaningless. Despite this, the best-fit parameters give a $w(\theta)$ that represents the data quite well (except for the highest-redshift bin). More importantly, they may fit the data better than would a power-law model as previously assumed throughout this thesis, although the large error bars remain a more significant problem. This may be promising for the application of this technique to more extensive data in the future, but would appear to rule out a meaningful analysis here.

Deep radio surveys, similar to that used in Chapter 5, over larger areas (e.g. MeerKAT and ASKAP and eventually SKA), overlapping with optical/IR photometric surveys can improve upon both the errors on $w(\theta)$ and the range over which it can be fitted with an HOD model, perhaps more sophisticated even than the 3–parameter model shown here in equation 6.2. Directly expanding upon the data used in this thesis, EVLA B-array observations carried out in December 2013 over 4.5 square degrees of the VIDEO XMM–LSS field, using the full 1–2 GHz bandwidth available at L-band, will reach a 5σ depth of 50 μJy. This is a further improvement on the 90 μJy used over the 1 square degree studied in Chapter 5, with $\sim 8,000$ detections expected as a result.

Bridging the gap in both area and depth between this survey and FIRST, for example, another
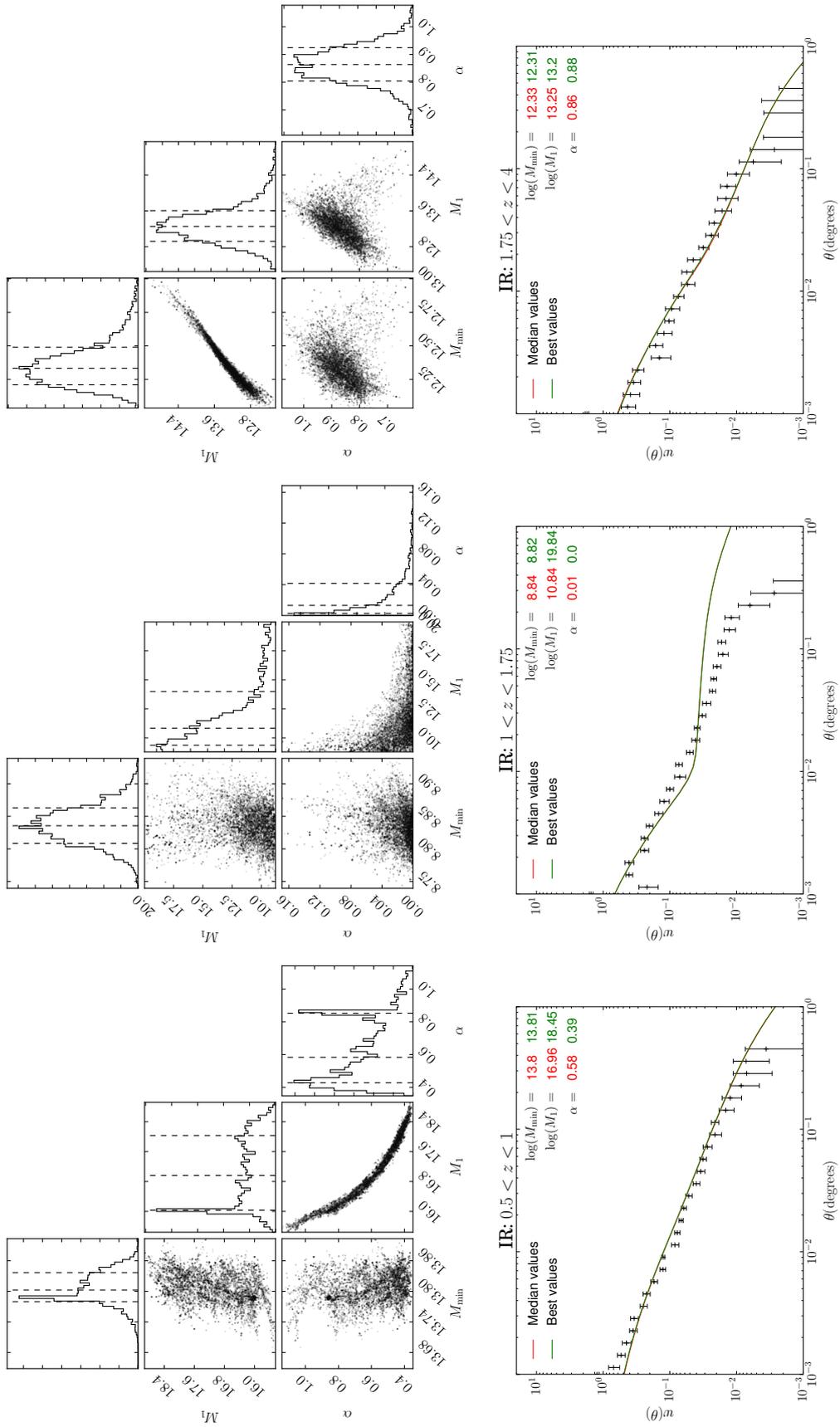
FIGURE 6.7: HOD model fits to the $0.5 < z < 1$ (left), $1 < z < 1.75$ (centre) and $1.75 < z < 4$ (right) angular correlation function of the K_s-selected sources.
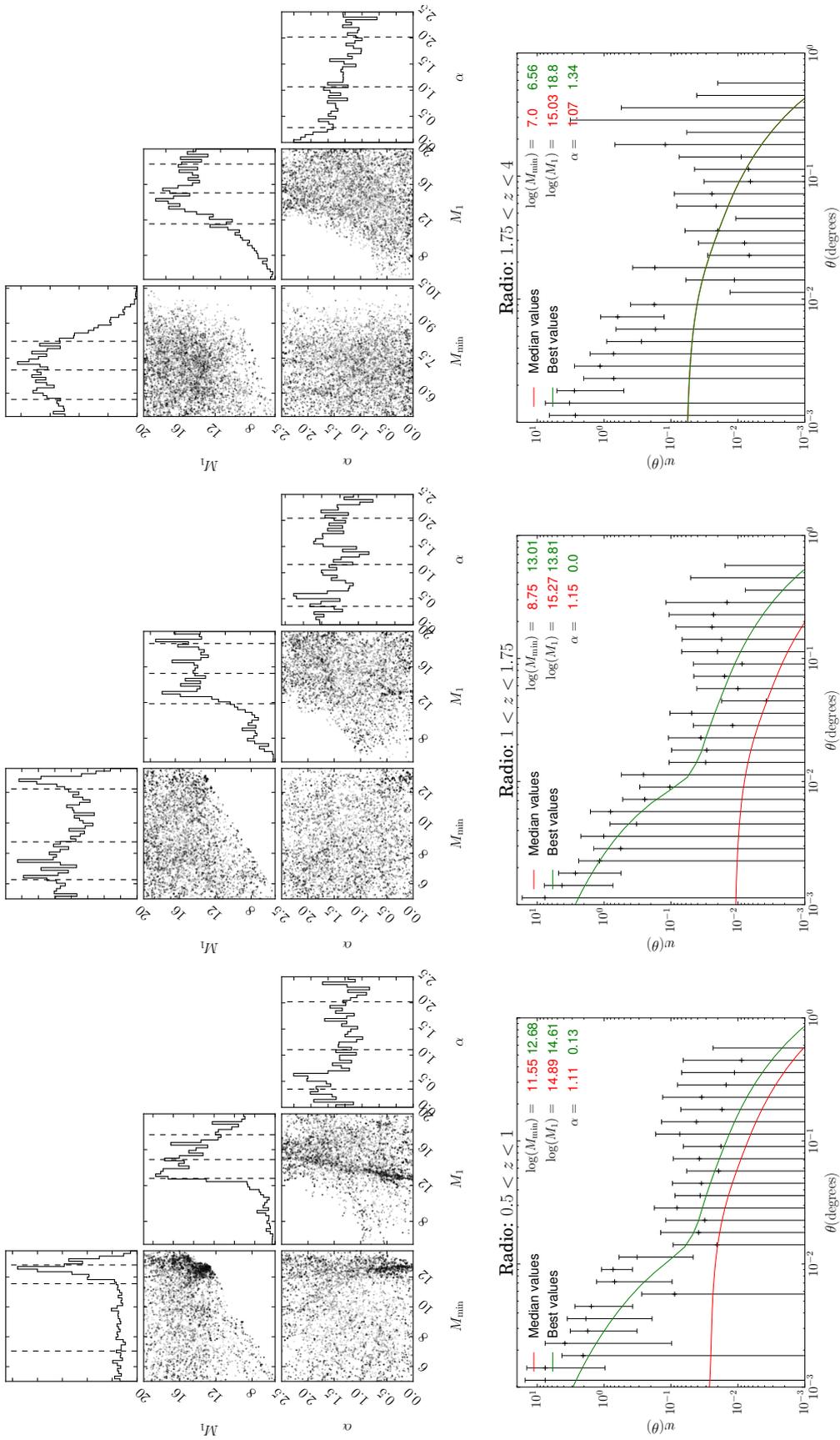
FIGURE 6.8: HOD model fits to the $0.5 < z < 1$ (left), $1 < z < 1.75$ (centre) and $1.75 < z < 4$ (right) angular correlation function of the radio-selected sources (inferred from the cross-correlation function; equation 6.13).
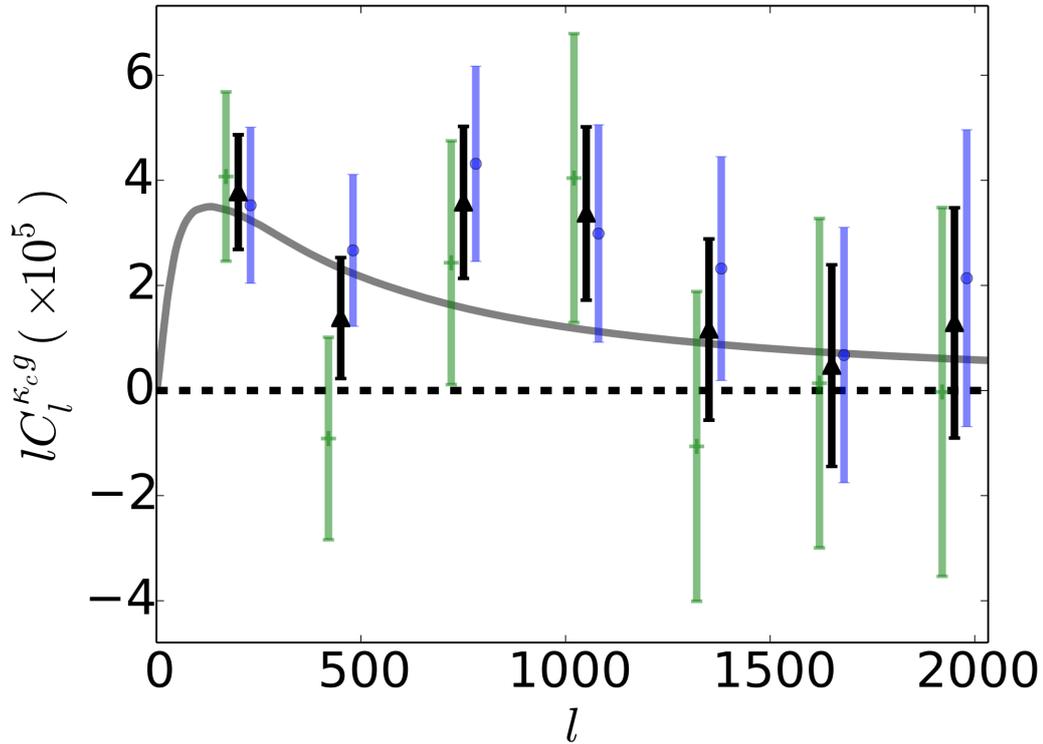
FIGURE 6.9: Galaxy–lensing cross-spectrum $C_\ell^{\kappa g}$ for ACT×FIRST (blue), ACTPol×FIRST (green) and their combination (black). The fiducial cross-spectrum is in grey; note this model is not fitted to the data. Scaling the fiducial bias model by an amplitude b, the combined significance of the cross-spectrum detection is 4.5σ (Allison, Lindsay, Sherwin et al., in prep.).

L-band survey is also being conducted, to a 5σ depth of 250 μJy, with the EVLA over the 270-square-degree SDSS Stripe 82 region ($-50° < \alpha < 59°$, $-1.25° < \delta < 1.25°$). This region has been repeatedly observed in SDSS $ugriz$ bands (reaching $g \sim 24.5$), attracting multi-wavelength surveys including $YJHK$ band near-infrared imaging in UKIDSS, *Herschel* far-infrared imaging from HerMES-HELMS (Oliver et al., 2012), and Atacama Cosmology Telescope (ACT; Fowler et al., 2010) observations at mm wavelengths. Hodge et al. (2011) have already surveyed 92 square degrees of this field to a similar depth before the VLA was upgraded, but the large bandwidth now offered expands this new survey to the full area. This Stripe 82 survey should see some 84,000 detections, with photometric redshifts being readily available to carry out similar work to that presented in this thesis, with the intended extensions and improvements.

6.3.3   Radio-CMB cross-correlation

A major science goal for the EVLA Stripe 82 survey is to measure the bias of radio sources using the cross-power spectrum with CMB lensing maps from the Atacama Cosmology Telescope (ACT; Das et al., 2011). The redshift distributions of radio surveys conveniently peak at $1 < z < 2$, coinciding with the range over which the CMB lensing signal is strongest. Lensing maps directly trace the density field, while radio sources do so indirectly, with radio emission being generated at high redshifts by massive AGN. Similar work has been undertaken in measuring the bias of quasars, by Sherwin et al. (2012) using SDSS-selected quasars with ACT lensing maps, and by Geach et al. (2013) using WISE-selected quasars with lensing maps from the South Pole Telescope. Figure 6.9 shows the galaxy–lensing cross-correlation of FIRST sources with ACT and ACTPol lensing maps over 300 square degrees, detected at 4.6σ significance (Allison, Lindsay, Sherwin et al., in prep.) – an indication of what will be done with this deeper Stripe 82 radio data, and in the longer term with the SKA.

With the latest CMB data from the Planck satellite (Planck Collaboration et al., 2013b), there is also potential progress to be made with existing radio surveys in observing a significant detection of the ISW effect (Section 1.2.5). Existing measurements, across radio and all wavelengths, are summarized by Giannantonio et al. (2012), including those with NVSS. However, cross-correlating an all-sky radio survey, using a combination of NVSS and SUMSS (Section 1.3.4), with the latest Planck CMB maps could still yield interesting results.

References

Afonso J., Mobasher B., Koekemoer A., Norris R. P., Cram L., 2006, AJ , 131, 1216

Alfvén H., Herlofson N., 1950, Physical Review, 78, 616

Alpher R. A., Herman R., 1948, Nature , 162, 774

Antonucci R., 1993, ARA&A , 31, 473

Archibald E. N., Dunlop J. S., Hughes D. H., Rawlings S., Eales S. A., Ivison R. J., 2001, MNRAS , 323, 417

Baade W., Minkowski R., 1954, ApJ , 119, 206

Bahcall N. A., Soneira R. M., 1983, ApJ , 270, 20

Baldry I. K. et al., 2010, MNRAS , 404, 86

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ , 304, 15

Baugh C. M., Efstathiou G., 1993, MNRAS , 265, 145

Becker R. H., White R. L., Helfand D. J., 1995, ApJ , 450, 559

Beckmann V., Shrader C. R., 2012, Active Galactic Nuclei

Bell E. F., 2003, ApJ , 586, 794

Bennett A. S., 1962, MmRAS , 68, 163

Berkhuijsen E. M., 1984, A&A , 140, 431

Best P. N., Heckman T. M., 2012, MNRAS , 421, 1569

Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS , 362, 25

Bielby R. M. et al., 2013, arXiv:1310.2172

Blake C., Collister A., Lahav O., 2008, MNRAS , 385, 1257

Blake C., Ferreira P. G., Borrill J., 2004, MNRAS , 351, 923

Blake C. et al., 2011, MNRAS , 418, 1707

Blake C., Mauch T., Sadler E. M., 2004, MNRAS , 347, 787

Blake C., Wall J., 2002a, MNRAS , 329, L37

Blake C., Wall J., 2002b, MNRAS , 337, 993

Bock D. C.-J., Large M. I., Sadler E. M., 1999, AJ , 117, 1578

Bolton J. G., Stanley G. J., Slee O. B., 1949, Nature , 164, 101

Bondi M. et al., 2003, A&A , 403, 857

Bonfield D. G. et al., 2011, MNRAS , 416, 13

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS , 370, 645

Brand K., Rawlings S., Hill G. J., Tufts J. R., 2005, MNRAS , 357, 1231

Camera S., Santos M. G., Bacon D. J., Jarvis M. J., McAlpine K., Norris R. P., Raccanelli A., Röttgering H., 2012, MNRAS , 427, 2079

Ciliegi P. et al., 2005, A&A , 441, 879

Cole S., Kaiser N., 1989, MNRAS , 237, 1127

Cole S. et al., 2005, MNRAS , 362, 505

Colless M. et al., 2001, MNRAS , 328, 1039

Condon J. J., 1992, ARA&A , 30, 575

Condon J. J., Cotton W. D., Broderick J. J., 2002, AJ , 124, 675

Condon J. J. et al., 2012, ApJ , 758, 23

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ , 115, 1693

Condon J. J., Odell S. L., Puschell J. J., Stein W. A., 1981, ApJ , 246, 624

Cooray A. et al., 2010, A&A , 518, L22

Cooray A., Sheth R., 2002, Phys. Rep. , 372, 1

Coupon J. et al., 2012, A&A , 542, A5

Cress C. M., Helfand D. J., Becker R. H., Gregg M. D., White R. L., 1996, ApJ , 473, 7

Croom S. M. et al., 2005, MNRAS , 356, 415

Croom S. M., Shanks T., Boyle B. J., Smith R. J., Miller L., Loaring N. S., Hoyle F., 2001, MNRAS , 325, 483

Croton D. J. et al., 2006, MNRAS , 365, 11

Curtis H. D., 1917, PASP , 29, 206

Daddi E., Broadhurst T., Zamorani G., Cimatti A., Röttgering H., Renzini A., 2001, A&A , 376, 825

Daddi E. et al., 2002, A&A , 384, L1

Dalton G. et al., 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Das S. et al., 2011, ApJ , 729, 62

Davis M., Peebles P. J. E., 1983, ApJ , 267, 465

de Jong T., Klein U., Wielebinski R., Wunderlich E., 1985, A&A , 147, L6

de Zotti G., Persic M., Franceschini A., Danese L., Palumbo G. G. C., Boldt E. A., Marshall F. E., 1990, ApJ , 351, 22

Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, ApJ , 142, 414

Driver S. P. et al., 2011, MNRAS , 413, 971

Dubois Y., Devriendt J., Slyz A., Teyssier R., 2010, MNRAS , 409, 985

Dunlop J. S., McLure R. J., Kukula M. J., Baum S. A., O'Dea C. P., Hughes D. H., 2003, MNRAS , 340, 1095

Dunlop J. S., Peacock J. A., 1990, MNRAS , 247, 19

Duyvendak J. J. L., 1942, PASP , 54, 91

Eales S. et al., 2010, PASP , 122, 499

Edge D. O., Shakeshaft J. R., McAdam W. B., Baldwin J. E., Archer S., 1959, MmRAS , 68, 37

Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS , 235, 715

Efstathiou G., Moody S. J., 2001, MNRAS , 325, 1603

Eisenstein D. J. et al., 2005, ApJ , 633, 560

El Bouchefry K., Cress C. M., 2007, Astronomische Nachrichten, 328, 577

Elyiv A. et al., 2012, A&A , 537, A131

Fanaroff B. L., Riley J. M., 1974, MNRAS , 167, 31P

Farrah D. et al., 2006, ApJ , 641, L17

Ferramacho L. D., Santos M. G., Jarvis M. J., Camera S., 2014, ArXiv e-prints

Fine S., Shanks T., Nikoloudakis N., Sawangwit U., 2011, MNRAS , 418, 2251

Foucaud S. et al., 2007, MNRAS , 376, L20

Fowler J. W. et al., 2010, ApJ , 722, 1148

Francis P. J., Hewett P. C., Foltz C. B., Chaffee F. H., Weymann R. J., Morris S. L., 1991, ApJ , 373, 465

Furusawa J., Sekiguchi K., Takata T., Furusawa H., Shimasaku K., Simpson C., Akiyama M., 2011, ApJ , 727, 111

Gamow G., 1948, Nature , 162, 680

Garrett M. A., 2002, A&A , 384, L19

Geach J. E. et al., 2013, ApJ , 776, L41

Geach J. E., Sobral D., Hickox R. C., Wake D. A., Smail I., Best P. N., Baugh C. M., Stott J. P., 2012, MNRAS , 426, 679

Giannantonio T., Crittenden R., Nichol R., Ross A. J., 2012, MNRAS , 426, 2581

Giannantonio T., Scranton R., Crittenden R. G., Nichol R. C., Boughn S. P., Myers A. D., Richards G. T., 2008, Phys. Rev. D , 77, 123520

Gonzalez-Solares E. A. et al., 2005, MNRAS , 358, 333

Grazian A. et al., 2006, A&A , 449, 951

Groth E. J., Peebles P. J. E., 1977, ApJ , 217, 385

Guo Q. et al., 2011, MNRAS , 412, 2277

Guzzo L., Strauss M. A., Fisher K. B., Giovanelli R., Haynes M. P., 1997, ApJ , 489, 37

Hamilton A. J. S., 1993, ApJ , 417, 19

Hardcastle M. J., Evans D. A., Croston J. H., 2007, MNRAS , 376, 1849

Hardcastle M. J. et al., 2010, MNRAS , 409, 122

Hartley W. G. et al., 2013, MNRAS , 431, 3045

Hatch N. A. et al., 2014, MNRAS , 445, 280

Hawkins E. et al., 2003, MNRAS , 346, 78

Helou G., Soifer B. T., Rowan-Robinson M., 1985, ApJ , 298, L7

Herbert P. D., Jarvis M. J., Willott C. J., McLure R. J., Mitchell E., Rawlings S., Hill G. J., Dunlop J. S., 2011, MNRAS , 410, 1360

Hey J. S., Parsons S. J., Phillips J. W., 1946, Nature , 158, 234

Hey J. S., Phillips J. W., Parsons S. J., 1946, Nature , 157, 296

Hildebrandt H., van Waerbeke L., Erben T., 2009, A&A , 507, 683

Hill D. T. et al., 2011, MNRAS , 412, 765

Hine R. G., Longair M. S., 1979, MNRAS , 188, 111

Hodge J. A., Becker R. H., White R. L., Richards G. T., Zeimann G. R., 2011, AJ , 142, 3

Hopkins P. F., Cox T. J., Hernquist L., Narayanan D., Hayward C. C., Murray N., 2013, MNRAS , 430, 1901

Hoyle F., 1948, MNRAS , 108, 372

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hubble E. P., 1925, Popular Astronomy, 33, 252

Huchra J. P., Geller M. J., 1982, ApJ , 257, 423

Huterer D., Knox L., Nichol R. C., 2001, ApJ , 555, 547

Ichikawa T. et al., 2007, PASJ , 59, 1081

Ilbert O. et al., 2006, A&A , 457, 841

Ivashchenko G., Zhdanov V. I., Tugay A. V., 2010, MNRAS , 409, 1691

Jansky K. G., 1933a, Popular Astronomy, 41, 548

Jansky K. G., 1933b, Nature , 132, 66

Jansky K. G., 1935, Proc. IRE, 21, 1387

Jarvis M. J., 2012, African Skies, 16, 44

Jarvis M. J. et al., 2013, MNRAS , 428, 1281

Jarvis M. J., Rawlings S., 2004, NewAR, 48, 1173

Jarvis M. J., Rawlings S., Eales S., Blundell K. M., Bunker A. J., Croft S., McLure R. J., Willott C. J., 2001, MNRAS , 326, 1585

Jarvis M. J. et al., 2010, MNRAS , 409, 92

Jennison R. C., Gupta M. K. D., 1956, Philosophical Magazine, 1, 65

Johnston S. et al., 2007, PASA, 24, 174

Kaiser N., 1984, ApJ , 284, L9

Kaiser N., 1987, MNRAS , 227, 1

Kerscher M., Szapudi I., Szalay A. S., 2000, ApJ , 535, L13

Khachikian E. Y., Weedman D. W., 1974, ApJ , 192, 581

Kiepenheuer K. O., 1950, Physical Review, 79, 738

Kim J.-W. et al., 2014, MNRAS , 438, 825

Kim J.-W., Edge A. C., Wake D. A., Stott J. P., 2011, MNRAS , 410, 241

Komatsu E. et al., 2011, ApJS , 192, 18

Kooiman B. L., Burns J. O., Klypin A. A., 1995, ApJ , 448, 500

Kovač K., Somerville R. S., Rhoads J. E., Malhotra S., Wang J., 2007, ApJ , 668, 15

Lacki B. C., Thompson T. A., 2010, ApJ , 717, 196

Lacki B. C., Thompson T. A., Quataert E., 2010, ApJ , 717, 1

Landy S. D., Szalay A. S., 1993, ApJ , 412, 64

Lawrence A. et al., 2007, MNRAS , 379, 1599

Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1

Lemaître G., 1927, Annales de la Societe Scietifique de Bruxelles, 47, 49

Lewis A., Challinor A., 2011, CAMB: Code for Anisotropies in the Microwave Background. Astrophysics Source Code Library

Limber D. N., 1953, ApJ , 117, 134

Lindsay S. N., Jarvis M. J., McAlpine K., 2014, MNRAS , 440, 2322

Lindsay S. N. et al., 2014, MNRAS , 440, 1527

Ling E. N., Barrow J. D., Frenk C. S., 1986, MNRAS , 223, 21P

Lisenfeld U., Voelk H. J., Xu C., 1996, A&A , 314, 745

Liske J., 2000, MNRAS , 319, 557

Loan A. J., Wall J. V., Lahav O., 1997, MNRAS , 286, 994

Lynden-Bell D., 1969, Nature , 223, 690

Lynden-Bell D., 1978, Phys. Scr , 17, 185

Magliocchetti M., Maddox S. J., 2002, MNRAS , 330, 241

Magliocchetti M. et al., 2004, MNRAS , 350, 1485

Magliocchetti M., Maddox S. J., Lahav O., Wall J. V., 1999, MNRAS , 306, 943

Mandelbaum R., Li C., Kauffmann G., White S. D. M., 2009, MNRAS , 393, 377

Masson C., 1979, MNRAS , 188, 261

Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS , 286, 115

Matthews T. A., Sandage A. R., 1963, ApJ , 138, 30

Mauch T., Sadler E. M., 2007, MNRAS , 375, 931

McAlpine K., Jarvis M. J., 2011, MNRAS , 413, 1054

McAlpine K., Jarvis M. J., Bonfield D. G., 2013, MNRAS , 436, 1084

McAlpine K., Smith D. J. B., Jarvis M. J., Bonfield D. G., Fleuren S., 2012, MNRAS , 423, 132

McCarthy P. J. et al., 2001, ApJ , 560, L131

McCracken H. J. et al., 2010, ApJ , 708, 202

McCracken H. J. et al., 2012, A&A , 544, A156

McEwen J. D., Vielva P., Hobson M. P., Martínez-González E., Lasenby A. N., 2007, MNRAS , 376, 1211

McLure R. J., Jarvis M. J., 2004, MNRAS , 353, L45

McLure R. J., Willott C. J., Jarvis M. J., Rawlings S., Hill G. J., Mitchell E., Dunlop J. S., Wold M., 2004, MNRAS , 351, 347

Mo H. J., White S. D. M., 1996, MNRAS , 282, 347

Mocz P., Fabian A. C., Blundell K. M., 2013, MNRAS , 432, 3381

Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23

Muzzin A. et al., 2013, ApJS , 206, 8

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ , 490, 493

Nilsson K. K., Meisenheimer K., 2009, NewAR, 53, 37

Norberg P. et al., 2002, MNRAS , 332, 827

Norris R. P. et al., 2013, PASA, 30, 20

Norris R. P. et al., 2011, PASA, 28, 215

Oke J. B., 1990, AJ , 99, 1621

Oliver S. J. et al., 2012, MNRAS , 424, 1614

Orr M. J. L., Browne I. W. A., 1982, MNRAS , 200, 1067

Overzier R. A., Röttgering H. J. A., Rengelink R. B., Wilman R. J., 2003, A&A , 405, 53

Owsianik I., Conway J. E., 1998, A&A , 337, 69

Passmoor S., Cress C., Faltenbacher A., Johnston R., Smith M., Ratsimbazafy A., Hoyle B., 2013, MNRAS , 429, 2183

Peacock J. A., Nicholson D., 1991, MNRAS , 253, 307

Peebles P. J. E., 1973, ApJ , 185, 413

Peebles P. J. E., 1980, The large-scale structure of the universe. Princeton University Press

Penzias A. A., Wilson R. W., 1965, ApJ , 142, 419

Percival W. J. et al., 2007, ApJ , 657, 645

Percival W. J. et al., 2010, MNRAS , 401, 2148

Perlmutter S. et al., 1999, ApJ , 517, 565

Planck Collaboration et al., 2013a, ArXiv e-prints

Planck Collaboration et al., 2013b, ArXiv e-prints

Prandoni I., Gregorini L., Parma P., de Ruiter H. R., Vettolani G., Zanichelli A., Wieringa M. H., Ekers R. D., 2001, A&A , 369, 787

Press W. H., Schechter P., 1974, ApJ , 187, 425

Priddey R. S., Isaak K. G., McMahon R. G., Robson E. I., Pearson C. P., 2003, MNRAS , 344, L74

Quadri R. et al., 2007, ApJ , 654, 138

Raccanelli A., Bonaldi A., Negrello M., Matarrese S., Tormen G., de Zotti G., 2008, MNRAS , 386, 2161

Raccanelli A. et al., 2012, MNRAS , 424, 801

Rawlings S., Willott C. J., Hill G. J., Archibald E. N., Dunlop J. S., Hughes D. H., 2004, MNRAS , 351, 676

Reber G., 1940, ApJ , 91, 621

Reber G., 1944, ApJ , 100, 279

Refregier A., 2003, ARA&A , 41, 645

Rengelink R. B., Tang Y., de Bruyn A. G., Miley G. K., Bremer M. N., Roettgering H. J. A., Bremer M. A. R., 1997, A&AS , 124, 259

Riess A. G. et al., 1998, AJ , 116, 1009

Roche N., Eales S. A., 1999, MNRAS , 307, 703

Roche N. D., Almaini O., Dunlop J., Ivison R. J., Willott C. J., 2002, MNRAS , 337, 1282

Ross A. J., Percival W. J., Brunner R. J., 2010, MNRAS , 407, 420

Ross N. P. et al., 2007, MNRAS , 381, 573

Ross N. P. et al., 2009, ApJ , 697, 1634

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ , 238, 471

Ryle M., 1950, Reports on Progress in Physics, 13, 184

Sachs R. K., Wolfe A. M., 1967, ApJ , 147, 73

Salpeter E. E., 1964, ApJ , 140, 796

Saunders W., Rowan-Robinson M., Lawrence A., 1992, MNRAS , 258, 134

Sawangwit U., Shanks T., Abdalla F. B., Cannon R. D., Croom S. M., Edge A. C., Ross N. P., Wake D. A., 2011, MNRAS , 416, 3033

Scheuer P. A. G., Readhead A. C. S., 1979, Nature , 277, 182

Schinnerer E. et al., 2010, ApJS , 188, 384

Schmidt M., 1963, Nature , 197, 1040

Scranton R. et al., 2005, ApJ , 633, 589

Seljak U., Warren M. S., 2004, MNRAS , 355, 129

Serjeant S. et al., 2010, A&A , 518, L7

Seyfert C. K., 1943, ApJ , 97, 28

Seymour N. et al., 2011, MNRAS , 413, 1777

Sherwin B. D. et al., 2012, Phys. Rev. D , 86, 083006

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS , 323, 1

Shklovsky I. S., 1949, Astron. Zh., 25, 237

Simon P., 2007, A&A , 473, 711

Simpson C. et al., 2012, MNRAS , 421, 3060

Skrutskie M. F. et al., 2006, AJ , 131, 1163

Slipher V. M., 1914, Popular Astronomy, 22, 19

Smith D. J. B. et al., 2011, MNRAS , 416, 857

Smith R. E. et al., 2003, MNRAS , 341, 1311

Smolčić V. et al., 2009, ApJ , 696, 24

Sramek R. A., Weedman D. W., 1980, ApJ , 238, 435

Sullivan M., Hopkins A. M., Afonso J., Georgakakis A., Chan B., Cram L. E., Mobasher B., Almeida C., 2004, ApJS , 155, 1

Sutherland W., Saunders W., 1992, MNRAS , 259, 413

Swinbank A. M., Chapman S. C., Smail I., Lindner C., Borys C., Blain A. W., Ivison R. J., Lewis G. F., 2006, MNRAS , 371, 465

Takeuchi T. T., Yoshikawa K., Ishii T. T., 2003, ApJ , 587, L89

Taylor P., Kobayashi C., 2015, ArXiv e-prints

Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I., 2005, ApJ , 631, 41

Townes C. H., 1947, ApJ , 105, 235

Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ , 598, 886

Urry C. M., Padovani P., 1995, PASP , 107, 803

van Albada T. S., Bahcall J. N., Begeman K., Sancisi R., 1985, ApJ , 295, 305

van de Hulst H., 1945, Ned.Tijd.Natuurkunde, 11, 210

van Haarlem M. P. et al., 2013, A&A , 556, A2

Van Waerbeke L., Mellier Y., 2003, ArXiv Astrophysics e-prints

Voelk H. J., 1989, A&A , 218, 67

Vogelsberger M. et al., 2014, MNRAS , 444, 1518

Wake D. A., Croom S. M., Sadler E. M., Johnston H. M., 2008, MNRAS , 391, 1674

Wang L. et al., 2011, MNRAS , 414, 596

Webster A., 1976, MNRAS , 175, 61

Whitaker K. E. et al., 2011, ApJ , 735, 86

White R. L., Becker R. H., Helfand D. J., Gregg M. D., 1997, ApJ , 475, 479

White S. D. M., Davis M., Efstathiou G., Frenk C. S., 1987, Nature , 330, 451

White S. V., Jarvis M. J., Häußler B., Maddox N., 2014, ArXiv e-prints

Willmer C. N. A., da Costa L. N., Pellegrini P. S., 1998, AJ , 115, 869

Willott C. J., McLure R. J., Jarvis M. J., Rawlings S., Hill G. J., Mitchell E., Dunlop J. S., 2003, NewAR, 47, 187

Willott C. J., Rawlings S., Archibald E. N., Dunlop J. S., 2002, MNRAS , 331, 435

Willott C. J., Rawlings S., Blundell K. M., 2001, MNRAS , 324, 1

Wilman R. J., Jarvis M. J., Mauch T., Rawlings S., Hickey S., 2010, MNRAS , 405, 447

Wilman R. J. et al., 2008, MNRAS , 388, 1335

Wilman R. J., Röttgering H. J. A., Overzier R. A., Jarvis M. J., 2003, MNRAS , 339, 695

Yang A., Saslaw W. C., 2011, ApJ , 729, 123

York D. G. et al., 2000, AJ , 120, 1579

Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ , 554, 803

Zehavi I. et al., 2005, ApJ , 630, 1

Zel'dovich Y. B., Novikov I. D., 1965, Soviet Physics Doklady, 9, 834

Zheng Z. et al., 2005, ApJ , 633, 791