UNIVERSITY *of the*
WESTERN CAPE

A Large-Scale Clustering Analysis of Radio Galaxies

by

Nikhita Madhanpall

October 2019

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

Doctor of Philosophy

in the

Department of Physics and Astronomy

UNIVERSITY OF THE WESTERN CAPE

Supervisor: Prof. Matt Jarvis

The financial assistance of the South African Radio Astronomy Observatory towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Abstract

The distribution of the visible structure in the Universe is in part dependent on the influence of the dark matter structure. Measurements of large-scale structure can therefore be related to the dark matter distribution by calculating the galaxy bias b(z), which indicates the extent to which the baryonic matter traces out the underlying dark matter distribution. Radio sources are easily detected out to higher redshifts and are thus an effective probe of the large-scale structure of the Universe, however very few clustering studies have been carried out at sub-mJy levels, as well as for various radio source types. In this Thesis, I measure the angular two-point correlation function of radio populations using 11,431 sources from a Jansky Very Large Array (JVLA) CnB $S_{1.4} > 440$ μJy survey over $\sim 100\text{deg}^2$ of the Sloan Digital Sky Survey (SDSS) Southern Equatorial Stripe (Stripe 82), having a resolution of 10×16 arcseconds. Sources are also matched with deep coadded data from the SDSS in order to obtain redshift information. The $S_{1.4} > 440$ μJy flux density limit allows for a sample which is active galactic nuclei (AGN) dominated and contains a number of higher-redshift AGN, the clustering of which is not well studied. The correlation length (r_0) and bias (b) values are measured for radio samples which have been split by redshift, radio luminosity, as well as radio source type. These clustering properties are also used to estimate the typical dark matter halo environments of the radio sub-populations. The measured clustering properties for the various sub-samples range from $r_0 = 3.81^{+1.25}_{-2.10} h^{-1}$Mpc and $b = 0.91^{+0.70}_{-0.50}$ for a sample of star-forming galaxies (SFGs) at $z < 1$, showing that SGFs are likely a low mass population which occupy less massive haloes of $M_{\text{halo}} \sim 10^{11} - 3 \times 10^{12} h^{-1} \mathcal{M}_{\odot}$, to $r_0 = 11.93^{+1.04}_{-2.96} h^{-1}$Mpc and $b = 2.50^{+0.83}_{-0.81}$ for a sample of high-redshift ($0.45 \leq z \leq 0.7$) low excitation radio galaxy (LERG) sources. The latter are found to occupy higher mass haloes of $M_{\text{halo}} \sim 1 - 5 \times 10^{13} h^{-1} \mathcal{M}_{\odot}$, providing evidence for the link between accretion mode and halo mass, as inefficient accreters are thought to reside in highly clustered environments where hot gas is accreted slowly, unlike SFGs and high excitation radio galaxies (HERGs) which reside in lower mass haloes which have a larger supply of cold gas. Relatively high clustering values of $r_0 = 8.02^{+1.73}_{-1.88}$ h^{-1}Mpc and $b(z) = 1.62^{+0.45}_{-0.38}$ are measured for a $L_{1.4\text{GHz}} < 10^{23}$ WHz^{-1} sample with a median redshift of $z_{\text{med}} = 0.25$, suggesting the presence of a significant

number of low luminosity AGN in the sample, which is lacking in the S^3 simulation. The bias evolution with redshift measured for the sample is found to closely follow the S^3 bias model for FRI galaxies.

Acknowledgements

I'd like to offer deep thanks and gratitude to my supervisor Matt Jarvis, for the years of academic and personal guidance. Thank you for being patient, available and supportive throughout this bicontinental, unconventional journey and for never losing confidence in my ability to reach the finish line, despite challenges put before me.

Thank you to Sam for your assistance during the initial stages and for providing clarity whenever needed. Thanks to Matt Prescott for useful discussions and to Imogen for being ever willing to help. Thank you to Prof. Roy Maartens, Prof. Mario Santos, Marta, Kim and Stephen for the support and guidance during my visits to UWC. Thank you to Tom for the many uplifting chats and advice over coffee during my visits to Cape Town. I'd like to thank SARAO for the PhD bursary, as well as travel and equipment grants, and Oxford University for the hospitality during my visits.

Thank you to Rajin and Iniyan for always lending an ear and giving honest, much needed advice. Thank you to Susan for the check-ins and pep talks when struggles with juggling the PhD and a full-time job occurred. Thank you Sarah for the wonderful company and momos during my Oxford visits. Much love and thanks to Karin for your invaluable company during our overlapping stays at Exeter House, for imparting so much knowledge and for your sincere love and kindness.

Big hugs and huge thanks to Kiki and Pearl for your fierce support during my years at the SARAO Rosebank office and giving me the strength to persevere. Thank you to my manager Carel for being supportive of my PhD and for allowing me time to work on it whenever possible. Thank you to my RFI teammates for your continued support.

Thank you to my sister Kesh for your kind and encouraging words throughout this journey. Thank you to my Nani and Nana for your kind words and prayers over the years. Thank you to my Aji-ma for your wisdom and long, enlightening phone calls. Thank you to my parents for their solid faith and prayers. Thank you to my Dad for supporting my astronomy dream since my teen years, for my telescope and the many lovely books. Thank you to my Mum for being my greatest supporter, for your unwavering belief in me and for being there through every high and every low. Finally, thank you to my husband Ari for providing the support, time, patience, counselling, humour and many cups of tea necessary to achieve this goal.

Declaration of Authorship

I declare that *A Large-Scale Clustering Analysis of Radio Galaxies* is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Full name: Nikhita Madhanpall Date: 25/10/2019

Signed.....

Contents

List of Figures

List of Tables

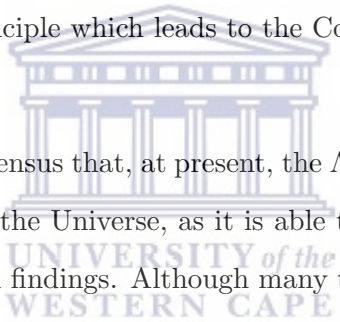
Chapter 1

Introduction

Approaching an era of next-generation telescopes and big data, there has never been a time when one should be more driven and optimistic about the advancement of cosmology than the present. Only in the last two decades have we developed the concordance model of cosmology, which consists of several ideas that together, are best able to describe the Universe as we understand it thus far. This model, termed the Lambda Cold Dark Matter (ΛCDM) Model, is the simplest model that is able to best explain the majority of observed phenomena in the Universe. However, the two main constituents of this model, dark energy and dark matter, are still poorly understood, as they cannot be studied via direct observation. It is evident that we have a great task at hand if we are to obtain a full understanding of the physics of this "dark" Universe. Having said this, the present generation has seen rapid improvement in observational technology and techniques, as well as computational advancement. Cosmology has thus become exceedingly data driven and theoretical models are tested using a huge amount of essential data.

1.1 ΛCDM Cosmology

In order to arrive at this model which is used to explain what we observe in the Universe, we must first assume that the Universe is uniform on sufficiently large scales. That is, that the Universe has no preferred direction, i.e. is isotropic and no preferred location, i.e. is homogeneous. This is known as the Cosmological Principle. The Universe is clearly not exactly homogeneous, thus cosmologists define homogeneity in an average sense, that is, the Universe is taken to be identical at different locations when one looks at sufficiently large pieces. Isotropy, however, does not imply homogeneity without the additional assumption that the observer is not in a special place. The Copernican Principle asserts that we are in no preferred location in the Universe, this is a built-in assumption of the favored solutions to Einstein's field equations, the Friedmann-Robertson-Walker space-times. It is therefore the observed isotropy together with the Copernican Principle which leads to the Cosmological Principle.



There is a mostly global consensus that, at present, the ΛCDM Model of cosmology provides the best description of the Universe, as it is able to explain and incorporate a number of major observational findings. Although many theories in cosmology have been developed, the ΛCDM model is still a benchmark against which other models or theories are compared. The key observational successes of this model are in fields such as the cosmic microwave background (CMB), the expansion of the Universe, the formation of galaxies and large scale structure and also the abundance of light elements in the early Universe. In each of these areas of cosmology, the ΛCDM Model is able to make accurate and testable hypotheses, which are found to be in remarkable agreement with observations. The goal of modern cosmology is to constrain various cosmological parameters by the analysis of large amounts data, from surveys carried out at different wavelengths, each probing different source types and different epochs in the Universe. The ΛCDM model is based on the observational constraints of various parameters which describe the Universe, such as the baryon density $\Omega_b h^2$, dark matter density $\Omega_c h^2$, dark energy density Ω_Λ, curvature fluctuation amplitude Δ_R^2 and age of the Universe t_0. In the aforementioned parameters, h is the dimensionless equivalent

of the Hubble constant, i.e. when the Hubble constant is expressed as $H_0 = 100h$ kms^{-1} Mpc^{-1}. Note, for example, that the matter density of the Universe Ω_m refers to the fraction of the total mass-energy density of the Universe attributed to matter. The physical baryon density and physical dark matter density give the fraction of the critical density made of baryons and dark matter, respectively, where the critical density is the total mass-energy density required for a flat Universe. Other parameters of great importance are the Hubble constant H_0, which gives the expansion rate of the Universe and σ_8 which measures the amplitude of the linear power spectrum on the scale of $8h^{-1}$ Mpc. Some of the key observations and experiments which have been used to successfully constrain the cosmological parameters relevant to large-scale clustering are described below.

1.1.1 Cosmic Expansion

The Hubble Constant

A foundation of the ΛCDM model is the observed expansion of the Universe, which was predicted as early as 1927 from General Relativity equations derived by Georges Lemaître. In 1929, Edwin Hubble found, after measuring velocities of extragalactic objects, that these objects were moving away from us, and that the further away a galaxy, the faster its recessional velocity (Hubble 1929). This movement of galaxies is termed the 'Hubble flow'. In order for isotropy and homogeneity to hold for Hubble's law, the Universe cannot be a fixed volume. Thus, space itself must be expanding, which is the basis for the Big Bang model. The rate at which space is expanding is termed the Hubble constant, H_0, known as the Hubble parameter at present day. Hubble's law is therefore,

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

where v, the recession velocity, is linearly proportional to the distance, d, to the galaxy. Thus, by inverting Hubble's law we are able to estimate the age of the Universe to be ~ 13.8 billion years (Planck Collaboration et al. 2018), assuming a constant rate

of expansion.

In the past, measurements of the Hubble constant yielded values $\sim$ 70 kms^{-1} Mpc^{-1}. For example, the Hubble Space Telescope (HST) Key Project measured the distance to Cepheid variable stars, which are stars that pulse at regular intervals corresponding to their peak brightness, and found a value of $H_0 = 72 \pm 8$ kms^{-1} Mpc^{-1} (Freedman et al. 2001). However, in more recent studies, by measuring small fluctuations in the temperature of the Cosmic Microwave Background (to be further discussed in this Section) using the Planck spacecraft, the Hubble constant was found to be $H_0 = 67.4 \pm 0.5$ kms^{-1} Mpc^{-1} (Planck Collaboration et al. 2018) in disagreement with previous measurements of $\sim$ 70 kms^{-1} Mpc^{-1}. To further investigate this discrepancy, Riess et al. (2019) calculated the Hubble constant by once again measuring the distance to Cepheid stars, but refined to very high precision, and find a value of $H_0 = 74.03 \pm 1.42$ kms^{-1} Mpc^{-1}. The most recent study by Freedman et al. (2019) made use of red giant stars as distance measures, i.e using the Tip of the Red Giant Branch (TRGB) method. Red giants undergo an event called a Helium flash, in which the temperature of the star rises drastically and the structure of the star is altered, resulting in a very significant decrease in the star's luminosity. Thus, the apparent brightness of red giants at this life-stage, in various galaxies, can provide distance measures. Using this method, a value of $H_0 = 69.8 \pm 0.8$ kms^{-1} Mpc^{-1} is found, which is in between those found by Planck Collaboration et al. (2018) and Riess et al. (2019). Tests by Riess et al. (2019) suggest that the discrepancy in H_0 is not due to the sources used or to an error in measurements, but rather an aspect of cosmology beyond the ΛCDM model. However, Freedman et al. (2019) conclude that more work towards reducing systematic errors in the local distance scale must be done before a revision on the standard model is necessary and motivate the need for independent local measurements to test the standard model of cosmology and its extrapolation to the present day. This issue is at present one of the most crucial in cosmology.

Dark Energy

Advances in the field of cosmic expansion were made in 1998, when the Hubble Space Telescope (HST) allowed for observations of distant supernovae. Type Ia supernovae result from the explosion of white dwarf stars which have a primary composition of carbon and oxygen, as opposed to white dwarfs composed mostly of magnesium, neon and oxygen, which usually collapse to form neutron stars. A slowly rotating carbon-oxygen white dwarf, which is accreting mass from a binary star, is limited by a mass known as the Chandrasekhar limit, about 1.38 $\mathcal{M}_{\odot}$. At this limit, electrons are pressed as close together as possible (electron degeneracy pressure). While accreting mass, the core of the white dwarf reaches the ignition temperature for carbon fusion. Once this nuclear fusion is initiated the star undergoes a reaction releasing immense amounts of energy (brighter than billions of stars) in just a few seconds, resulting in a supernova explosion. Due to the uniformity in the mass of exploding white dwarfs, these supernovae all have similar peak luminosities and are all very bright. This allows for the use of these explosions as standard candles across cosmological distances, which in turn allows for accurate distance measurements. The absolute magnitude, which is the intrinsic brightness of type Ia supernovae, is known, which allows for distance measures using the apparent magnitude, or, observed brightness.

The goal of early type Ia supernova projects was therefore to determine the distance of galaxies as well as their redshifts, that is, measure the Hubble relationship. The findings of the *High-z Supernova Search Team* in 1998 showed startling evidence that, in the past, the Universe was expanding more slowly than it is today (Riess et al. 1998). This was, at the time, a very unexpected finding due to the fact that gravity pulls matter together, which should be causing the expansion to slow down. However the results of Riess et al. (1998) would later be supported by further observations of type Ia supernovae. It was found that the luminosity of the observed supernovae were about 25% fainter than expected, suggesting that they were further away than they should be (Tonry et al. 2003; Wang et al. 2003b). Thus, based on the analysis of 42 type Ia supernovae at $0.18 < z < 0.83$, the work by members of the Supernova Cosmology Project indicated that the Universe is expanding at an increasing rate. Us-

ing additional measurements from the Calan/Tololo Supernova Survey (Hamuy et al. 1996) of supernovae at z $<$ 0.1, they were able to constrain the values of Ω_m and Ω_Λ, results of which are shown in Figure 1.1 (Perlmutter et al. 1999). This work was awarded the 2011 Nobel Prize in Physics.

The observation that the expansion of the Universe is accelerating is one of the greatest mysteries in modern cosmology. The conclusion is that there must be some dominant factor of the Universe, with a negative pressure and which is acting strongly against the pull of gravity, causing the fabric of space-time to accelerate in its expansion. This phenomenon which we know very little about is termed dark energy and in the past decade, has become one of the most important issues in cosmology. For a Universe comprising of both dark matter and dark energy, we would expect that at some stage, when the Universe was smaller and more dense, that the gravitational pull of dark matter would be the dominant force, resulting in a decelerating expansion of the Universe. At a later stage dark energy would then become dominant and the expansion would accelerate. The work of Riess et al. (2004) showed that this jerk in the rate of expansion of the Universe is exactly what is observed and is believed to have occurred about 7.5 billion years ago. The nature of dark energy is still unknown, but various possibilities have been proposed. It could be a property of space itself, that is, the energy density of a vacuum, known as the cosmological constant (Λ - Einstein 1917), or a new type of dynamical energy fluid that fills all of space (quintessence - Caldwell et al. 1998; Kant Goswami et al. 2019), or may be explained by an alternate theory of gravity on cosmic scales (Milgrom 1983). The way forward is through larger and better surveys. One recent survey aimed at uncovering the nature of dark energy is the Dark Energy Survey (DES - The Dark Energy Survey Collaboration 2005), which operated over 2013-2019. Using the measurements of 207 spectroscopically confirmed type Ia supernovae they constrain the baryon density to $\Omega_b \sim 0.069$ (Abbott et al. 2019). Subsequently, by combining measurements of these supernovae with 122 supernovae from literature they measure a value for the matter density of $\Omega_m = 0.331 \pm 0.038$ (Abbott et al. 2019).

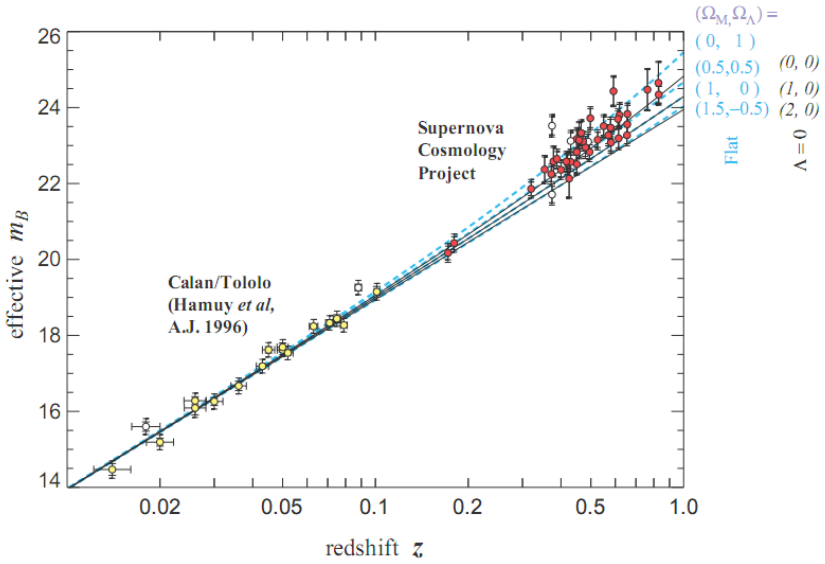


Figure 1.1: Evidence for the accelerated expansion of the Universe from type Ia supernovae data, obtained by members of the Supernova Cosmology Project. This image is taken from the work of Perlmutter et al. (1999) and shows the apparent magnitude of the supernovae with redshift. It is shown that the Universe is clearly not matter dominated.

1.1.2 Dark Matter

Prior to the late 1970's it was believed that the Universe consisted of only normal, baryonic matter: protons, neutrons and electrons, as well as radiation. Observations have since suggested that not only does matter account for only 26% of the mass-energy density of the Universe, but that 84% of the matter content is thought to comprise of an electromagnetically non-interacting material known as dark matter, which only interacts with other particles via gravitational forces and possibly the weak force if dark matter is composed of weakly interacting massive particles (WIMPs), such as a type of neutrino (Jungman et al. 1996). The reason for the term cold dark matter arises from the need for a matter whose primordial velocities were low enough to maintain the very small-scale fluctuations in matter density which are required to seed galaxy formation. The first suggestion that this mysterious, unseen matter was thought to exist arose after measurements made by Dutch astronomer Jan Oort, who when studying the motion of stars in the local galactic neighbourhood, deduced that in order to explain the stellar motions there must be more mass in the galactic plane than that derived from luminous material alone. Just a year later in 1933, Swiss astrophysicist Fritz Zwicky discovered further evidence for the existence of this dark matter while

observing the Coma cluster of galaxies (Zwicky 1933). Zwicky studied the motions of galaxies near the cluster's edge and found that their orbits were very fast. He then deduced that the gravitational force caused by the mass of the visible matter alone would not be large enough to account for such high speed orbits and that the actual mass of the cluster must be about 400 times larger than previously estimated. This became known as the "missing mass problem", which led Zwicky to believe that there was some form of non-visible matter in the cluster.

The most striking evidence for the existence of dark matter arose after observing the motion of stars in nearby galaxies, in orbit around their galactic center. The resulting galaxy rotation curves, which illustrate the rotational velocity of stars as a function of their distance from the center of the galaxy, drew surprising conclusions. It was found that stars located far from the centres of galaxies have much higher velocities than previously expected. It was understood that the rotational velocities of stars in a galaxy should be highest near the galactic center, due to the large mass concentration, and that the velocities should decrease with increasing distance from the galactic center. However, in 1975 Vera Rubin found that most of the stars observed in spiral galaxies orbit their galactic center at roughly the same speed, implying that the mass densities in the galaxy were more or less uniform, even at a great distance from the galactic bulge. This was confirmed by observations of the rotational properties of spiral galaxies (Rubin et al. 1980). In Figure 1.2 a result from this influential paper is shown, which presents the superposition of the galaxy rotation curves of 21 spiral galaxies. It is shown that the velocities remain relatively constant with increasing distance from the galactic nucleus, instead of falling as previously expected (Rubin et al. 1980). The results suggested that most galaxies must contain about 10 times as much dark mass as that accounted for by visible matter.

Strong evidence for the existence of dark matter also comes from the gravitational lensing studies of galaxy clusters such as the Bullet cluster and Abell 1689 (Clowe et al. 2004; Markevitch et al. 2004). The Bullet cluster image, shown in Figure 1.3, is actually the result of two colliding galaxy clusters. The baryonic matter, i.e. the

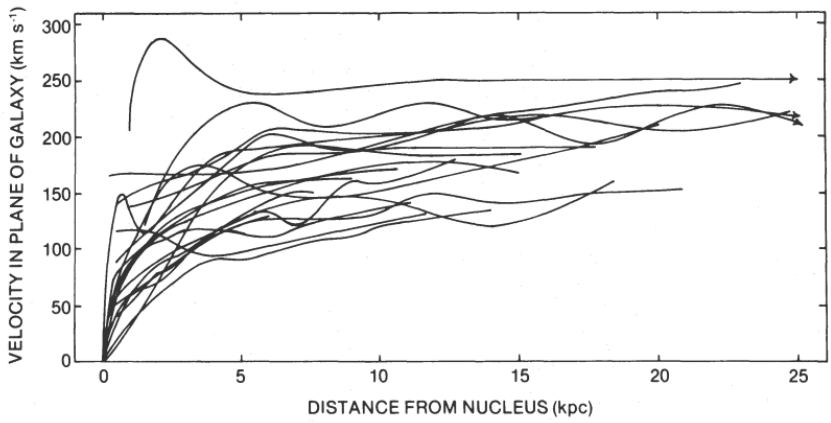


Figure 1.2: Superposition of the galaxy rotation curves of 21 spiral Sc galaxies, taken from the work of Rubin et al. (1980). Clear evidence is shown that velocities remain fairly constant with increasing distance from their galactic nucleus.

stars and gas of the cluster, behaves differently to the dark matter component as the smaller of the two clusters passes "through" the larger. The stars which make up the clusters are not greatly affected by the collision and most pass through completely, as observed in the optical regime. However, the hot gas component of each cluster makes up most of the baryonic mass and this gas interacts due to a drag effect, e.g. ram-pressure stripping, and emits in the X-ray. This causes the hot gas to be slowed down significantly during the collision and is represented by the red regions of the image. The distribution of the dark matter component, shown in blue, was detected indirectly due to the gravitational lensing effect of the Bullet cluster on background objects (Clowe et al. 2006). This strongly suggests the existence of a collisionless dark matter and provides evidence against theories of Modified Newtonian Dynamics (MOND - Milgrom 1983), as the latter predicts that the lensing should follow the large baryonic component, that is, the hot gas.

1.1.3 Cosmic Microwave Background Radiation

The strongest observational evidence of isotropy in the Universe is the almost exact uniformity of the cosmic microwave background (CMB) radiation. This is a faint background glow of radiation which is observed in every direction in the Universe and is believed to be the remnant radiation from the proposed Big Bang. After the Big Bang, the Universe was very hot and dense. As a consequence, photons were constantly being

Figure 1.3: Image of the Bullet cluster showing clear evidence of a non-interacting dark matter component (blue). The baryonic, hot gas component (red), as seen in the X-ray, remains in the central region of the collision. Image courtesy of nasa.gov.

scattered off the ions and electrons present in the ionized gas, resulting in an opaque fog. Approximately 379,000 years after the Big Bang the Universe had expanded, and thus cooled, to a temperature where electrons and protons were now able to combine to form neutral Hydrogen atoms. At this point, radiation decoupled from matter. No longer able to interact with the background gas, radiation was free to propagate and has been traveling towards us ever since. The CMB thus provides a snapshot of the Universe at a very early stage of its existence and is said to have been emitted from the hypothetical spherical surface called the 'surface of last scattering'. The time period in which neutral atoms were first formed is known as the epoch of recombination which occurred at the time of photon decoupling. Precise measurements of the CMB radiation are critical to cosmology, as any proposed model of the Universe must be able to explain this radiation.

The spectrum of the CMB radiation is the best representation of a black body spectrum ever observed, more accurate than any black body made in a laboratory. From this black body curve, the CMB radiation is found to be at a temperature of $\sim$2.726 K in almost every direction (Planck Collaboration et al. 2018). The almost

perfect uniformity of the CMB has also reinforced the cosmological theory that the Universe underwent a period of inflation (Liddle 1999), in which the Universe expanded exponentially at $\sim 10^{-36}$ seconds after the Big Bang for $\sim 10^{-4}$ seconds, and that all structure in the Universe arose due to this process. The radiation reaching us today has stretched out to longer wavelengths, due to the expansion of the Universe and therefore now falls in the microwave region of the electromagnetic spectrum.

Although the CMB is nearly perfectly uniform, it does contain tiny temperature fluctuations, or, anisotropies. When looking for temperature variations of the order of μK, it is very difficult to obtain accurate data from the ground due to noise caused by water vapour in the atmosphere, which absorbs radiation in the microwave region. Thus, the first accurate measurements of the CMB temperature anisotropies were measured by the Cosmic Microwave Background Explorer satellite (COBE - Smoot et al. 1992) satellite. COBE found the scale of these anisotropies to be one part in 100 000, corresponding to temperature fluctuations of only 0.00005 K. These minute over and under densities of the CMB radiation are imprints of tiny irregularities in the energy density of the early Universe, which grew over time under the influence of gravity and have resulted in the galaxies and galaxy clusters we see today. This provides further support for the ΛCDM model as it is able to explain the origin of galaxies and structure from the growth of these primordial fluctuations.

Another property that can be investigated via CMB observations is the geometry, or, curvature of the Universe, which is essentially the local shape of the observable Universe. In General Relativity, the local geometry of the Universe is modeled by the the Friedmann–Lemaître–Robertson–Walker (FLRW) model. This model asserts that if we ignore all possible forms of dark energy, then the curvature of the Universe can be determined by measuring the average density of matter within it. The density parameter, Ω, is the average density of the Universe divided by the critical energy density, which is the energy required for a flat Universe, or, a Universe with zero curvature. Density parameter values of $\Omega = 1$, $\Omega < 1$ and $\Omega > 1$ correspond to a Universe of zero, negative or positive curvature respectively, which in turn corresponds to a Universe

having either a flat, closed or open geometry. The curvature of the Universe can be inferred from the location of the first acoustic peak (to be further discussed in Section 1.1.4) featured in the CMB angular power spectrum, shown in Figure 1.4. The CMB angular power spectrum gives the distribution of matter as a function of angular scale and is essentially the Fourier transform of the angular correlation function of the CMB anisotropies. By assuming a value for the Hubble constant we can infer the curvature of the Universe from the angular size of the first peak in the CMB power spectrum.

Due to the fact that the COBE satellite could only achieve an angular resolution of 7 degrees, it could therefore only observe the largest angle fluctuations, which were larger than the angle corresponding to the first acoustic peak in the CMB power spectrum. Thus, after the COBE mission, a number of ground and balloon-based experiments, such as BOOMERanG, DASI and MAXIMA, were designed to measure CMB anisotropies on smaller angular scales than was possible with COBE, with a primary goal of measuring the angular scale of the first acoustic peak (Crill et al. 2003; Leitch et al. 2002; Rabii et al. 2006). MAXIMA and BOOMERANG clearly show strong fluctuations on a scale very close to one degree. As the size of the sound horizon is well known, it is simple geometry which shows that an angular scale of one degree gives evidence of a flat Universe. If the variations were any smaller or larger it would imply either a positive or negative curvature, respectively. In 2001, a second CMB satellite mission called the Wilkinson Microwave Anisotropy Probe (WMAP) was launched, which was 45 times more sensitive and armed with 33 times the angular resolution of its predecessor satellite COBE. The goal of WMAP was to follow from the observations of COBE and make higher precision measurements of large scale CMB anisotropies over the whole sky, with its much improved resolution of 13 arcminutes. The results from WMAP were presented after the 9-year data capture, which produced a 0.2 degree resolution map of the CMB. These measurements were able to provide precise constraints for the geometry, content and evolution of the Universe and as a resulting milestone, WMAP was a crucial element in developing the the ΛCDM model of cosmology. In particular, the data coincides very well with a flat Universe that is dominated by dark energy (71.4%) in the form of a cosmological constant (Hinshaw

et al. 2013). The WMAP measurements found the curvature of space to be within 0.4% of flat, Euclidean geometry and also supports the theory of cosmic inflation. The WMAP mission was also able to constrain various key cosmological parameters: The age of the Universe was measured to be 13.77 billion years to within 0.5%. They constrain the Hubble constant to $H_0 = 69.7 \pm 2.4 \text{kms}^{-1}$ Mpc^{-1}, the dark energy density to $\Omega_\Lambda = 0.721 \pm 0.025$ and the baryon density to $\Omega_b = 0.0463 \pm 0.0024$ (Hinshaw et al. 2013). Also of key importance was WMAP's precision determination that baryons make up only 4.64% of the Universe, measured to within 0.2% and finds the percentage of dark matter to be 24%. Figure 1.4 shows the angular power spectrum measurements made by the WMAP 9-year results (Hinshaw et al. 2013). The second acoustic peak gives information about the baryon density of the Universe, while the third provides information about the dark matter density. The second peak tells us about the amount of matter in the universe. The amount of radiation pressure, i.e. the amount of baryonic matter, competing with the the pull of gravity is quantified by how much the second peak is damped, i.e the smaller the second peak, the more matter in the Universe. The third peak tells us about the amount of dark matter in the Universe. Unlike baryonic matter, dark matter clumping experiences no competing force of radiation pressure, thus the height of the third peak is a measure of the matter/radiation ratio. Due to the fact that light and baryonic matter are related, the matter/radiation ratio tells us the amount of dark matter in the Universe. In conclusion, it is clear that the CMB is able to provide a wealth of cosmological information and will continue to be a powerful probe for cosmology going forward. The Planck mission which ran over 2009-2013 measured the CMB radiation at even higher sensitivity and resolution (3 times that of WMAP) and constrained the Hubble constant to a value of $H_0 = 67.4 \pm 0.5$ and a matter density of $\Omega_m = 0.315 \pm 0.007$ (Planck Collaboration et al. 2018).

1.1.4 Baryon Acoustic Oscillations

Baryon Acoustic Oscillations are fluctuations in the density of the baryonic matter of the Universe. In the early Universe, prior to the decoupling of photons from matter (described in Section 1.1.3), the overdensities in the primordial plasma would try to

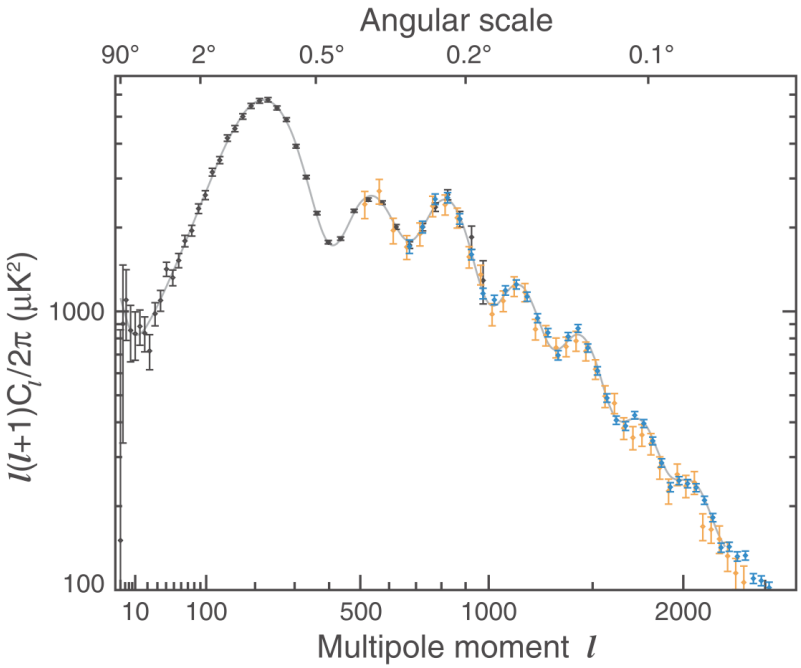


Figure 1.4: The angular power spectrum of the CMB from 9-year WMAP data (Hinshaw et al. 2013). WMAP results are shown in black, SPT data in blue and ACT data in orange. The grey line is the ΛCDM model fit to the WMAP data.

create clustering of the ionized gas by force of gravity. This was opposed by the outward radiation pressure resulting from the interaction of photons with matter. These competing forces result in oscillations in the primordial plasma, which are analogous to sound waves. Each sound wave, containing baryons and photons, moved outwards from the overdensity due to the radiation pressure, with a speed of just over half the speed of light. Dark matter interacts only via gravitational interaction, thus it remains at the origin of the overdensity while the photons and baryons move outward together. After decoupling occurs, photons which no longer interact with the baryons continue to move outwards, diffusing away. The result is an overdensity of matter at the origin of the wave, as well as in a shell at radius referred to as the sound horizon at the time of recombination. Thus, for each wave there remains a shell of baryons at a fixed radius, which provides a standard ruler for cosmology. Due to the fact that these over- and underdensities have developed into the large-scale structure and voids that we see today, we should be able to detect the signature of the BAO in the distribution of galaxies. Indeed, these baryon acoustic oscillations are seen in the CMB anisotropy data and also leave a faint imprint in the clustering of matter and galaxies today. The signature of baryon acoustic oscillations can be found directly from the galaxy auto-correlation

function or power spectrum. This phenomenon in galaxy clustering is found in the form of 'wiggles' in the power spectrum, or a single 'bump' in the correlation function (Eisenstein et al. 2005; Percival et al. 2010; Blake et al. 2011), as shown in Figure 1.5. This corresponds to a ∼1 percent 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. Thus, in much the same way that Type Ia supernovae (SNe) may provide a standard candle, the BAO signal give us a standard ruler, that is, a distance scale throughout the history of the Universe which is independent of cosmological model.

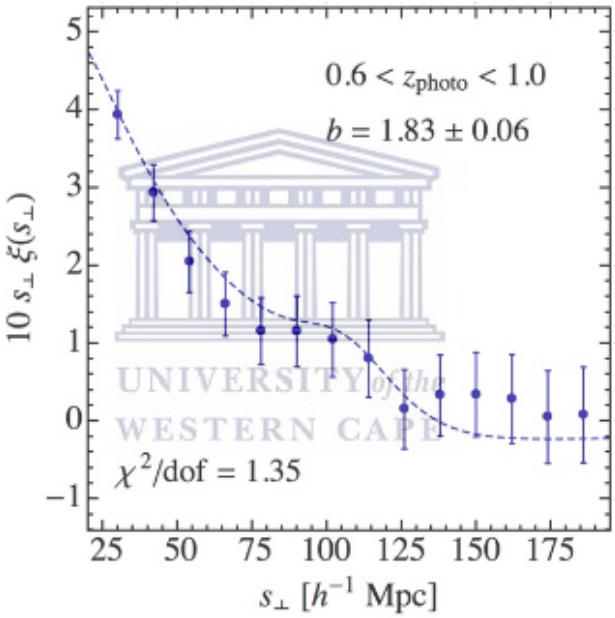


Figure 1.5: The signature of baryon acoustic oscillations at $\sim 110h^{-1}$Mpc, shown by the characteristic bump in the spatial correlation function of 1.3 million galaxies from the first year of data from the Dark Energy Survey (DES) (Crocce et al. 2019). The dashed line is the best fit model assuming linear bias and a smeared BAO feature as discussed in Abbott et al. (2019).

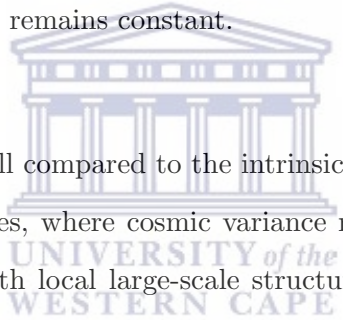
By the statistical measurement of a large number of galaxies, we find the BAO signal in an observed excess in the number of galaxy pairs separated by about 150 Mpc, which is the size of the sound horizon at the time of recombination. The resultant patterns in the distribution of distant galaxies can thus also be a vital tool for investigating the expansion history of the Universe. The standard ruler used is how-

ever, inconveniently large, and we therefore require wide surveys over a large redshift range in order to make use of this method. The largest and deepest redshift surveys to date have been those from the SDSS, undertaken using the Apache Point Observatory (APO) 2.5m telescope. The SDSS has measured the redshifts of ~ 4 million galaxies in spectroscopic observations from phases SDSS-I, II and III (2000-2014), described in the twelfth data release (DR12 - Alam et al. 2015). The Baryon Oscillation Spectroscopic Survey (BOSS) is a 5 year experiment (2009-2014) making up part of the SDSS-III (Eisenstein et al. 2011). BOSS has mapped out the distribution in redshift of $\sim$1.2 million Luminous Red Galaxies (LRGs) out to redshift ~ 0.75 over $\sim 9,300\text{deg}^2$ (Alam et al. 2017). Combined with supernova Ia data, their analysis yields a Hubble constant of $H_0 = 67.3 \pm 1.0\text{kms}^{-1}\text{Mpc}^{-1}$. The curvature of the Universe can be inferred from location of the first acoustic peak of the baryon acoustic oscillations featured in the CMB angular power spectrum. BOSS find a value for the curvature of the Universe to be $\Lambda_k = 0.0003 \pm 0.0026$, i.e. flatness at the 10^{-3} level and detect the BAO signal in three redshift slices to a significance of 8σ. The latest generation of large galaxy redshift surveys will allow for us to probe the giga-parsec volumes which are required to make precision measurements of the BAO signal. The Dark Energy Survey has measured the clustering of hundreds of millions of galaxies at various distances from us, but uses photometric redshift information. This will allow for the angular scale of the sound horizon to be determined for galaxy populations at different redshifts, which will provide information about the history of cosmic expansion to complement type 1a supernova measurements.

1.1.5 The Integrated Sachs-Wolfe Effect

An additional cosmological probe, which provides evidence for the existence of a dominant dark energy, comes from what is known as the Integrated Sachs-Wolfe Effect (ISW - Sachs & Wolfe 1967). Light traveling towards us from distant sources will pass through massive regions of over- and under-densities such as superclusters and supervoids, respectively. When falling into a gravitational potential well photons should gain energy, and lose the same amount of energy as it climbs out. However, at late times, due to the accelerated expansion of the Universe, these potential wells and voids

are being stretched/flattened out over time. This results in a net increase in energy for light passing through a large potential well, as less energy is required to climb out of this region. Similarly, there is a net loss in energy of a photon passing through a large void region. To measure this effect, the angular cross-correlation of the non-primordial CMB hot and cold spots and a large scale structure (LSS) map of similar resolution, is determined. This effect is important at late times ($z \leq 1$), once the Universe had begun its transition towards a dominant dark energy component and accelerated expansion. Thus, a Universe with a dominant constituent of dark energy will lead to a positive correlation at $z < 1$. This has been successfully detected by various groups, mostly with the use of data from WMAP, NRAO VLA Sky Survey (NVSS - Condon et al. 1998) and the Sloan Digital Sky Survey (SDSS - York et al. 2000) data (Scranton et al. 2003; Ho et al. 2008; Raccanelli et al. 2008). At early times, i.e. when the Universe is matter dominated, the gravitational potential wells do not evolve significantly, thus the gravitational potential remains constant.



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. Thus, cross-correlation with local large-scale structure, with extensive sky coverage, is required in order to produce a significant result. An example of such an analysis is presented in work by Giannantonio et al. (2012), who also discuss other ISW measurements and potential problems therewith. While several detections of the ISW effect had been made previously, cross-correlating the CMB maps from WMAP with radio, infrared, optical, and X-ray surveys, the results 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. In recent years, Nadathur & Crittenden (2016) performed ISW measurements using Planck CMB data and the CMASS galaxy sample from the BOSS DR12 measurements of ∼8,000 voids and ∼55,000 superclusters and found a significant detection of the ISW amplitude of $A = 1.64 \pm 0.53$, in agreement with the ΛCDM model. Making use of the latest CMB measurements and upcoming all-sky surveys of massive, low-redshift galaxies, as well as an understanding of their relationship

to the underlying dark matter distribution, the ISW could be a powerful cosmological tool.

1.1.6 Weak Lensing

Gravitational lensing is also a crucial probe for constraining ΛCDM cosmology. This is the process whereby the path of light rays from distant sources is bent by the intervening mass as it travels towards us. This idea that massive bodies could bend the path of light rays, creating a localised space-time curvature, was put forth by Albert Einstein at the beginning of the twentieth century as a prediction from his General Theory of Relativity. This theory was first confirmed during a solar eclipse in 1919 when Arthur Eddington observed that stars which passed close to the sun were slightly out of position (Dyson et al. 1920), which inferred that the light coming from them was being bent slightly. Thereafter, the theory of General Relativity received general attention. Einstein however, had only considered gravitational lensing by single stars and it was Fritz Zwicky who in 1937 put forth the notion of galaxies as gravitational lenses (Zwicky 1937). However, only in the year 1979 was the first gravitationally lensed object identified, and by accident at that. It became known as the 'Twin Quasar' due to the observed double images, separated by 6 arcminutes, of a single quasar (Walsh et al. 1979). The observing team noticed that the "two quasars" seemed to be unusually close to each other and had similar redshifts and observed spectra.

Gravitational lensing results in the image distortion and increased apparent magnitude of background sources, as well as multiple observed images of a single source. One of the most strongly lensed sources illustrating these effects is the very massive, distant Abell 2218 galaxy cluster (Kneib et al. 1996). The observed giant gravitational arcs and multiple images visible in Figure 1.6 are characteristic of the type of distortion and magnification resulting from strong gravitational lensing, a phenomenon where the lensing effects are caused by a single, large source. These strong gravitational lensing effects are very impressive, but also very rare. Most lensing of distant sources by the large scale structure which lies along the line of sight between us and

them, is in the weak lensing regime. Weak gravitational lensing results in distortions which are much weaker (∼10% of the distortion created by strong lensing) and thus may only be detected by analysis of a large number of sources in order to find coherent distortions and is thus an intrinsically statistical measurement. The effects of weak gravitational lensing are categorised by either the distortion or magnification of the background images (Bartelmann & Schneider 1999; Mellier 1999; Munshi et al. 2008). The distortion of galaxy images by weak lensing is referred to as cosmic shear. This effect induces correlated ellipticities in the images of galaxies that are not physically close to one another (Munshi et al. 2008; Hoekstra & Jain 2008). Sources behind a lens are also magnified in size, while the surface brightness of the source is conserved, this therefore results in an increase in the total observed flux of a source, a process known as cosmic magnification. Evidence for these effects are found in experiments where two spatially separated samples are found to be correlated. The induced correlations between the positions and magnitudes of high-redshift galaxies and foreground matter provides an excellent tool to study the intervening matter that lies between us and the distant lensed objects. Since the presence of dark matter can only be inferred indirectly from its gravitational effect on matter and radiation, gravitational lensing is a uniquely promising way to obtain a comprehensive knowledge of the nature and distribution of dark matter in the Universe and how this distribution has evolved over time. In the past two decades, conclusive measurements of the systematic distortion and magnification of sources in the distant Universe, such as high-redshift galaxies and quasars, due to large scale foreground structure at low redshifts, have been made by various groups (Scranton et al. 2005; Ménard et al. 2010; Wang et al. 2011; Ford et al. 2012).

Recent measurements of weak lensing were carried out by the Dark Energy Survey (DES). In the first study, they performed a cross-correlation between the weak lensing from galaxies measured with DES ($0.2 < z < 1.3$) with that of the CMB from Planck measurements (Omori et al. 2019). We expect the distortions from both the galaxy and CMB measurements to be correlated since the CMB photons pass through some of the same intervening gravitational potentials as the photons from distant galaxies. They

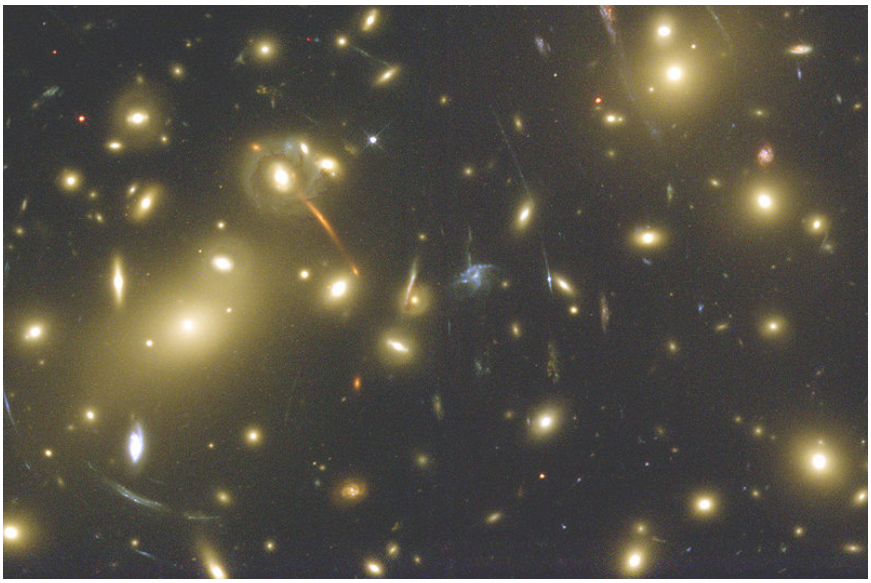


Figure 1.6: Observation of strong gravitational lensing effects in the Abell 2218 cluster. (Image courtesy of NASA-ESA, ERO team).

find a positive correlation with a clustering amplitude of $A = 0.99 \pm 0.17$, consistent with expectations from the ΛCDM model of $A = 1$ and a constraint on the matter density of $\Omega_m \sim 0.261$. Another survey which in recent years carried out measurements of weak gravitational lensing is the Kilo-Degree Survey (KiDS) (de Jong et al. 2013), a large optical imaging survey using the VLT Survey Telescope (VST), carried out over ~ 1500 deg^2 of the Southern sky overlapping with fields covered by the 2-degree-field (2dF) and SDSS and which was also imaged by the Visible and Infrared Survey Telescope for Astronomy (VISTA). The measurement of the weak gravitational lensing shear power spectrum, based on 450 deg^2 of KiDS imaging data. The cosmological parameter that is best constrained by the cosmic shear power spectrum data is a combination of the total matter density Ω_m and the clustering amplitude σ_8 called the parameter combination $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$ yielding $S_8 = 0.651 \pm 0.058$ (Köhlinger et al. 2017).

1.1.7 Combined Cosmological Probes

Using the combined information provided by the aforementioned studies and assuming a ΛCDM cosmology, it is possible to better constrain the matter and dark energy components of the Universe. An example of the advantage of one such combination

is shown in Figure 1.7 taken from Abbott et al. (2019), which shows constraints of Ω_Λ versus Ω_m. We see that the combined results from DES supernovae and CMB experiments are consistent with a flat Universe having a definite non-zero dark energy component. Thus, there is sound evidence for the presence of a dominant dark energy component of the Universe and hence, an accelerated expansion.

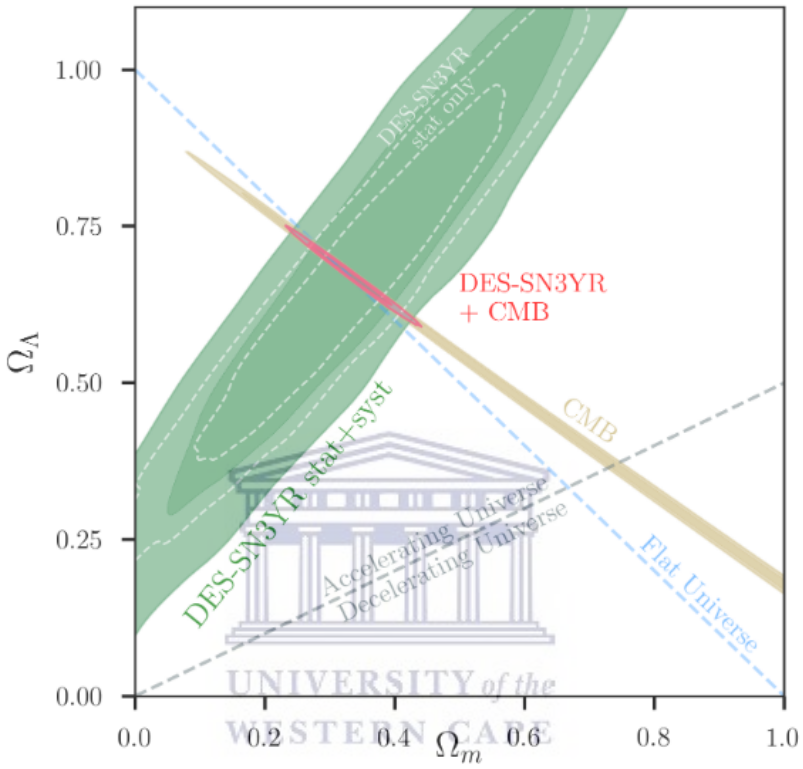


Figure 1.7: The confidence regions of 68.3% and 95.4% in the (Ω_m,Ω_Λ) using measurements of SNe from the Dark Energy Survey combined with constraints from the CMB, taken from (Abbott et al. 2019).

1.2 Large-Scale Structure

Any cosmological model must be able to account for the amount and distribution of observed structure in the Universe. The study of large-scale structure is that of the structure, or inhomogeneity, of the Universe on scales larger than that of a galaxy. The origin of all structure we observe in the Universe lies in pre-inflation quantum fluctuations, which were greatly amplified during inflation. Post inflation, the Universe was still nearly perfectly uniform, yet the slight (order 10^{-5}) over- and underdensities in

the very early Universe resulted in the galaxies and voids observed in the Universe today. The overdensities acted as seeds around which dark matter could gravitationally congregate, resulting in the potential wells. Thus, in ΛCDM cosmology it is believed that dark matter clusters over time causing regions of overdensity, which then gravitationally draw in the baryonic matter.

The CMB provides the earliest snapshot of what the primordial 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, which governs the distribution of baryonic matter in the form of the galaxies we are able to observe. By studying the Shapley-Ames catalogue, Hubble (1934) found an excess in the number count of galaxies, above Poisson, on angular scales less than 10°, but that the sample followed a Gaussian distribution on larger scales. Thus, while the Universe may appear to be homogeneous on the largest scales, it is clearly clustered.

Prior to 1989 it was widely assumed that viralized galaxy clusters were the largest existing structures, which were distributed almost uniformly throughout the Universe. Analysis of redshift survey data then revealed the existence of the first observed large-scale structure, a sheet of galaxies over 500 million light-years long and 200 million light-years wide, but only 15 million light years thick (Geller & Huchra 1989). In following years other redshift surveys showed evidence of similar structure, as well as galaxy superclusters and immensely large voids. From the 2dF (2-degree-field) Galaxy Redshift Survey (2dFGRS) (1997-2002) (Colless et al. 2001), we came to learn that large-scale structure exists as a network of filaments of galaxies, with huge, bubble like voids in between (see Figure 1.8 - Colless et al. 2001). Thus, it is suggested that galaxies form along the dark matter filamentary structure in the Universe, with galaxy superclusters forming at the nodes of these filaments.

Given that the distribution of structure in the Universe is dependent on the influence of the matter density and dark energy properties of the Universe, as well as the

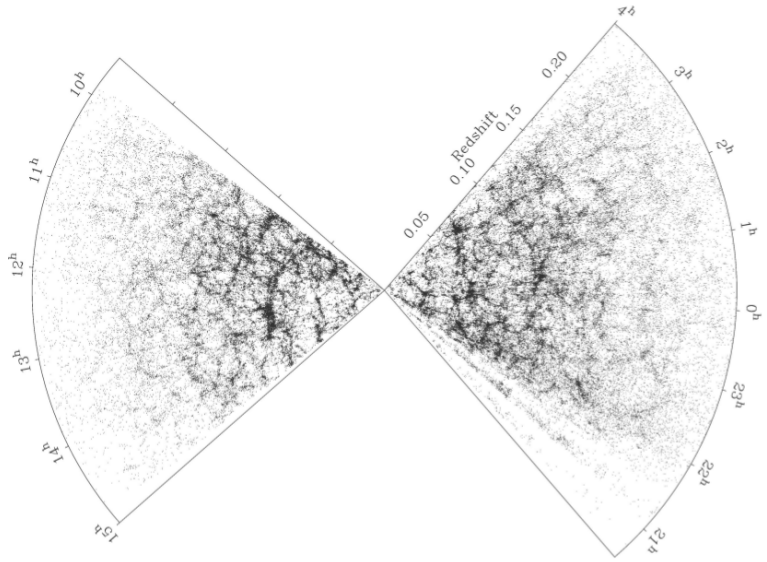


Figure 1.8: Wedge showing the projected 2dFGRS distribution of galaxies as a function of redshift and RA (Colless et al. 2001). The filamentary nature of the large-scale structure is clearly visible, with sheets separated by large voids and dense superclusters of galaxies at filament nodes.

physics of galaxy evolution, measurements of large-scale structure therefore enable us to place constraints on cosmology and the evolution of galaxies. The majority of experiments which aim to constrain cosmological parameters require information about the mass distribution of the Universe and how it evolves with cosmic time. In the ΛCDM model, this is essentially the dark matter distribution. However, as the dark matter distribution can not be observed directly, we instead quantify the clustering of the observed structure in the Universe and infer the dark matter clustering. Below I describe methods which may be used in order to accomplish this.

1.2.1 Methods for Measuring Large-Scale Structure

The use of the methods to measure the clustering properties of a sample have various applications in cosmology. Understanding the clustering properties of a population is a stepping stone toward an understanding of the underlying dark matter distribution, as well as properties of dark energy.

The Angular Correlation Function $\omega(\theta)$

The most commonly used quantitative measure of large scale structure, due to its simplicity and relation to the power spectrum, is the galaxy two-point correlation

function. The angular correlation function (ACF - (Peebles 1980)), which requires only the two-dimensional position of galaxies, has been employed extensively for data which lack accurate spectroscopic information. It is a simple measure of the angular distribution of sources on the sky and is a projection of the spatial clustering of galaxies on the plane of the sky. The two-point correlation function (TPCF) measures the excess, or deficit, in clustering, as compared to a random sample. The random sample must have the same areal coverage and contain the same biases as the data sample, but will comprise of randomly distributed 'mock' sources. The probability δP of finding a pair galaxies at an angular separation θ is therefore given by

$$\delta P = N[1 + w(\theta)]\delta\Omega, \tag{1.2}$$

where $\omega(\theta)$ is the angular two-point correlation function, N is the mean number of galaxies per steradian and $\delta\Omega$ is the solid angle.

There are many estimators which may be used to calculate the two-point correlation function. All perform similarly at small scale separations, but edge effects and/or holes in the survey cause an increase in the variance at large scales. This is due to the fact that sources at the edges have fewer neighbours, thus pair counts would be systematically underestimated at large separations. To compensate for this, when counting the number of neighbours at a given angular separation θ, estimators omit sources which lie at a distance less than θ from the border/edge. Therefore, at large scales, only a small number of sources are used for the estimation, thus increasing the variance. It is for this reason that the favoured estimators are those which incorporate an edge-correction. These are known as the Hamilton (Hamilton 1993) and Landy-Szalay (Landy & Szalay 1993) estimators, the latter of which is used in this work, due to it having minimal statistical bias. It is least sensitive to the number of random sources used (Kerscher et al. 2000).

The form of the Landy-Szalay estimator is as follows,

$$\omega(\theta) = \frac{DD - 2DR + RR}{RR}, \tag{1.3}$$

where DD and RR represent the normalised number of galaxy pairs at a given separation θ that exist in the data and random catalogue, respectively. The additional cross correlation between the data and random sets makes for a better estimator by minimizing the effect of fluctuations in the mean galaxy density at small scales. This estimator is also found to have smaller uncertainties at large scales than other pairwise estimators (Kerscher et al. 2000). By using a random sample that is > 10 times more densely populated than the data set, we can assume the statistical error in the random set to be negligible. The uncertainty in $\omega(\theta)$ is thus naively the Poisson variance for the data (DD) pairs,

$$\Delta\omega = \frac{1 + \omega(\theta)}{\sqrt{DD}}. \tag{1.4}$$

However, the assumption of Poisson statistics has been shown to underestimate the true variance in the TPCF (Norberg et al. 2009; Cress et al. 1996). The uncertainty can be calculated in a more rigorous fashion using the Jackknife re-sampling technique (Lupton 1993). In this technique, the data is split into N bins in RA and the correlation function is calculated N times, each time leaving out a different subset of the data sample. The Jackknife errors are then obtained from the variance in ω,

$$\sigma_{\omega_{mean}} = \sqrt{(N-1)\sum_{i=1}^{N}\frac{(\omega_i - \omega)^2}{N}}. \tag{1.5}$$

The measured angular correlation function does however suffer from an induced negative offset, which arises due to the finite size of the survey. As discussed by Roche & Eales (1999), when the angular scale becomes comparable with that of the survey area, edge effects become important. We correct for this offset by adding a constant known as the *integral constraint* expressed as follows

$$\omega_{obs}(\theta) = \omega(\theta) - \sigma^2, \tag{1.6}$$

where $\omega_{obs}(\theta)$ is the ACF measured from observations, $\omega(\theta)$ is the true ACF and the integral constraint σ^2 can be approximated with

$$\sigma^2 = \frac{\sum RR(\theta)w(\theta)}{\sum RR(\theta)}. \tag{1.7}$$

Spatial Clustering

In order to relate the distribution of galaxies to that of the underlying dark matter, we first require information about the spatial clustering of the sample. Direct computation of the 3D/spatial two-point correlation function $\xi(r)$ requires accurate redshift information for each source. However, for surveys which lack accurate spectroscopic redshifts, the angular correlation function $\omega(\theta)$ described above can be easily de-projected to obtain information about the spatial correlation function $\xi(r)$ using a technique called *Limber Inversion* (Limber 1953) and information regarding the redshift distribution $N(z)$ of the sample. Limber Inversion allows for the calculation of the correlation length r_0. This is not a physical length, but rather represents a separation length of the spatial correlation function at which $\xi(r) = 1$. Thus, at scales below this separation length, the number of data-data pairs are more than twice that of the random-random pairs. Therefore, a larger correlation length would correspond to a more clustered sample.

As in the case of the angular two-point correlation function (discussed further in Section 2.2), the spatial two-point correlation function can also be expressed in the form of a power law (Peebles 1973) as follows,

$$\xi(r, z) = \left(\frac{r_0}{r}\right)^{\gamma}, \tag{1.8}$$

where r is the separation distance in comoving units and $\gamma = 1 - \eta$, where η is the slope of the measured correlation function. If we assume a redshift dependent form of the spatial correlation (Overzier et al. 2003) we then have

$$\xi(r, z) = \left(\frac{r_0}{r}\right)^{\gamma} \times (1 + z)^{\gamma - (3+\epsilon)}. \tag{1.9}$$

Here the evolution of clustering is parameterised by ϵ and the chosen clustering

model for this Thesis is $\epsilon = \gamma - 3$, the case in which the clustering is constant in comoving coordinates. To relate the angular and spatial correlation functions, the clustering amplitude A is expressed as a function of r_0 (in comoving coordinates) as follows, in what is termed the *Limber Equation* (Overzier et al. 2003; Kim et al. 2011; Elyiv et al. 2012)

$$A = r_0^{\gamma} H_{\gamma} \left(\frac{H_0}{c} \right) \frac{\int_0^{\infty} N^2(z)(1+z)^{\gamma-(3+\epsilon)} \chi^{1-\gamma}(z) E(z) \mathrm{d}z}{\left[\int_0^{\infty} N(z)\, \mathrm{d}z \right]^2}, \tag{1.10}$$

where H_{γ} is related to the Gamma function, $H_{\gamma} = \Gamma(\frac{1}{2})\Gamma(\frac{\gamma-1}{2})/\Gamma(\frac{\gamma}{2})$, $N(z)$ is the redshift distribution and $\chi(z)$ is the comoving line-of-sight distance to an object at a redshift z.

$$\chi(z) = \frac{c}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')}. \tag{1.11}$$

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

$$E(z) = \left[\Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0} \right]^{\frac{1}{2}}. \tag{1.12}$$

The Limber equation (Equation 1.10) is then inverted to calculate the correlation length r_0 as a function of the measured angular clustering amplitude A, the slope η of the angular correlation function and the redshift distribution $N(z)$ of the respective sample.

The Galaxy Bias

The galaxy bias provides information regarding the relationship between the clustering of dark matter and that of luminous, baryonic matter, i.e. how strongly a sample of galaxies traces out the underlying dark matter distribution. Having inferred the spatial clustering of the sample, we are able to use the correlation length r_0 to estimate the galaxy bias $b(z)$ at the median redshift of the sample. The galaxy bias is defined as the ratio of the spatial clustering of the galaxies to that of the dark matter halos, as

introduced by Kaiser (1984) and Bardeen et al. (1986):

$$b^2(z) = \frac{\xi_{\rm gal}(r,z)}{\xi_{\rm DM}(r,z)}, \tag{1.13}$$

where the numerator and denominator are the galaxy and dark matter spatial correlation functions, respectively. Given that the cosmological parameter σ_8 measures the amplitude of the linear power spectrum on the scale of 8 h^{-1}Mpc, we set r=8 h^{-1}Mpc. Thus, using Equation 1.8 we can write the spatial galaxy correlation function as

$$\xi_{\rm gal}(8,z) = \left[\frac{r_0(z)}{8}\right]^{\gamma}. \tag{1.14}$$

The dark matter correlation function is given by Peebles (1980) as

$$\xi_{\rm DM}(8,z) = \sigma_8^2(z)/J_2 \tag{1.15}$$

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. Therefore, given only the correlation length r_0 and slope η of the angular correlation function, we are able to calculate the bias:

$$b(z) = \left[\frac{r_0(z)}{8}\right]^{\gamma/2} \frac{J_2^{1/2}}{\sigma_8 D(z)/D(0)} \tag{1.16}$$

where $D(z)$ is the linear growth factor at a given redshift. The growth factor determines the normalisation of the amplitude of fluctuations in the large-scale structure, relative to those in the CMB, and is calculated using the formula from Hamilton (2001) as follows

$$D(z) = \frac{5\Omega_m}{2} \times \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda} \times \int_0^z \frac{1+z}{(\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda})^3}. \tag{1.17}$$

In this Thesis, all bias values will be calculated at the median redshift of the

respective sample.

1.3 Radio Astronomy

In this Thesis, I use radio sources in order to measure the large-scale structure. I therefore give a brief overview of radio astronomy, discuss the types of galaxies found in radio surveys and present past, current and future radio surveys which are relevant to this work.

1.3.1 Overview

The radio window was the first new spectral range to become available to astronomy after visible and near-infrared. This advancement in astronomy began in 1931 when Karl Guthe Jansky discovered radiation at 20.5 MHz (14.6m) that was being emitted by an extraterrestrial source which was not the sun, using a direction sensitive antenna array. He was able to conclude that this radiation was coming from the constellation of Sagittarius (Jansky 1933b; Jansky 1933a). Subsequent observations showed radio sources throughout the Milky Way. Due to the fact that the Sun was a very weak emitter, Jansky concluded that the "noise" was from non-stellar sources (Jansky 1935), and also speculated that the emission was generated in an ionized interstellar medium (ISM). The full significance of radio astronomy was only realised in the 1950s and 1960s, with discoveries of the 21-cm hydrogen line (Ewen & Purcell 1951), quasars (Matthews & Sandage 1963), the CMB (Penzias & Wilson 1965) and pulsars (Bell Burnell 1977).

There are several advantages to radio astronomy that highlight why it has and will play a vital role in the progression of astronomy and cosmology. Observations of radio signals are Earth based and can operate in most conditions. The telescope can operate during during day or night and the longer wavelengths of radio emission mean that these signals from space are not hampered by cloud or poor weather conditions on Earth. Likewise, the emission from distant sources in the Universe are not obscured by dust, and can travel through these regions without being reflected or absorbed,

making many more sources accessible. The Universe is primarily composed of neutral hydrogen, which is only observable in the radio spectrum, as hydrogen atoms produce radio emission at a wavelength of 21cm (frequency of 1420 MHz). Observations of hydrogen are vital in understanding galaxy formation and the distribution of matter in the Universe.

Early radio surveys such as the Third Cambridge Catalogue (3C - Edge et al. 1959) and the revised version 3CR (Bennett 1962) established the field of radio astronomy and highlighted the variety in extragalactic radio sources by allowing for identification with optical sources due to their improved resolution. A large improvement in sensitivity was then required in order to observe more than just the most powerful radio sources and to sample higher-redshift populations. Single-dish facilities, such as the Green Bank Telescope at the National Radio Astronomy Observatory (NRAO) and the Arecibo Observatory in Puerto Rico have performed well in producing large radio surveys of tens of thousands of sources. However, in the last few decades, it is large radio interferometer arrays which have carried the majority of the workload. Radio telescopes must have very large aperture diameters (D) to achieve good angular resolution, as $\theta \sim \lambda/D$. Even the biggest precision radio telescopes, such as the previously mentioned Green Bank Telescope (GBT), with $D = 100$m are still limited to poor resolution. Large multi-array radio telescopes are able to implement long baselines (distance between antennas) and thus achieve aperture diameters up to 1000s of km using very long baseline interferometry (VLBI). Interferometers also allow for extremely accurate astrometry, due to the fact that interferometric positions depend on measuring time delays between telescopes rather than on the mechanical pointing errors of telescopes, and clocks are far more accurate than rulers. Radio astronomy now strives for measurements with higher sensitivity, resolution and at shorter wavelengths.

1.3.2 Active Galactic Nuclei

Many galaxies observed in the radio have very bright nuclei, which can be more luminous than the remaining galaxy light. These are galaxies which have an Active Galactic Nucleus. Active Galactic Nuclei (AGN) are associated with the accretion of

material onto supermassive black holes. Most normal galaxies have a supermassive black hole at their center, and in an active galaxy, it is understood that the AGN is powered by the accretion of material onto a central supermassive black hole from the galaxy's dense central region (Salpeter 1964; Zel'dovich & Novikov 1965; Lynden-Bell 1978). As the material falls in towards the black hole, angular momentum will cause it to spiral in and form a disk. This disk, called an accretion disk, heats up due to the gravitational and frictional forces at work. Most of the energy output from AGN, at radio wavelengths, is non-thermal synchrotron emission, which is generated by charged particles spiralling around magnetic field lines and thus accelerating. However, many AGN are also strong emitters in the X-ray, ultraviolet, as well as optical. The in-fall of material at an extremely rapid rate results in an immense amount of released energy and radiation, making AGN highly luminous sources. AGN are also often characterised by radio jets which can extend to the scale of a megaparsec and emanate from the AGN at the centre of the host galaxy.

Seyfert Galaxies

The first optical spectrum of an active galaxy was obtained by E. A. Fath in 1908, noting strong emission lines in NGC 1068 which were later resolved to have widths of several hundred km s^{-1}. Subsequently a study of six spiral nebulae by Carl Seyfert (Seyfert 1943) set out the template for a new morphological classification of galaxies. Seyfert galaxies are identified primarily by their appearance, that is, ordinary spiral galaxies with high surface brightness nuclei. Their spectra are also characterised by strong, high-ionisation emission lines which further subdivided Seyferts into two categories, Seyfert 1 and Seyfert 2 galaxies, 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 km s^{-1} , but the permitted emission lines are very broad, corresponding to velocities of 1,000–5,000 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, sug-

gesting an origin from the same region in the host galaxy. Example spectra for Seyfert 1 and Seyfert 2 galaxies taken from Ho et al. (1993) are shown in Figure 1.9. A further class of AGN related to Seyferts are low-ionisation nuclear emission regions (LINERs - Heckman 1980). These are similar to Seyfert 2 galaxies, except with forbidden lines arising from less ionised atoms.

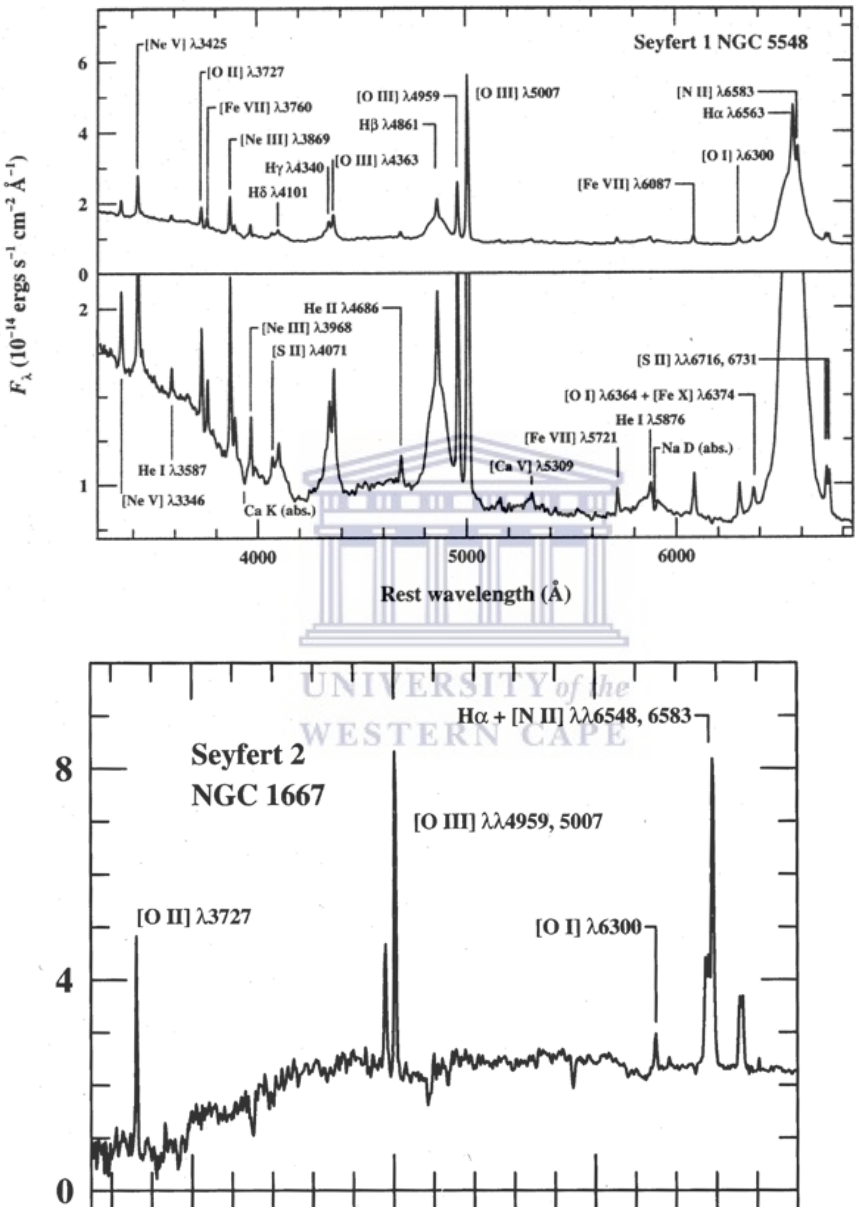


Figure 1.9: Example spectra of Seyfert 1 (NGC 1275 - Top) and Seyfert 2 (NGC 1667 - Bottom) galaxies taken from Ho et al. (1993). The Seyfert 1 spectrum is the mean of several observations made in 1993 with the 3m Shane Telescope at the Lick Observatory. In the Seyfert 1 spectrum, the prominent broad and narrow emission lines, as well as the absorption features are labeled. In the Seyfert 2 spectrum we see the identified emission lines.

Quasars

Early radio surveys such as 3C (Edge et al. 1959) and 3CR (Bennett 1962) observed some sources at high galactic latitude which were coincident with objects that looked like stars at visible wavelengths, 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 realised when 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 highest measured at the time, and an order of magnitude larger than the original Seyfert galaxies. The extreme distances of these sources implied very high luminosity, with 3C 273 being the brightest known quasar ($B = 13.1$mag) which is $\sim$100 times more luminous than Milky Way type galaxies.

While the discovery of quasars was based on their radio emission, the majority of quasi-stellar objects are in fact not powerful radio emitters (Sramek & Weedman 1980; Condon et al. 1981). The terms 'quasar' and 'QSO' are now used interchangeably, but can be subdivided into radio-loud quasars (RLQs - Barr et al. 2003; Wang et al. 2003a) and radio-quiet quasars (RQQs - Kukula et al. 1998; White et al. 2017). The commonly accepted technical definition of a radio loud source is one which has a ratio of 5 GHz radio flux to B-band ($\sim 4000\mathring{A}$) optical flux ≥ 10 and < 1 for RQQs (Kellermann et al. 1989). It is in fact only a small percentage of AGN, approximately 5-10%, which are radio-loud. The radio emission of these radio-loud sources sometimes results from synchrotron emission associated with powerful relativistic jets of plasma which are ejected from a region very close to the supermassive black hole (SMBH) and which interact with a magnetic field. Despite efforts to understand the mechanisms responsible for the launch of relativistic jets, their origin remains unclear. It is also unclear why they appear to occur in so few of known AGN. Further subclasses of quasars are BL Lacertae (BL Lac) objects and optically violent variables (OVVs), collectively known as 'blazars'. Blazars are thought to be AGN which have one radio jet aimed at the observer. While most AGN exhibit variability in their continuum emis-

sion 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 that the regions responsible for this emission are less than 1 light-day across. This variation is also 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 by the lack of emission or absorption lines in their spectra. This may be due to intrinsic variability in the sources, with emission lines being visible only when the continuum is relatively faint.

The Unified Model of AGN

The above descriptions are of independently discovered classes of objects, however the suggestion is that they are all different types of AGN which are part of a unified model which states that the central regions of all AGN are thought to be similar and that the variation in properties is related to the line of sight into the central region of the AGN, i.e., the angle from which we observe them. This would account for the different characteristics in their observed spectra (Barthel 1989). Thus, the different AGN subclasses represent different orientations of an intrinsically similar set of objects. 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, where in the case of RLQs, the relativistic beaming is directed toward us. 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).

Figure 1.10 taken from Beckmann & Shrader (2012) 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 source observed. The emission is powered by the supermassive black hole at the centre of the galaxy, accompanied by a thin, hot accretion disk occupying the region of ~ 1 parsec around the black hole. Vast amounts of gravitational energy are released as matter falls towards 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 (corresponding to Seyfert 1s, Type 1 QSOs and Broad Line Radio Galaxies (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 Narrow-Line Radio Galaxies (NLRGs).

An orientation-based model, however, is unable to explain certain observations, such as 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 $10^8 \mathcal{M}_\odot$ (McLure & Jarvis 2004), while radio-quiet QSOs may have black hole masses as low as $10^7 \mathcal{M}_\odot$, suggesting this mass to be a possible limiting factor in radio emission. Alternatively, the smaller numbers of radio-loud AGN as opposed to radio-quiet may lend credit to the suggestion that as part of the life cycle of an AGN, a phase of radio loudness occurs (Urry 2003). Recent studies by Kellermann et al. (2016) find that the radio luminosity function of optically selected quasars and the extended radio emission from RLQs are inconsistent with a unified model's differentiation between RLQs and RQQs.

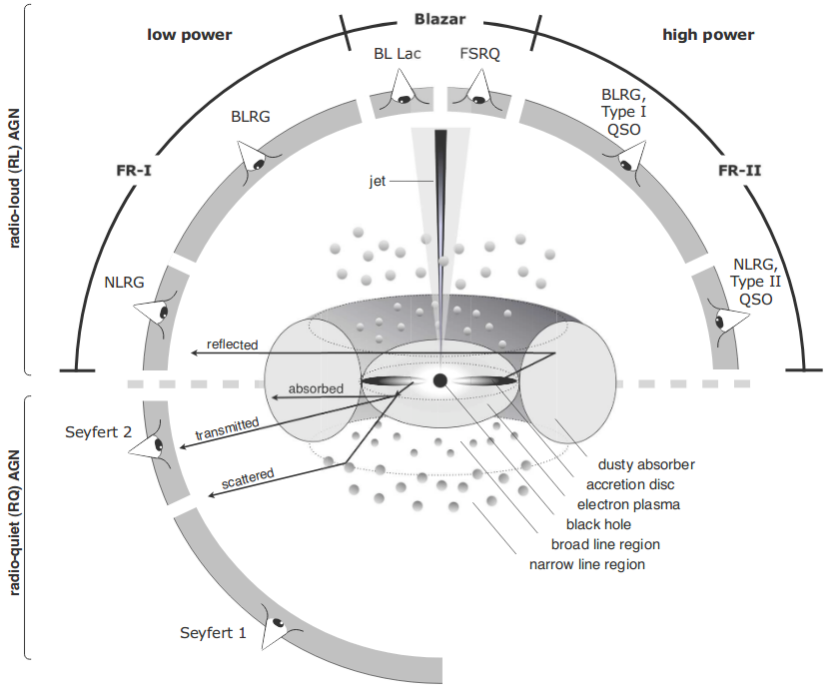


Figure 1.10: A schematic diagram of the archetypal AGN, showing the physical components and orientation giving rise to particular observed AGN types (Beckmann & Shrader 2012).

1.3.3 The Radio Source Population at > mJy Flux Densities

The radio source population observed at flux densities greater than 1 mJy is comprised almost entirely of radio-loud AGN sources. This refers to all AGN which have a high radio-to-optical flux ratio. The two main classification schemes for these radio-loud AGN sources are described below.

• Fanaroff-Riley classification: AGN were originally classified into one of two groups, namely Fanaroff-Riley I (FRI) and Fanaroff-Riley II (FRII) (Fanaroff & Riley 1974). These two subclasses are based on the morphology of the radio galaxy, in particular, the characteristics of the emission from the radio lobes. FRI galaxies have radio lobes which are diffuse and emission which is brightest at their centers. FRII radio galaxies have characteristic large jets which are more collimated, with hotspots located at the edges of the lobes. FRII sources are generally the more luminous of the two subclasses, with a dividing luminosity of about $L_{1.4\mathrm{GHz}}$= 10^{25} W Hz^{-1}. There are however, some

radio galaxies which fall into a hybrid class in which each of the two lobes displays a different Fanaroff-Riley morphology. This observation supports the popular notion that the nature of these relativistic jets are primarily dependent on the nature of the external media with which they interact (Gopal-Krishna & Wiita 2000). 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 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. Spectral index, which measures the dependence of the flux density of a source on frequency ($S_\nu \propto \nu^\alpha$) is thus used as another common means of subdividing radio galaxies into steep-spectrum or flat-spectrum categories, with the division made somewhat arbitrarily at $\alpha = -0.4$. Examples of the spectra of various radio source types are shown in Figure 1.11 taken from Verschuur & Kellermann (1988). Extended sources tend to have a steep spectrum, while more compact sources have a flat spectrum. AGN are also observed with one jet often fainter than the other, or missing entirely. 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.

• High and low excitation radio galaxies: In recent years it has been found that there could be a division of subclasses which is semi-independent of morphology and is instead related to different modes of black hole accretion within the AGN. This classification is dependent on the presence or absence of narrow, high-excitation, emission

lines in the spectra of the AGN host galaxies (Hine & Longair 1979a; Jackson & Rawlings 1997; Willott et al. 2001). AGN with or without these high-excitation emission lines are classified as either high-excitation (HERG) and low-excitation (LERG) radio galaxies, respectively. The predominant case is that low luminosity FRI galaxies are LERGs and HERGs are the more powerful FRII sources. However, there are discrepancies where many FRII galaxies have been found to be low-excitation radio galaxies (Evans et al. 2006). It has thus been argued that the HERG and LERG characteristic spectra are attributed to different modes of accretion onto the SMBH of the AGN. HERGs and LERGs are further discussed and studied in chapter 4 of this Thesis.

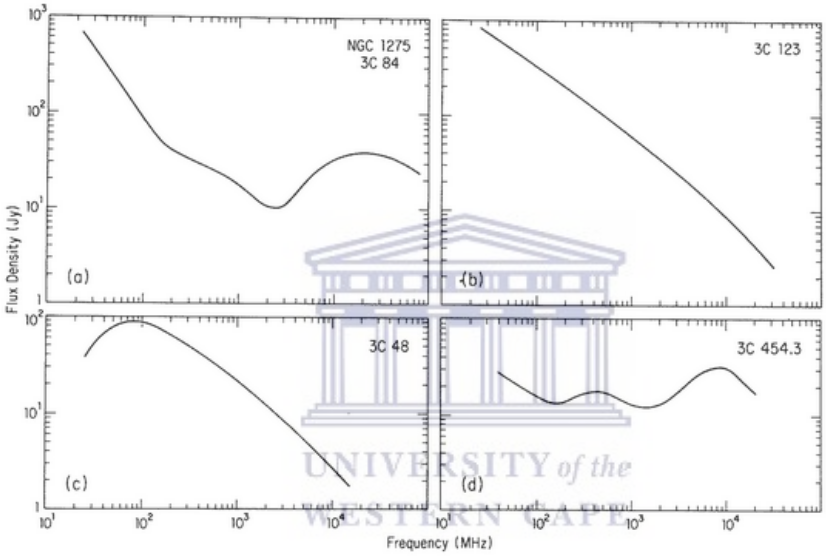


Figure 1.11: Spectra of various radio galaxies from Verschuur & Kellermann (1988) (a) shows the spectrum of a compact source, while (b) and (c) show the steep spectrum of extended sources and (d) shows the spectrum of a quasar which contains structures of different sizes which become opaque at some frequencies.

1.3.4 The Sub-mJy Radio Source Population

At sub-mJy flux densities, the radio source counts are dominated by ordinary, non-active, star-forming galaxies (Condon 1992). Thus, the radio population at these flux levels is thought to comprise of a mixture of star-forming galaxies and a number of high-redshift and lower redshift radio-quiet AGN (Jarvis & Rawlings 2004). The contributions of each of these source types to the sub-mJy population has been a subject of debate for many years (Simpson et al. 2006; Seymour et al. 2008; Smolčić

et al. 2009a; Padovani et al. 2009; Kimball et al. 2011; Bonzini et al. 2012; Condon et al. 2013; White et al. 2017). In the past, progress in investigating the relative proportions of radio source types which make up the sub-mJy population has been hampered due to the compromise of deep surveys and survey area. Study of the sub-mJy radio population has also been hampered by the scarcity of spectroscopic information which is necessary for the accurate classification of these sources, as these are usually only available for brighter sources. Results from recent work investigating the nature of the faint radio population however, have shown evidence that at the level of a few tens of μJy, the radio source population comprises of a fairly equal contribution of AGN and star-forming galaxies, with radio-quiet AGN making up $\sim$25% of all AGN at this faint end (Simpson et al. 2006; Padovani et al. 2009; Magliocchetti et al. 2017; Hale et al. 2018). There are also galaxies in which the processes of star formation and black hole accretion each make significant contributions to the observed radio emission. These objects are of considerable interest in terms of galaxy evolution studies, as it is thought that the interplay between bursts of star formation and black hole accretion could be of particular importance in understanding the fueling mechanisms of AGN, as well as the quenching of star formation over cosmic time.

1.3.5 Radio Surveys

Research in the field of radio astronomy has contributed largely to our current understanding of the Universe. Progress in this field has seen the discovery of various phenomena and objects such as the cosmic microwave background, quasars, pulsars and black holes, some a major milestone in the field of cosmology. In this section I provide a description of various current and near-future radio continuum surveys, as well as some of their key science goals which are relevant to cosmology and large-scale structure.

Current, near-future and next-generation radio continuum surveys

Radio astronomy is now entering the era of very large, sensitive and high resolution surveys. As a result, our view of the Universe at radio wavelengths is changing dramatically and precision cosmology using radio continuum surveys is within reach. Presented

below are some of the current and next-generation radio telescopes and their proposed continuum surveys.

• ASKAP- The Australian Square Kilometer Array Pathfinder was built at a radio quiet site in Western Australia and has begun doing science (Johnston et al. 2008). ASKAP consists of 36 12m diameter dishes spread over a region 6km in diameter. The key feature of ASKAP is its wide field of view of 30 square degrees, generated by its unique phased array feed (PAF) receivers. With the use of specialised digital systems, the PAFs create 36 separate (simultaneous) beams on the sky which are then mosaicked together into a large single image. The planned continuum survey project of the ASKAP telescope is EMU-Evolutionary Map of the Universe (Norris 2012). The EMU survey is a deep, wide survey which will observe objects down to a rms flux of 10μJy beam^{-1} at 1.3GHz over 75% of the entire sky. The survey will probe star forming galaxies out to redshift ~ 0.5, powerful starbursts at higher redshifts, AGN to the edge of the visible Universe and is also highly likely to find new classes of objects. There are three key science goals for EMU (Norris 2009): (i) To measure the distribution of radio sources in the Universe and explore large-scale structure in order to estimate cosmological parameters (ii) To trace the evolution of star forming galaxies from $z = 2$ to the present day, using a wavelength that is unbiased by dust or molecular emission. (iii) To trace the evolution of massive black holes throughout the history of the Universe and understand their relationship with star formation.

To trace the evolution of star-forming galaxies from z=2 to the present day, using a wavelength unbiased by dust or molecular emission. •To trace the evolution of massive black holes throughout the history of the Universe, and understand their relationship to star-formation. •To use the distribution of radio sources to explore the large-scale structure and cosmological parameters of the Universe. •To explore an uncharted region of observational parameter space, almost certainly finding new classes of object.

• LOFAR- The Low Frequency ARray (van Haarlem et al. 2013) is a multi-national telescope with its core situated in the east of the Netherlands. It has an innovative design based on an array of simple omni-directional antennas instead of large mechanical

dish antennas. This allows for digital beam forming towards any part of the sky. In the past, low frequency radio astronomy has been neglected due to the poor resolving power of the available radio telescopes. This is due to the fact that the spatial resolution of a telescope is proportional to its operating wavelength. Thus, at frequencies below 250MHz, most dish-based radio telescopes would have resolutions of arcminutes and the radio images would be blurred by a factor of several thousand compared to images in the optical. LOFAR is able to address this problem at frequencies below 250 MHz by building dipole antennas and combining their signals to simulate large dishes. In addition, LOFAR makes use of very long baselines, as it has stations spanning across Europe. The resulting resolution is of the order $\sim$1 arcsecond. The telescope will carry out large-sky surveys between 10-250 MHz, a part of the electromagnetic spectrum that is relatively unexplored (Röttgering 2003) and will thus be sensitive to different source populations. There are four key science goals for the LOFAR surveys: (i) To use $z \geq 6$ radio galaxies as probes to study the formation of massive galaxies, clusters and black holes. (ii) To use diffuse radio emission detections and Faraday rotation measures in galaxy clusters as probes to study intracluster magnetic fields. (iii) Study star formation in the early Universe by probing starburst galaxies. (iv) To explore new parameter space with the hope of making new and unexpected discoveries.

• WSRT- The Westerbork Synthesis Radio Telescope situated in the Netherlands underwent a major upgrade in 2013 as part of the APERTIF (APERture Tile In Focus) project, which consists of a new phased array feed receiver system (Röttgering et al. 2011b). It consists of a linear array of 14 antennas, each with a diameter of 25 metres and arranged on a 2.7 km East-West line. The telescope can operate at several frequencies between 120 MHz and 8.3 GHz, is part of the European VLBI Network and is operated by ASTRON. The phased array feeds will increase the survey speed of the telescope by a factor of 25, allowing for an 8 square degree field of view. The planned survey for the WSRT is the Westerbork Observations of the Deep Apertif Northern (WODAN - Röttgering et al. 2011a) sky survey and will be conducted using APERTIF. WODAN will survey the entire northern sky above +30° at 1.4 GHz, down to a proposed rms flux density of 10μJy beam^{-1}. This survey is essentially the Northern

version of EMU, as WODAN and EMU have almost precisely the same sensitivity, will probe the same redshift distribution and yield similar galaxy numbers. The combined data therefore has the potential to provide a deep, homogeneous, full-sky catalogue. The combination of the EMU and WODAN would yield unprecedented ISW measurements due to it being an all-sky survey.

• MeerKAT- The MeerKAT array is an SKA precursor telescope built in the remote Karoo area of South Africa and will be integrated into the mid-frequency component of SKA Phase 1. MeerKAT consists of 64 dishes, each 13.5m in diameter. The dishes are distributed such that 70% is contained in an inner core component distributed over 300m, with the longest baseline being 1km, while the remaining 30% is contained in an outer core which covers 2500m and a longest baseline of 8km. The frequency bands available with MeerKAT are as follows: 580 - 1015 MHz, 1000 - 1750 MHz and 2000 - 4000 MHz. The unprecedented sensitivity and high resolution of this telescope guarantee many important uses for cosmology. The science goals of the MeerKAT surveys are in line with the key science goals for the first phase of the SKA. The planned deep continuum survey of MeerKAT is the MeerKAT International Giga-Hertz Tiered Extragalactic Exploration (MIGHTEE) survey (Jarvis et al. 2018). This survey will make use of MeerKAT's high resolution and fast survey speed to conduct a tiered survey at 1.4GHz, over several well-studied fields, totaling 20 square degrees to μJy sensitivities and an ultra-deep image of a single 1 square degree field. The MIGHTEE project was designed to pathfind SKA key science in cosmology and galaxy evolution. Thus, MIGHTEE will explore dark matter and large scale structure, the evolution of galaxies, including AGN activity and star formation as a function of cosmic time and environment, the emergence and evolution of magnetic fields in galaxies, and the magnetic counterpart to large scale structure of the Universe.

• SKA- The Square Kilometre Array project is an international effort to build the world's largest radio telescope, with a square kilometre of collecting area. Majority of the SKA will be hosted in South Africa, with the remainder located in Australia. In Australia, the SKA low-frequency telescope will comprise 512 stations in a large core

and three spiral arms creating a maximum baseline of 65km. Each of the stations will contain ∼250 individual antennas, thus a resulting ∼130,000 will be installed in total. In South Africa, 133 antennas will be added to the existing the 64-dish MeerKAT precursor telescope, forming an array of nearly 200 dishes. Some of the antennas will be used to form three spiral arms with a maximum baseline of 150km. The ultimate goal is to expand the SKA to 10 times this size, with a million low-frequency antennas in Australia and ∼2000 high- and mid-frequency dishes and aperture arrays extending into the African partner countries. Early science observations are expected to start in the mid-2020s with a partial array. The unprecedented sensitivity of the SKA telescope will allow for studies of the formation and evolution of the first stars and galaxies, will shed light on dark energy and dark matter and enable tests of gravity using pulsars and black holes.

In summary, 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. New telescopes are being built while others such as the VLA have been upgraded. LOFAR, for example, has begun observations of the entire northern hemisphere ($\delta < +30^{\circ}$). As part of a three-tiered survey, LOFAR will carry out a 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 first of which is the LOFAR Two-metre Sky Survey (LoTSS - Shimwell et al. 2017) is a 120-168 MHz survey of the entire northern sky, with a sensitivity of 100 μJy/beam. The combination of EMU and WODAN will observe the entire sky, producing a catalogue of almost 100 million sources. Discussion of the applications of some of these surveys to cosmology is given by Raccanelli et al. (2012). The analysis includes simulated catalogues which emulate the planned surveys and are used in order to model possible source distributions and the respective science achievable with these surveys, focusing on three main cosmological probes: The angular two-point correlation of radio sources, the late-time Integrated Sachs-Wolfe effect and the cosmic magnification bias. In work by Raccanelli et al. (2012), it is assumed that no redshift information is available for individual radio sources. Work by Cam-

era et al. (2012) investigates the resulting improvement in cosmological measurements when redshift information is included.

Science teams using the aforementioned surveys will ensure a maximum scientific return by combining the deep radio continuum data with the best multi-wavelength data sets from X-ray through to far-infrared. In Figure 1.12 taken from Jarvis (2012), a comparison of existing and near-future radio surveys, in terms of sky coverage and limiting flux density, is shown. The ideal survey would sit at the bottom right corner of this plot, having both a very large sky coverage and very low limiting flux density. It is shown that definitive improvement in both depth and sky coverage will be made with the near-future surveys, with narrow, but very deep surveys such as the MeerKAT MIGHTEE Tier-3 survey and very large, relatively shallow surveys such as EMU and WODAN. The SKA surveys will push closer to the bottom right of this plot, probing a new, unexplored volume of the Universe.

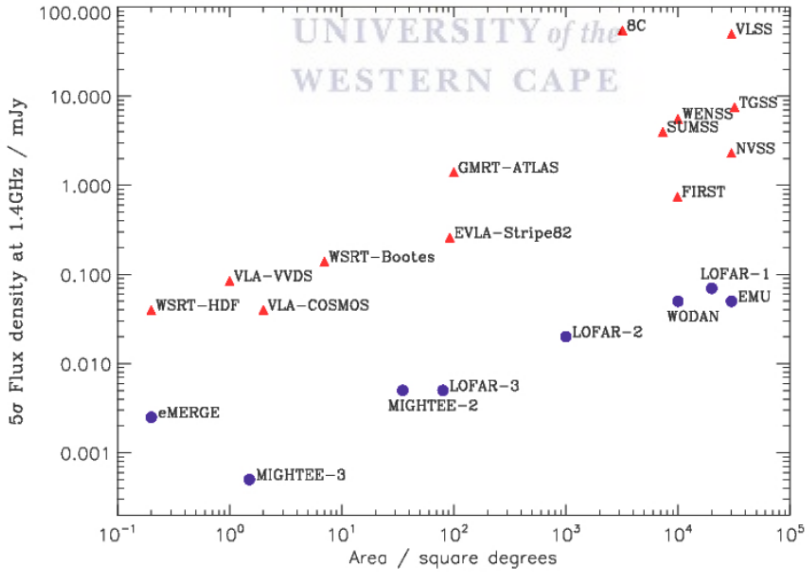


Figure 1.12: Comparison of existing and near-future 1.4GHz surveys in terms of sky coverage and limiting flux density, where triangles represent past and existing surveys and filled circles represent surveys which will be underway in the near future. Image taken from the work of Jarvis (2012).

1.4 Previous Measurements of Large-Scale Clustering at Radio Wavelengths

Radio observations are unaffected by dust extinction and relatively, less effected by our atmosphere, therefore we are able to observe any region of the sky relatively unobscured and therefore obtain reliably complete catalogs over large areas of sky. Radio sources are also detected over wide areas of sky and out to high redshifts (z$\sim$6) and therefore sample much larger volumes of space than optical surveys. Radio galaxies also span a very broad range in redshift, which allows for the study of the evolution of clustering over cosmic time. However, this is also a disadvantage if redshift information is unavailable. The broad redshift distribution of radio surveys makes the detection of clustering difficult, as many "close" pairs of sources are actually widely separated in the line of sight direction, thus the projection on the sky of the observed sources can result in a distribution which is almost random and therefore smearing out information on the large-scale structure. Thus, measuring the faint clustering imprint in the radio sky was only made possible with the advent of wide-area radio surveys of mJy depth, such as the 325 MHz Westerbork Northern Sky Survey (WENSS - Rengelink et al. 1997), the 1.4 GHz VLA surveys, such as the Faint Images of the Radio Sky at 20 cm (FIRST - Becker et al. 1995) survey and NVSS (Condon et al. 1998), as well as the Sydney University Molonglo Sky Survey (SUMSS- Bock et al. (1999)). Kooiman et al. (1995) were able to detect marginally positive ACFs in their analysis of 54,579 sources with S $>$ 35mJy from the Green Bank 4.85 GHz survey. Marginal detections were also made by Loan et al. (1997) using data from the Parkes-MIT-NRAO survey. However, the first significant detection of the clustering of radio galaxies was measured by Cress et al. (1996) by analysis of $\sim$138,000 sources from the VLA FIRST survey. Cress et al. (1996) inferred a correlation length of $r_0 \sim 10h^{-1}$Mpc, confirming that even this faint radio population was more strongly clustered than the previously mentioned optically selected samples. Further analyses of this data set were carried out by Magliocchetti et al. (1998), Blake & Wall (2002a) and Overzier et al. (2003). Strong detections of clustering in the WENSS and SUMSS data were found by Rengelink & et al. (1998) and Blake et al. (2004), respectively.

The use of data at other wavelengths in order to obtain spectroscopic or photometric information allowed for accurate constraints of the clustering properties of these radio samples. Magliocchetti et al. (2004) perform a clustering analysis of 820 objects from FIRST, with additional information from 2dFGRS, this yielded a $S_{1.4\mathrm{GHz}} > 1$ mJy sample at z < 0.3 and find a correlation length of $r_0 \sim 6.7h^{-1}$Mpc for the full sample. For a sample of 536 AGN, they measure $r_0 \sim 10.9h^{-1}$Mpc. A similar analysis by Brand et al. (2005) of $S_{1.4\mathrm{GHz}} > 3$ mJy sources from the Texas-Oxford NVSS Structure (TONS) survey find a correlation length of $r_0 \sim 8.7h^{-1}$Mpc for z < 0.5 sample. Wake et al. (2008) measure a high correlation length of $r_0 \sim 12.3h^{-1}$Mpc for a sample of radio-loud luminous red galaxies (LRGs) at $z \sim 0.55$ from the 2SLAQ LRG survey. These results show the clustering strength of radio galaxies to be high relative to Milky Way type galaxies ($r_0 \sim 5h^{-1}$Mpc), however the bright limits of the optical sources used biases the studied samples towards low redshift radio sources.

In more recent years, deep, sub-mJy radio surveys carried out over smaller but well studied fields in which the radio sources have a wealth of multi-wavelength information, have opened a new era for the direct investigation of the environmental and clustering properties of radio sources to higher redshifts. In these cases, photometric or in some cases spectroscopic redshift information is acquired for the radio sources. A few clustering analyses have made use of such surveys in order achieve more accurate measurements of clustering. Lindsay et al. (2014a) find strong evolution in the clustering and bias of faint radio sources from the analysis of FIRST sources over the Galaxy And Mass Assembly (GAMA - Driver et al. 2011) survey area having a flux density limit of $S_{1.4\mathrm{GHz}} > 1$ mJy. A correlation length of $r_0 \sim 8.2h^{-1}$Mpc is found for the sample, as well as an evolution of r_0 from $\sim 6 - 14h^{-1}$Mpc between redshift $z = 0.3$ and $z = 1.55$. Magliocchetti et al. (2017) perform a clustering analysis of a sample of $S_{1.4\mathrm{GHz}} > 0.15$ mJy AGN and SFGs from a VLA survey over the COSMOS field. They find values of $r_0 \sim 6.8h^{-1}$Mpc and $r_0 \sim 11.2h^{-1}$Mpc for SFGs and AGN sources, respectively. Hale et al. (2018) perform a similar analysis of a 3 GHz VLA sample over the COSMOS field and measure correlation length values of $r_0 \sim 6.1h^{-1}$Mpc and $r_0 \sim 9.6h^{-1}$Mpc

for SFGs and AGN sources, respectively.

All the aforementioned works converge at indicating that radio sources are a very strongly clustered population. This suggests that they inhabit more massive dark matter haloes and are thus excellent traces of the underlying dark matter distribution. Therefore, studying the distribution of various radio source populations should provide useful measurements of the galaxy bias and how the bias evolves over cosmic time.

1.5 Aims of the thesis

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 provide information about the baryonic matter, from which we must infer the dark matter distribution. Radio galaxies have been shown to be more massive and therefore good tracers of the underlying dark matter, however very few clustering studies have been carried out at sub-mJy levels. At these depths the radio source population is less dominated by active galactic nuclei (AGN), and a greater fraction of star-forming galaxies is observed. However, the lower flux density limit also allows for the detection of AGN beyond z∼1, at which the bias of radio sources is poorly understood. It is advantageous to investigate the bias of radio sources out to higher redshift, as we are then able to learn about the evolution in clustering of the radio source population. The purpose of this Thesis is to investigate the clustering properties of the sub-mJy radio source population and how the clustering evolves over cosmic time, as well as to determine the clustering properties of the various radio galaxy sub-populations. I will also investigate the evolution of the galaxy bias $b(z)$ with redshift for the various radio samples.

In Chapter 2 I perform a clustering analysis of a $S_{1.4\mathrm{GHz}} > 440$ μJy sample of 11,431 sources from a JVLA survey over the SDSS Stripe 82 region. The analysis is performed for the full sample of radio sources for completeness, and to allow for comparison with previous clustering analyses. A clustering analysis is then performed

for a subset of the radio sources which have been cross-matched with optical sources from the SDSS, allowing for sources to be split into bins in redshift. Studying the clustering of sources in each redshift bin therefore allows for the investigation of the evolution of clustering with redshift. In Chapter 3, the sources from the cross-matched sample are split by radio luminosity, allowing for a study of the clustering properties of the low and high $S_{1.4\mathrm{GHz}}$ luminosity populations, which in turn provides insight into the types of sources which make up the respective radio luminosity populations. A clustering analysis of $S_{1.4\mathrm{GHz}} > 440$ μJy sources from the S^3 simulation (Wilman et al. 2008), at the same luminosity cuts, is performed for comparison. In Chapter 4 a clustering analysis is performed for samples of the JVLA Stripe 82 sources which have been cross-matched with optical SDSS sources via visual inspection in order to obtain reliable host galaxy matches and thus, spectroscopic information. This allows for the investigation of the clustering properties of sources which have spectroscopic matches and have been classified as either AGN, star-forming galaxies, HERGs and LERGs, providing insight into how these various radio sub-populations cluster, how biased each sample is and how the bias evolves with cosmic time. In Chapter 5 I present a summary of the results and highlight future work that would increase our understanding of the way in which radio sources trace the large-scale structure of the Universe.

Throughout this Thesis I assume a ΛCDM cosmology with $H_0 = 100h$ kms^{-1} Mpc^{-1}, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $\sigma_8 = 0.8$. For all clustering measurements we leave our values in terms of h^{-1}, where $h = 0.7$.

Chapter 2

Angular correlation function analyses & the evolution of bias with redshift

In this chapter I investigate the large-scale clustering of radio galaxies by measuring the two-point angular correlation function $\omega(\theta)$ using a new radio survey over the Sloan Digital Sky Survey (SDSS) Stripe 82 region. This is measured for the full radio sample, as well as for a sample which is cross-matched with sources from deep, coadded SDSS data from the Stripe 82 survey. After obtaining redshift information from the host galaxies, the sources with trusted optical matches are split into three bins in redshift and an angular correlation function analysis is performed for each, allowing for an investigation of the evolution of clustering with redshift. From the angular correlation function, one is able to infer the spatial clustering of the sample, given an estimate of the sample's redshift distribution N(z). The galaxy bias b(z) is then calculated for each subsample and its evolution with redshift is investigated. The results are also compared with bias evolution models, simulated by the SKADS S^3 semi-empirical simulation, for various types of radio source populations.

2.1 Data

2.1.1 JVLA

In recent years there have been many deep continuum surveys carried out using the Karl G. Jansky Very Large Array (JVLA) telescope. Examples are a survey of The Chandra Deep Field South (Padovani et al. 2011) and a 3GHz observations over the Cosmic Evolution Survey (COSMOS) field (Smolčić et al. 2017). In the past, wider, more shallow surveys such as Faint Images of the Radio Sky at Twenty-Centimeters (FIRST- Becker et al. 1995) and NRAO VLA Sky Survey (NVSS- Condon et al. 1998). In order to bridge this gap, a medium-deep survey over ~ 100 deg^2 was carried out by Heywood et al. (2016) using the JVLA telescope and this is the dataset used for this Thesis.

The survey was carried out at a centre frequency of 1.4 GHz over the SDSS Southern Equatorial Stripe (Stripe 82) region (Heywood et al. 2016). The survey was carried out with the array in a compact, hybrid CnB configuration yielding an angular resolution of 16x10 arcseconds and a mosaic image with an effective 1σ noise level of 88 μJy/beam (See Figure 2.1). The survey is composed of 1,026 snapshot observations, each with a duration of 2.5 minutes. The hybrid CnB configuration allows for a survey which is sensitive to diffuse, low surface brightness structures and extended radio emission, which allows for the detection of low luminosity lobes from AGN. The Stripe 82 field was previously observed in a survey conducted by Hodge et al. (2011), which was carried out with the JVLA in A-configuration with supplemental B-configuration data, yielding a high-resolution (1.8 arcsecond) sample. The SDSS Stripe 82 field is also rich in complimentary multi-wavelength data. The Heywood et al. (2016) survey is split into an eastern and western field, each having approximately equal sky area. A horizontal gap exits in the western field, spanning the whole range in RA, and is due to unobserved scheduling blocks. For this Thesis, I use a 5σ ($S_{1.4} > 440\mu$Jy) catalogue consisting of 5,490 and 5,941 sources for the eastern and western strips, respectively. For details of the observations, data reduction and catalogue generation see Heywood et al. (2016).

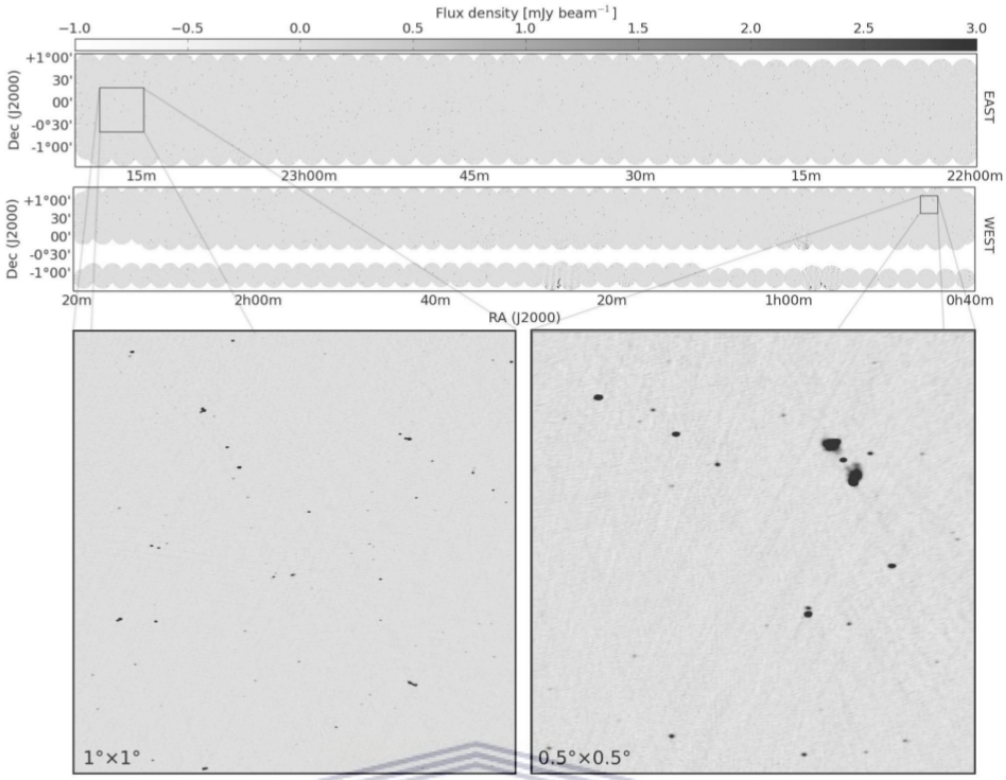


Figure 2.1: Total intensity mosaics of the JVLA Stripe 82 survey for the eastern and western regions taken from Heywood et al. (2016). The zoomed-in regions are $1° \times 1°$ and $0.5° \times 0.5°$ in size for the eastern and western fields respectively.

2.1.2 SKADS Simulated Skies

The mock data used in this work is provided by the SKA Design Study (SKADS) initiative, in particular, the SKADS Simulated Skies Semi-Empirical eXtragalactic (S^3-SEX; Wilman et al. 2008), which is a simulation of extragalactic radio continuum sources. This is a semi-empirical simulation, that is, the sources simulated were derived from observed and extrapolated luminosity functions and distributed according to an underlying dark matter density field, with biases corresponding to their measured large-scale clustering properties. This was developed in order to estimate the extragalactic radio continuum and line-emitting data we might obtain with the SKA, with an emphasis on modeling the large-scale distribution of radio sources in the Universe. The simulation covers a sky area of 20×20 deg^2 out to a redshift of $z = 20$, and down to a flux density limit of 10 nJy at five frequencies in the range 151 MHz - 18 GHz, as well as providing K-band magnitudes and classifications of 5 different radio source types

(Wilman et al. 2010): star-forming galaxies, starburst galaxies, 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, as well as the Mo & White (1996) bias model. In this Thesis, the S^3 catalogue is used to obtain approximate redshift distributions of radio samples, at a given limiting flux density, which then allows us to obtain information regarding the spatial/three-dimensional clustering of the sample. The simulated redshift distributions are found to agree with deep field radio observations down to a flux density limit of 100 μJy (Simpson et al. 2012; Smolcic et al. 2015).

2.1.3 Random Source Catalogue

As described in Section 1.2.1, in order to measure the two-point angular correlation function, the distribution of the radio data sample is compared to that of a random source sample. It is important that the random sample generated contain the same characteristics/biases as the data set. The random data is generated over exactly the same areal coverage as the survey data. However, the random sample must also be masked according to the noise distribution of the field. The flux density threshold for detection of a source depends on the local rms noise, which varies slightly across the survey area. This must therefore be taken into account when distributing the random sources, as a random source cannot be placed at a location at which a real source, of the same flux, would fall below the flux density limit. To mask the random sample to account for noise, I assign each source a random location within the survey area and then assign a flux (chosen randomly from SKADS sources with $S_{1.4} > 50\mu$Jy) to each of these mock sources. For each source, I then obtain the rms noise at the assigned location. This noise value is used to create a Gaussian probability, from which the noise value is randomly sampled. The noise is then added to the randomly chosen SKADS flux, and the mock source is discarded if the resulting new flux falls below our flux cut of 440μJy. In Figure 2.2, the distribution of the masked random sample is shown in relation to the sky coverage of the real sources. It is evident, particularly from the fringes of the observed fields, that random sources have not been placed in positions at which real sources would not be detected due to the associated noise level.

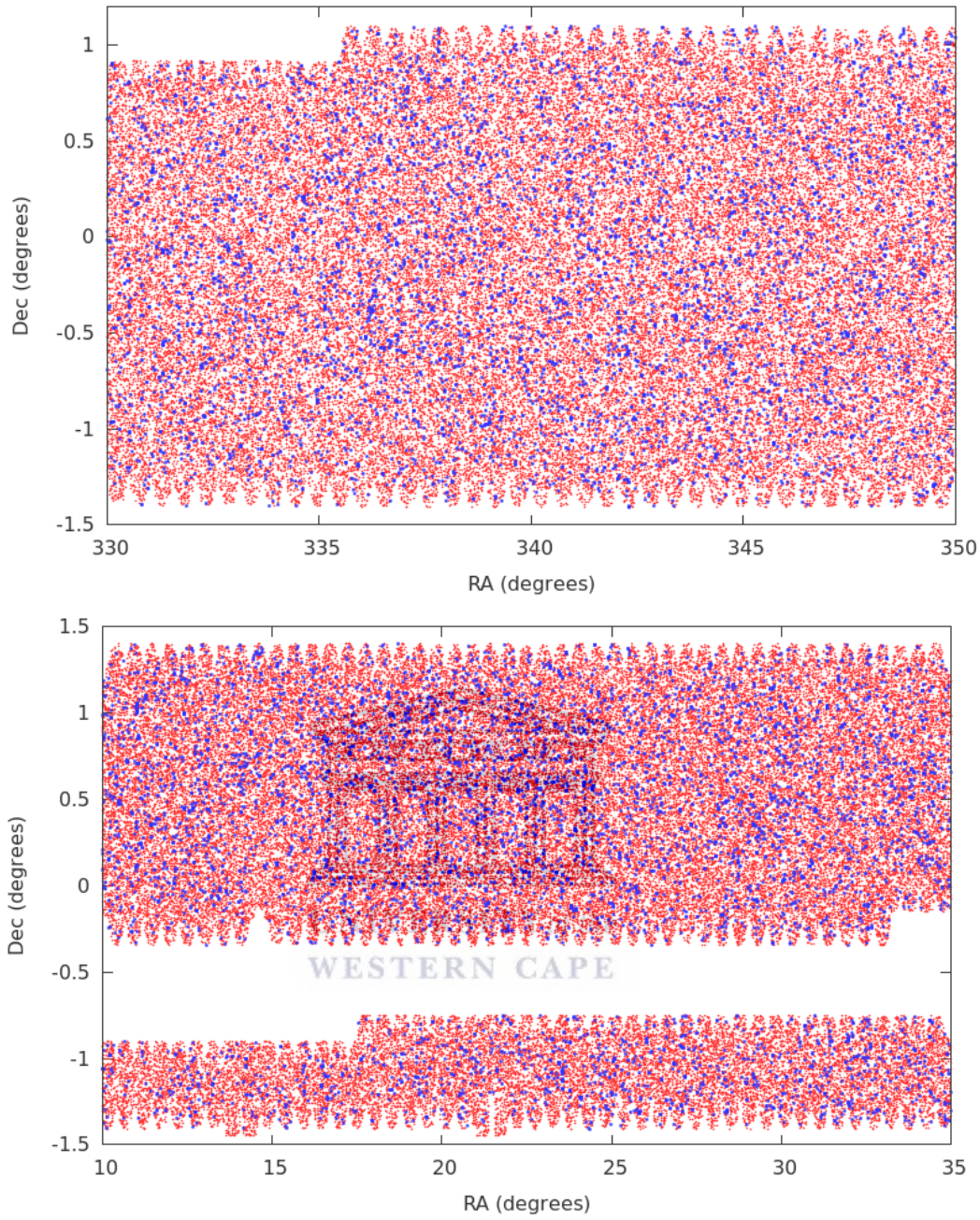


Figure 2.2: The distribution of the randomly generated sample (red) masked according to the properties of the data in relation to the distribution of the real sources (blue), in both the eastern (top) and western (bottom) fields of the survey.

2.2 Measured ACFs

The two-point angular correlation function ($\omega(\theta)$) is calculated using the Landy-Szalay estimator presented in Section 1.2.1. When implementing the Jackknife method (also

presented in Section 1.2.1) to calculate the $\omega(\theta)$ errors, the sample is split into 50 bins in RA. Theory predicts power law distributions for the clustering of dark matter, and hence for galaxies over a broad range of scales (Peebles 1973). However, the two-point correlation function of a sample of radio galaxies usually follows a double power law form, i.e. the sum of two power laws, each of which dominates over different angular scales (Blake & Wall 2002a). This is written as $\omega(\theta) = A\theta^{\eta} + B\theta^{\kappa} - \sigma^2$, where A and B are the clustering amplitudes of the large- and small-scale power laws, respectively. The slope of the large- and small-angle power laws are equal to $\eta = 1 - \gamma_1$ and $\kappa = 1 - \gamma_2$, respectively and σ^2 is the integral constraint, described in section 1.2.1. However, the power law which dominates over small angular scales does not represent true cosmological clustering. In high-resolution radio surveys, extended sources such as FR-I and FR-II galaxies (Fanaroff & Riley 1974) are resolved into their multiple radio components, resulting in some single sources being counted as two or more sources. We therefore observe an upturn in the angular correlation function at small angular scales. Blake & Wall (2002a) and Overzier et al. (2003) find that even a small number of multi-component sources can greatly affect the ACF on scales of up to several arcminutes. The issue of components of extended radio galaxies being recognised as individual sources is to some extent mitigated by the source finding algorithm used by Heywood et al. (2016) as follows: The *Python Blob Detection and Source Finder* (PyBDSF) source finder (Mohan & Rafferty 2015) first estimates the spatial variation of the image noise by moving overlapping boxes across the image, in which the root mean square (rms) of all pixels are calculated and values from each step are interpolated. The boxes are 250×250 pixels in size and are moved across the image in steps of 50 pixels. The detection threshold is thereby determined and positions which exceed 5σ, where σ is the local value of the rms noise, are identified. These peaks are then grown into islands, i.e. regions where the emission exceeds 3σ. The algorithm will then declare components to be of the same source if two criteria are met: (i) The pixels along the line which join the centres of two components must all be above the 3σ island threshold. (ii) The length of the line joining the component centres must be less than the sum of the full width half maxima of the two components. See Heywood et al. (2016) for further information regarding the catalogue generation. However, we

still fit two separate power laws, over different angular scales, in order to obtain a better measurement of the large scale clustering amplitude and η, as this is what we are interested in. From this point onwards, when I refer to γ it will be that of the large-angle power law.

The measured angular two-point correlation functions for all sources in the eastern and western survey fields, respectively, are presented in Figures 2.3 and 2.4. We see the angular correlation function following the double power law form as expected, shown by the blue curve, with an upturn around theta $\sim 0.1^{\circ}$, consistent with the upturn found by Blake & Wall (2002b). The Jackknife errors are seen to increase at theta $\sim 2.0^{\circ}$, this is due to the sky coverage of the sample being in the form of a strip, which is about 2° wide in declination. Thus, the number of pair counts at angular scales larger than this is deficient. The grey dashed lines in Figures 2.3 and 2.4 denote the negative values of the angular correlation function, which were omitted when fitting the small- and large-scale power laws. The large angle power law, which measures the clustering of galaxies, is fitted for $\theta > 0.1$, as this is the angular separation above which we can be confident that $\omega(\theta)$ measures the cosmological clustering of individual galaxies and does not include clustering power from large, extended radio sources that have been resolved into more than one source. The chosen scale is motivated by studies from Lara et al. (2001) on the occurrence and size of giant radio galaxies, which shows that some AGN extend up to scales of several arcminutes. This best fit, using the χ^2_{min} statistic, is shown by the cyan coloured line, the parameters of which are presented in Table 2.1 along with corresponding 1σ uncertainties. These uncertainties are calculated by determining the probability distribution of the respective χ^2 sampling and then integrating under the curve to find the parameter value corresponding to an area of 34% on either side of the best fit parameter value. The best fit parameters to the large-angle power law differ slightly for the measured ACFs of the eastern and western fields, with clustering amplitudes of $A = 3.91^{+1.85}_{-1.54} \times 10^{-3}$ and $A = 9.82^{+7.28}_{-3.44} \times 10^{-3}$ measured for the western and eastern fields, respectively. These results are still consistent within the uncertainties at the 1.5σ level. The measured γ values, i.e. 1-η, of $\gamma = 1.54^{+0.24}_{-0.26}$ and $\gamma = 1.98^{+0.20}_{-0.22}$ for the eastern and western fields, respectively, are also consistent

within the uncertainties, and consistent with the commonly found value of ∼1.8 for a variety of galaxy classes (Bahcall & Soneira 1983; Blake & Wall 2002a).

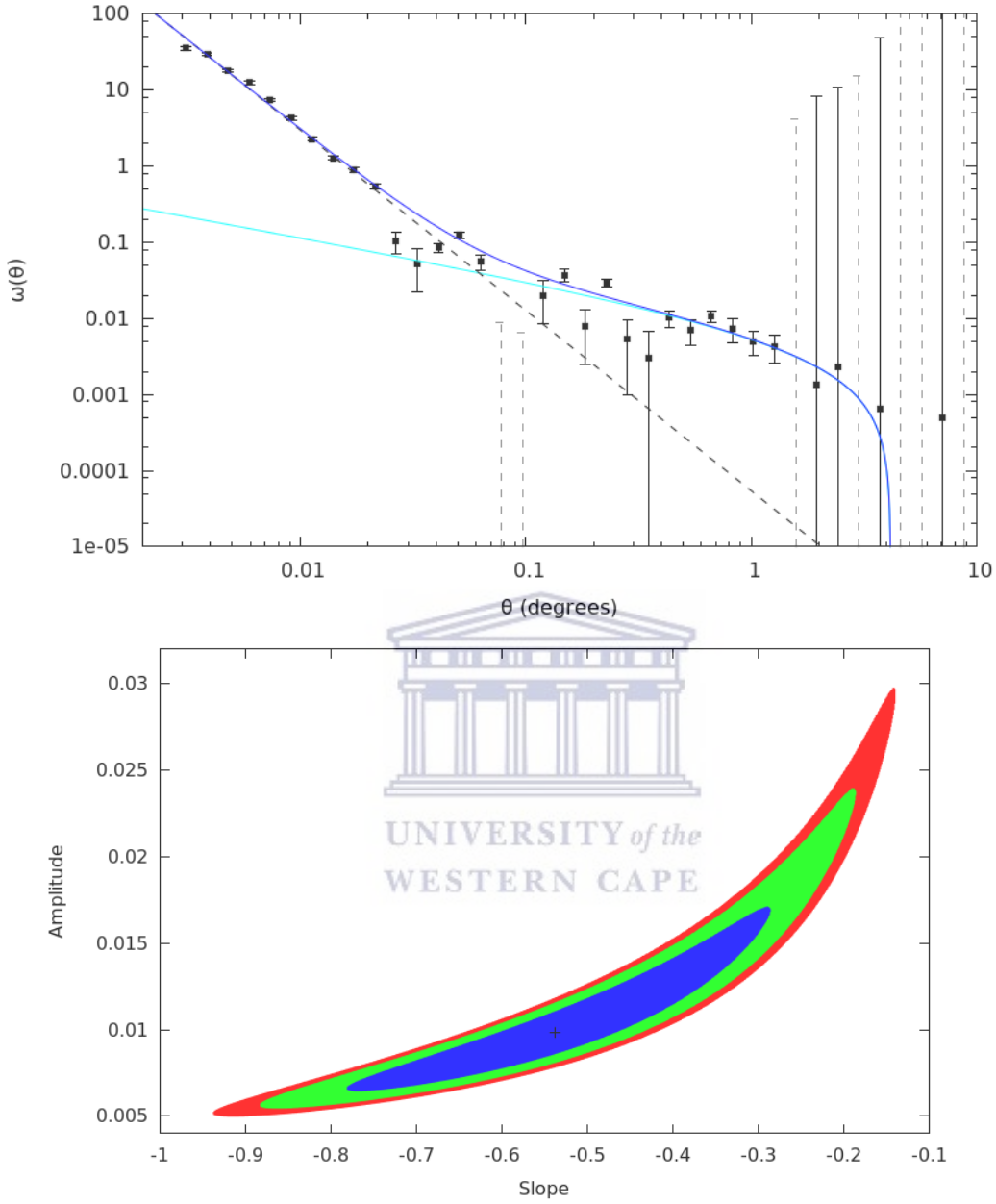


Figure 2.3: Top: Full angular two-point correlation function of sources in the eastern field of the survey. The cyan coloured line and dashed black line indicate the best fit power-laws to the data, which dominate over large ($\theta > 0.1°$) and small angular scales, respectively. The blue line depicts the best fit double power-law, i.e. the sum of the small and large angle power laws minus an integral constraint σ^2 and the grey dashed lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the large-angle single power-law fits.

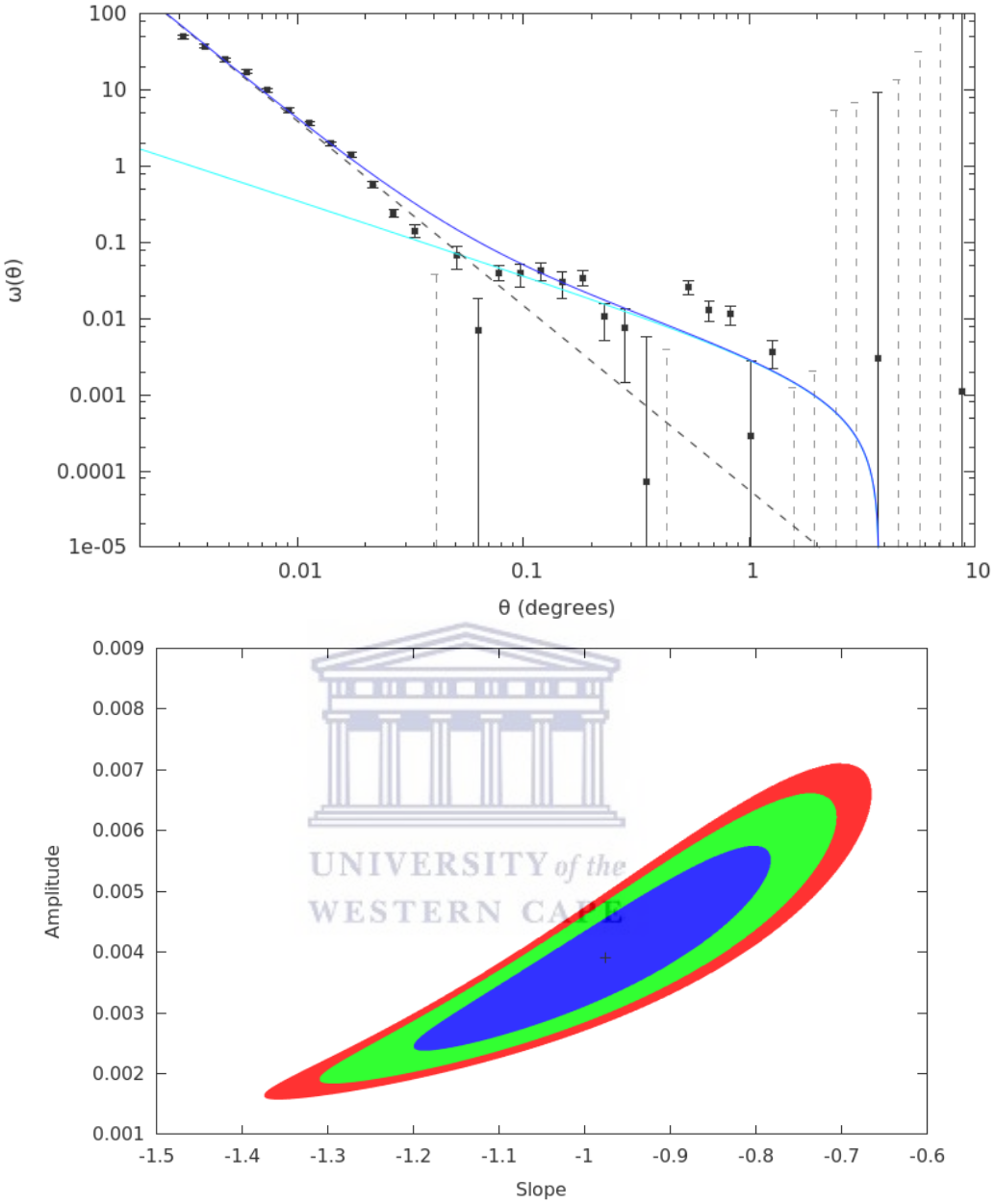


Figure 2.4: Top: Full angular two-point correlation function of sources in the western field of the survey. The cyan coloured line and dashed black line indicate the best fit power-laws to the data, which dominate over large ($\theta > 0.1°$) and small angular scales, respectively. The blue line depicts the best fit double power-law, i.e. the sum of the small and large angle power laws minus an integral constraint σ^2 and the grey dashed lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the large-angle single power-law fits.

2.2.1 Collapsing Technique

An alternative to fitting a double power law to account for the contribution of multicomponent radio sources to the TPCF, is to mitigate this contribution altogether. The approach was proposed by Cress et al. (1996), whereby all sources within 0.02° of each other are collapsed into a single source, as this is the scale at which a break in the power law was observed (Cress et al. 1996). The collapsed sources will have a new location at the flux weighted average of the individual positions of the sources in the respective group. By applying the collapsing technique to the data, as well as the random sample, and calculating the ACF, the double power-law form of the angular TPCF is mitigated, i.e. there is no longer an upturn. Using this technique I fit a single power-law to the ACF over the theta range $0.02° < \theta < 15°$. The angular correlation functions of the collapsed samples, are shown in Figures 2.5 and 2.6 for the eastern and western fields, respectively. The best fit parameter values for the ACFs of the collapsed samples are presented in Table 2.1 and will be discussed in the following section.

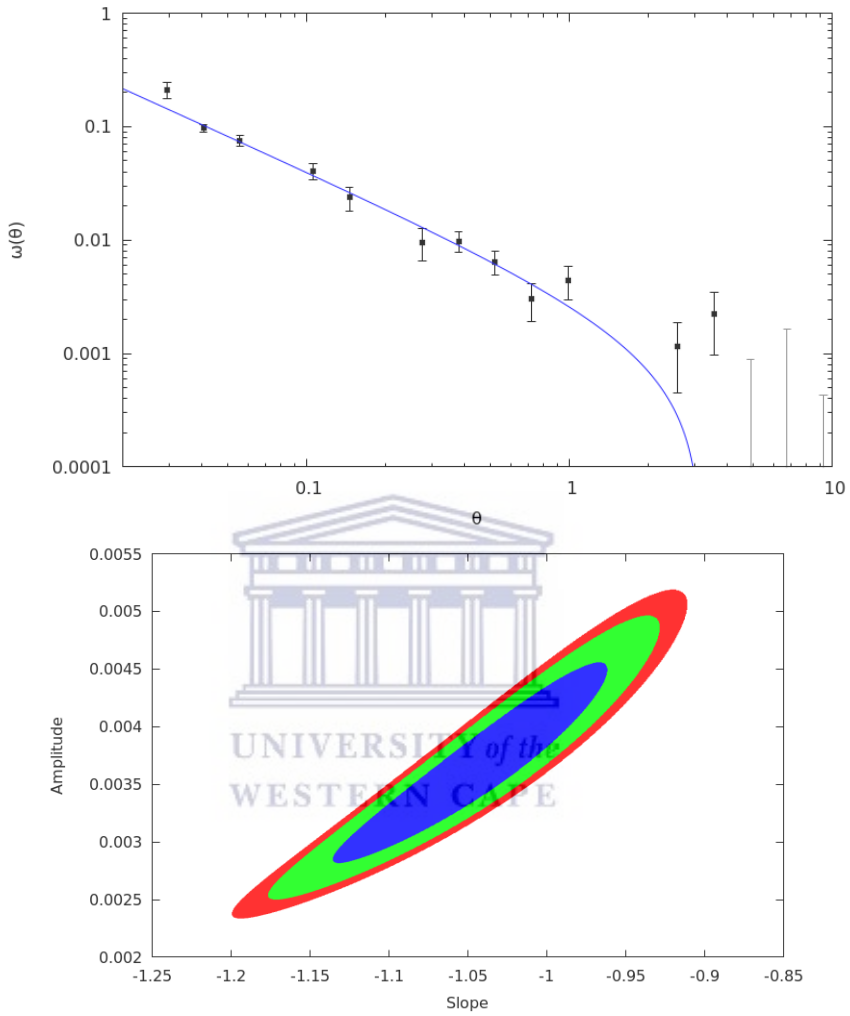


Figure 2.5: Angular two-point correlation function of sources in the eastern field after the collapsing technique is applied. The blue line indicates the best fit power-law to the data, minus an integral constraint σ^2 and the grey lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the large-angle single power-law fits.

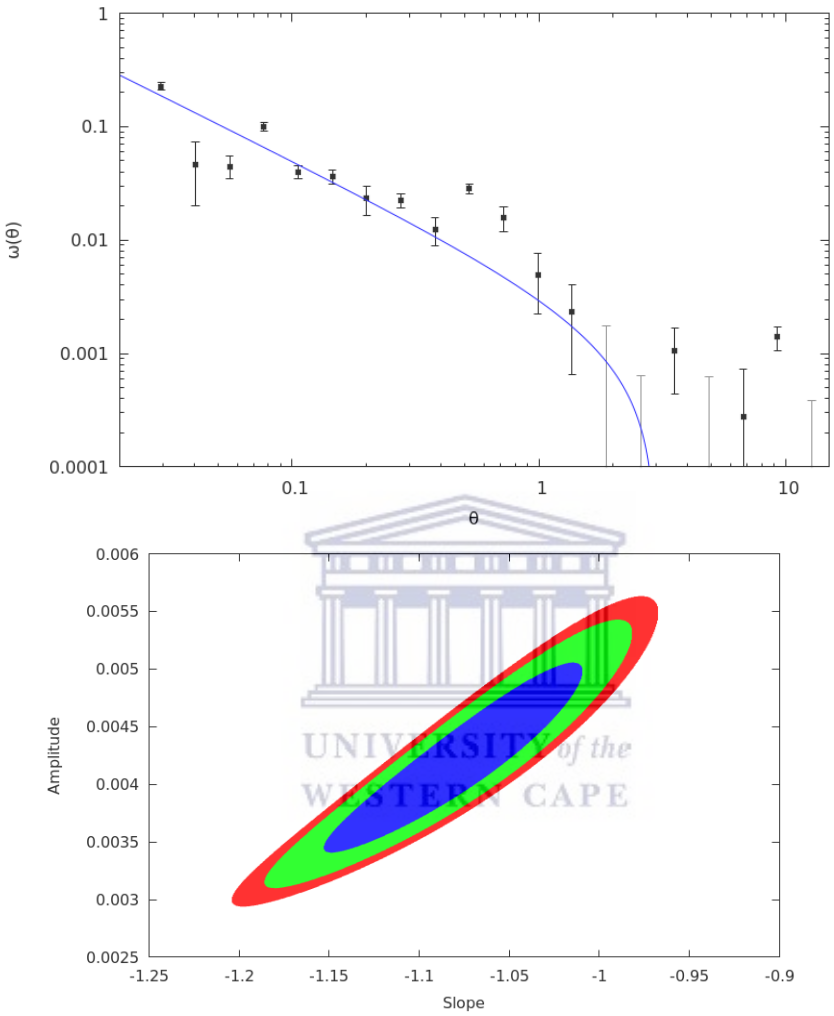


Figure 2.6: Angular two-point correlation function of sources in the western field after the collapsing technique is applied. The blue line indicates the best fit power-law to the data, minus an integral constraint σ^2 and the grey lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the large-angle single power-law fits.

2.2.2 Clustering Results & Discussion

Using the best-fit parameters of the measured ACFs, we then infer the spatial clustering and calculate the correlation length r_0 as well as the galaxy bias. As described in Section 1.2.1, the correlation length r_0 represents the separation length of the correlation function at which $\xi(r) = 1$ and the galaxy bias is defined as the ratio of the spatial clustering of a galaxy sample to that of the dark matter halos. In order to calculate these values, knowledge of the redshift distribution of the sample is required. To obtain this for the full JVLA sample, I use data from the S^3 semi-empirical simulation and impose a flux density cut of $S_{1.4} > 440\mu$Jy, as was done to obtain the JVLA radio sample used for the analysis. This yields a SKADS sample of 68,334 mock sources with a median redshift of z=1.21. This SKADS sample is then used to provide an estimate of the redshift distribution $N(z)$ of our radio sample, which is shown in Figure 2.7.

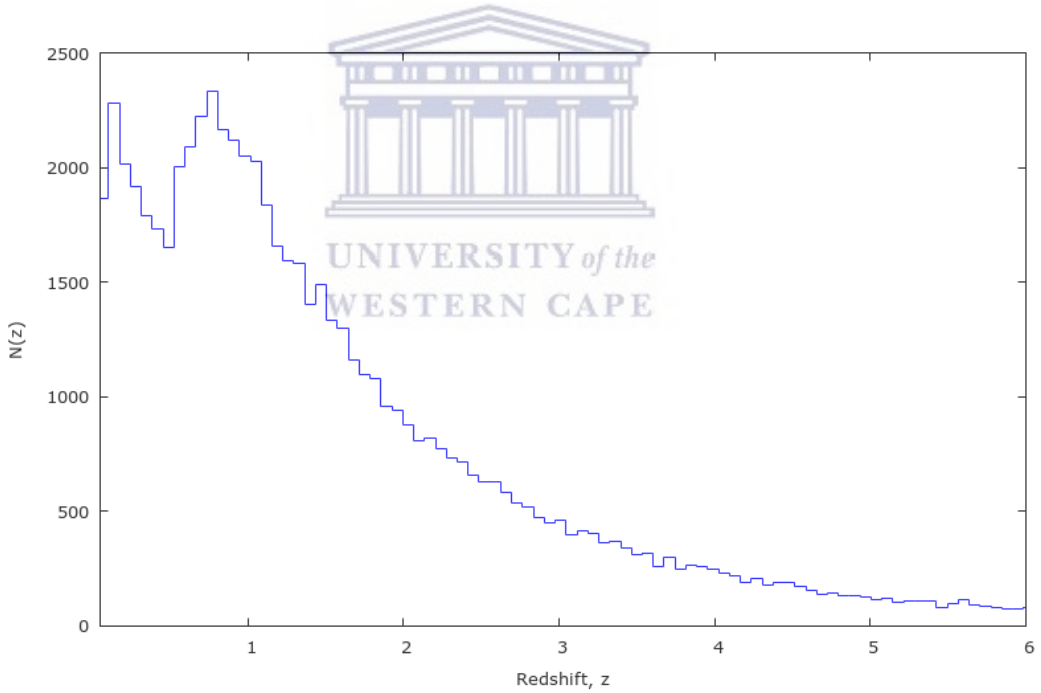


Figure 2.7: The redshift distribution of sources from the SKADS Simulated Skies, which have a flux density limit of $S_{1.4} > 440\mu$Jy.

When comparing the best-fit parameter results of the ACFs of both the eastern and western fields to that of previous measurements, we see that with regard to the ACF for the eastern field, the clustering amplitude is slightly higher than found in previous ACF analyses of $S_{1.4\text{GHz}} > 1$ mJy samples from the FIRST survey ($A = 4.87^{0.17}_{0.12} \times 10^{-3}$

- Lindsay et al. 2014b) and NVSS ($A = 1.04 \pm 0.09 \times 10^{-3}$ - Blake & Wall 2002b). This can be explained by two possible influences: (i) the probable lowering of the average redshift of the sources in the sample due to the expected large contribution of nearby faint star-forming galaxies in a sample at sub-mJy flux densities (Windhorst et al. 1985; Condon 1984). This is supported by the redshift distribution of samples with different flux-density cuts from the SKADS extragalactic continuum simulation. As shown in Figure 2.8, the normalised redshift distribution of the $S_{1.4} > 440\mu$Jy SKADS sample shows a much larger population of radio sources at low redshift (z < 0.5) than that of the $S_{1.4} > 1$ mJy SKADS sample. At lower redshifts, a given angle corresponds to a smaller physical scale, at which there is more clustering. (ii) The increased sensitivity of the survey allows for the detection of a larger number of powerful, but high-redshift AGN which generally reside in massive elliptical hosts (Jarvis et al. 2001; McLure et al. 2004; Seymour et al. 2007; Herbert et al. 2011) and in rich cluster environments (Mandelbaum et al. 2009; Donoso et al. 2014), thus yielding a higher clustering amplitude. The slope of 0.54 is more shallow than the previously measured values of $\geq$0.8, however is acceptable within the uncertainties with a 1σ upper-limit of -0.78. The best-fit parameters of the ACF measured for the western field of the survey show a slightly smaller clustering amplitude than that of the eastern ACF, which is closer to that found in ACF analyses for samples at the 1mJy level. The larger slope of -0.98 is in agreement with previous clustering studies in the radio ($\gamma = 1.85 \pm 0.1$ - Blake & Wall 2002b; $\gamma = 1.84 \pm 0.02$ Lindsay et al. 2014a).

Table 2.1: Best-fit parameters of the large-angle power law from the two-point angular correlation functions presented, correlation length r_0 and galaxy bias b(z) values. All uncertainties quoted are 1σ, obtained from χ^2_{min} fitting.

Sample	N_{obj}	A ($\times 10^{-3}$)	γ	χ^2_{red}	r_0 (h^{-1}Mpc)	b(z_{med}=1.21)
Full East	5,356	$9.82^{+7.28}_{-3.44}$	$1.54^{+0.24}_{-0.26}$	1.18	$15.67^{+10.30}_{-8.93}$	$2.45^{+2.10}_{-1.23}$
Full West	5,893	$3.91^{+1.85}_{-1.54}$	$1.98^{+0.20}_{-0.22}$	0.80	$12.68^{+3.30}_{-4.12}$	$2.74^{+1.51}_{-1.05}$
Collapsed East	3,846	$3.64^{+0.91}_{-0.83}$	$2.45^{+0.09}_{-0.08}$	1.52	$12.54^{+1.70}_{-1.89}$	$2.86^{+0.67}_{-0.57}$
Collapsed West	4,143	$4.20^{+0.90}_{-0.80}$	$2.78^{+0.08}_{-0.07}$	2.56	$13.56^{+1.50}_{-1.59}$	$3.18^{+0.62}_{-0.53}$

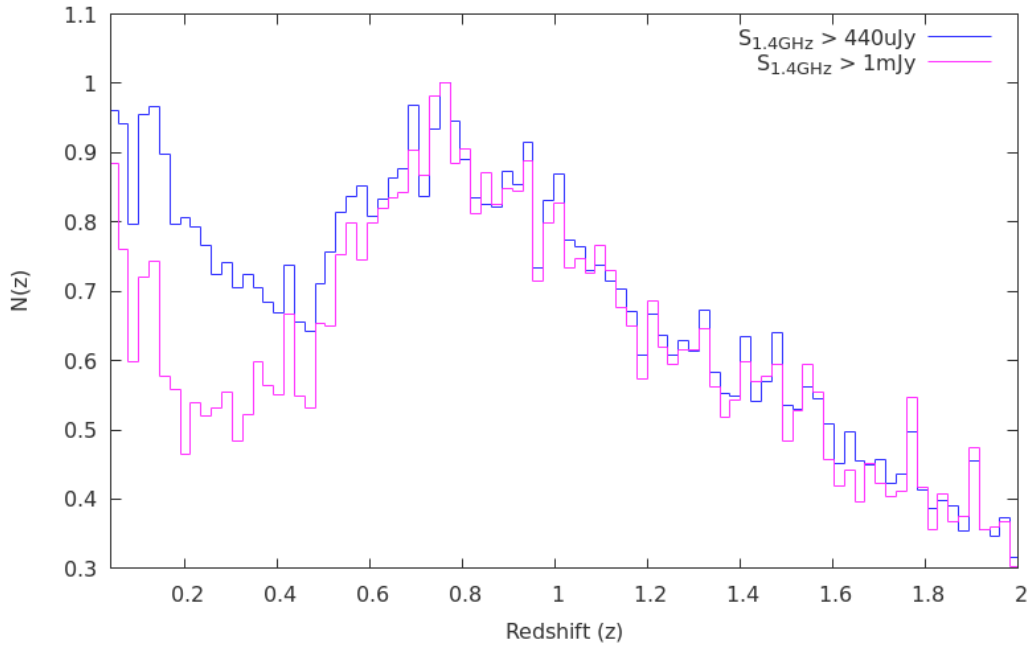


Figure 2.8: A comparison of the normalised redshift distribution of source samples from the SKADS Simulated Skies having a $S_{1.4} > 440\mu$Jy flux density limit (blue) versus a $S_{1.4} > 1$ mJy flux density limit (magenta), to highlight the difference in the redshift distribution of sources from surveys such as NVSS and FIRST and that of the Stripe 82 survey.

The results presented in Table 2.1 are the full extent of the information extracted from the angular two-point correlation functions measured for the full JVLA-CnB sample, as well as the sample resulting from the use of the collapsing technique presented in the previous section. Looking first at the results from the full samples, we find the correlation length and galaxy bias values to be r_0=$15.67^{+10.30}_{-8.93}$ h^{-1}Mpc, $b(z) = 2.45^{+2.10}_{-1.23}$ and r_0=$12.68^{+3.30}_{-4.12}$ h^{-1}Mpc, $b(z) = 2.74^{+1.51}_{-1.05}$ for the eastern and western fields, respectively. Radio sources generally trace a high-mass and therefore more biased galaxy population, the clustering strength is therefore expected to be somewhere between that of local Milky Way type galaxies ($r_0 \sim 5h^{-1}$Mpc; Davis & Peebles 1983) and galaxy clusters ($r_0 \sim 20 - 25h^{-1}$Mpc; Bahcall & Soneira 1983). Given that the SKADS 440μJy sample used to provide an estimate of our sample's redshift distribution has a high median redshift ($z = 1.21$) we find these large r_0 and bias values to be in agreement with the studies of Daddi et al. (2001), McCarthy et al. (2001), McLure et al. (2004), Rocca-Volmerange et al. (2004), Best et al. (2005a) and Seymour et al. (2007) which show that there is a tendency for AGN to be found in

larger host-galaxies such as Extremely Red Objects (EROs) and richer environments. EROs at z $\sim$ 1, thought to be the progenitors of local bright elliptical galaxies (Dunlop et al. 2003) are found to have r_0 values $\sim$12$\pm$3 h^{-1}Mpc (Foucaud et al. 2007), which is consistent with our results. The results of the galaxy bias b(z), for the full eastern and western samples, at the SKADS median redshift of $z_{med} = 1.21$ are also higher than that of normal Milky Way type galaxies and are in keeping with bias values found by Lindsay et al. (2014a) ($b(z) = 1.12 \pm 0.03$). The bias measurements and their evolution with redshift will be further discussed in Section 2.3.5. These results are promising and consistent with those found in previous clustering studies of radio galaxies. However, we keep in mind the effect of a sample with sources spanning a broad redshift range on clustering measurements, i.e. high- and low-redshift sources are superimposed, resulting in clustering measurements which are difficult to interpret. This will be addressed in the following section.

Results obtained after application of the collapsing technique, show similarly high values for both the correlation length and galaxy bias, i.e. r_0=$12.54^{+1.70}_{-1.89}$ h^{-1}Mpc and $b(z) = 2.86^{+0.67}_{-0.57}$ for the eastern field and r_0=$13.56^{+1.50}_{-1.59}$ h^{-1}Mpc and $b(z) = 3.18^{+0.62}_{-0.53}$ for the western field. Comparing the collapsing results to those derived from the angular TPCF of the original, full radio samples, we find that the correlation length and bias values agree well within the uncertainties.

2.3 Evolution of Clustering & Bias With Redshift

In order to measure the evolution of the clustering of radio sources, and thus the galaxy bias b(z), over cosmic time, we cross-match the JVLA radio sources with deep optical data from the SDSS Stripe 82 region in order to obtain redshift information for the radio sources.

2.3.1 SDSS Data

The optical Sloan Digital Sky Survey (SDSS - York et al. 2000) has observed $\sim$10,000 square degrees of sky. The survey uses a dedicated 2.5 m telescope, which is equipped

with a large format mosaic CCD camera, to image the sky in five optical bands (*ugriz*), as well as two digital spectrographs to obtain the spectra of about one million galaxies and 100,000 quasars selected from the imaging data. The SDSS data used in this Thesis is that obtained over the Southern Equatorial Stripe, or, Stripe 82. Stripe 82 is a region on the celestial equator that the SDSS has imaged 70-90 times in *ugirz*, giving co-added optical data which is 1.9-2.2 magnitudes deeper than single run SDSS observations, reaching g∼24.5 (Jiang et al. 2014). The SDSS Stripe 82 coadd photometric redshift catalogue was obtained from the SDSS Data Archive Server (DAS) via `http://das.sdss.org/va/coadd_galaxies/`. The spectroscopic redshifts were obtained by submitting a query to the SDSS SkyServer DR7 Stripe 82 service.

2.3.2 SDSS Cross-matching

The cross-matching is performed using the software TOPCAT (`http://www.star.bris.ac.uk/~mbt/topcat/`). In order to determine the maximum separation distance that should be allowed for a successful match between a radio source and an optical galaxy, we produce a histogram of the separation distance between each radio source and its corresponding best optical match within 10 arcseconds. We then repeat this for a case in which a random offset between 10 and 20 arcseconds has been added to the positions of all radio sources. This allows for confidence in the separation distance below which a match becomes a random coincidence. In Figure 2.10 it is shown that for the case of the original source positions (no offset), the number of matches levels off at separation distances > 3.0 arcseconds, and when the random offset is added to the positions of the radio sources, we find that the number of sources with best match pair separations of < 3.0 arcseconds drops significantly. The results indicate that any best matched optical source at a separation distance > 3.0 arcseconds is ∼60% more likely to be coincidental/random and therefore unreliable.

Radio sources are first checked for a match with SDSS sources having spectroscopic redshift information, thereafter, those lacking a spectroscopic match are checked for a match with sources which have photometric redshifts. Performing a cross-match of sources within 3.0 arcseconds, and applying a cut at $z = 1.0$, we find 4,511 SDSS

matches for the radio sources from the eastern and western fields combined. The SDSS photometric redshifts are derived using a machine learning code that requires a representative spectroscopic training set. This training set is generally robust at $z < 1$, but becomes unreliable at $z > 1$ (Zhang et al. 2013). Figure 2.9 shows the general reliability of the Stripe 82 photometric redshifts (Prescott et al. 2018). Of the 4,511 matched sources, 1,420 have spectroscopic redshift information, with a mean redshift of $z\sim0.33$. A further 3,091 have photometric redshift information. The redshift distributions of these matched samples are shown in Figure 2.11. As expected, the availability of spectroscopic information decreases with redshift, however the number of spectroscopic matches remains fairly good out to $z\sim0.6$.

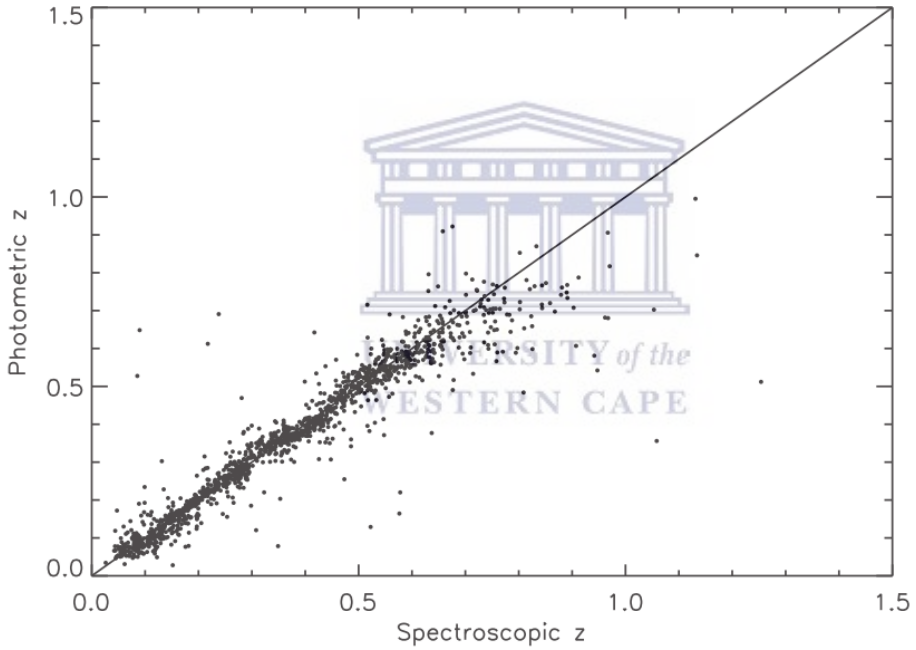


Figure 2.9: A comparison between the spectroscopic and photometric redshifts of 1445 Stripe 82 galaxies, taken from Prescott et al. (2018). The black solid line indicates where the two values are equal.

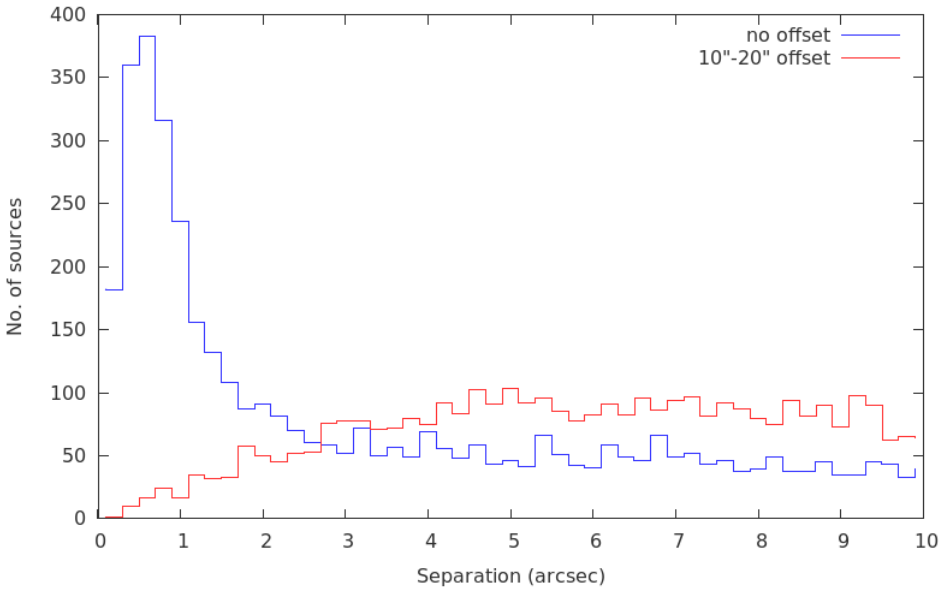


Figure 2.10: The blue histogram is that of the arcsecond separation between each JVLA source and its corresponding best SDSS match. The red histogram is that of the separation between best matched sources when a random positional offset between 10-20 arcseconds is added to all JVLA sources.

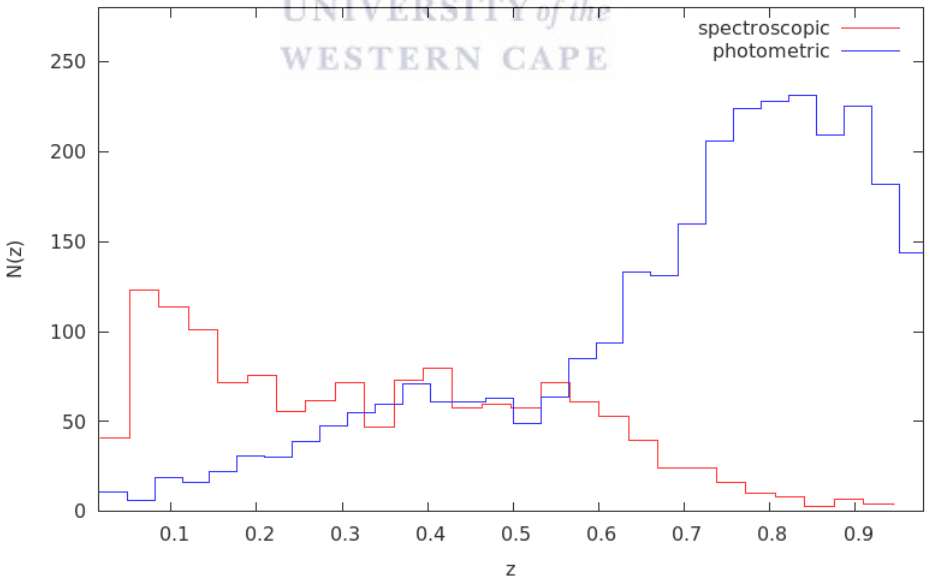


Figure 2.11: The redshift distributions of all JVLA sources, which have SDSS spectroscopic redshift matches (red) and photometric matches (blue). This sample of matched sources has a median redshift of $z = 0.62$.

2.3.3 ACF of Sources in Separate Redshift Bins

Having created a catalogue of JVLA sources with SDSS matches for $z \leq 1.0$ for both the eastern and western fields of the survey, I then perform a coarse binning of these sources into three bins in redshift, with the goal of investigating how the clustering of the radio galaxy population evolves over cosmic time. The chosen redshift bins are $z \leq 0.4$, $0.4 < \mathrm{z} \leq 0.7$ and $0.7 < \mathrm{z} \leq 1.0$, as these yield similar sized samples for each bin, with the high-redshift bin containing a slightly higher number of sources due to the larger volume of space sampled. The two-point angular correlation function is then calculated for each redshift-bin sample, as well as for the total matched sample, for both the eastern and western fields. The ACF analysis of the remaining unmatched sample will be presented in Chapter 3, where the data are used as an estimate for a high radio luminosity ($L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}) sample.

2.3.4 Results & Discussion

The resulting ACFs measured for the three redshift samples are presented in Figures 2.12 and 2.14 for the eastern and western samples, respectively. Due to the sources in these samples all having an optical match, i.e. being confirmed sources, the ACFs are now in the form of a single power-law. It is also observed that the power-law fit drops off around $\theta \sim 2.0°$. This is due to the subtraction of the integral constraint (described in Section 1.2.1) which corrects for the finite size of the survey and the width of the Stripe 82 field being $\sim 2.0°$. A χ^2_{min} fit is performed for each ACF while varying values for the slope and clustering amplitude. The negative values of the measured ACF, denoted by the dashed grey lines in Figures 2.12 and 2.14, were excluded during the fitting process. The best-fit parameters to the measured ACFs, as well as the resulting correlation length (r_0) and galaxy bias (b) values for the three redshift bins for both the eastern and western fields, are shown in Tables 2.2 and 2.3, respectively. From these results we see that the best-fit parameter values corresponding to the measured ACFs of the eastern and western fields agree well within the uncertainty margins. These results, along with those presented in Figures 2.12 and 2.14, show that the clustering amplitude, A, of the measured ACF is found to decrease from the first to the second redshift bin, but then increases slightly for the third bin. This is the case for both

the eastern and western field results. Figure 2.16 shows the redshift distributions of different radio source types from the $S_{1.4} > 440\mu$Jy S^3 sample. From this we see that the third redshift bin is most likely to be dominated by FR-I, Radio-Quiet and FR-II AGN galaxies, thus the increased clustering amplitude in the third bin is likely due to the presence of higher-redshift, strongly clustered AGN. The clustering amplitudes are larger than that found for studies of mJy samples (Cress et al. 1996, Blake & Wall 2002a, Overzier et al. 2003, Blake et al. 2004, Lindsay et al. 2014a). This is to be expected as the lower flux-density limit allows for the detection of powerful, but high-redshift FR-I and FR-II type galaxy populations which are strongly clustered. Looking at the γ= 1-η values of the measured ACFs for all redshift bins, these range from $\gamma = 1.82^{+0.08}_{-0.08}$ to $\gamma = 2.21^{+0.05}_{-0.06}$ for both the eastern and western field results, which are consistent with the findings from most studies of the clustering properties of radio galaxies.

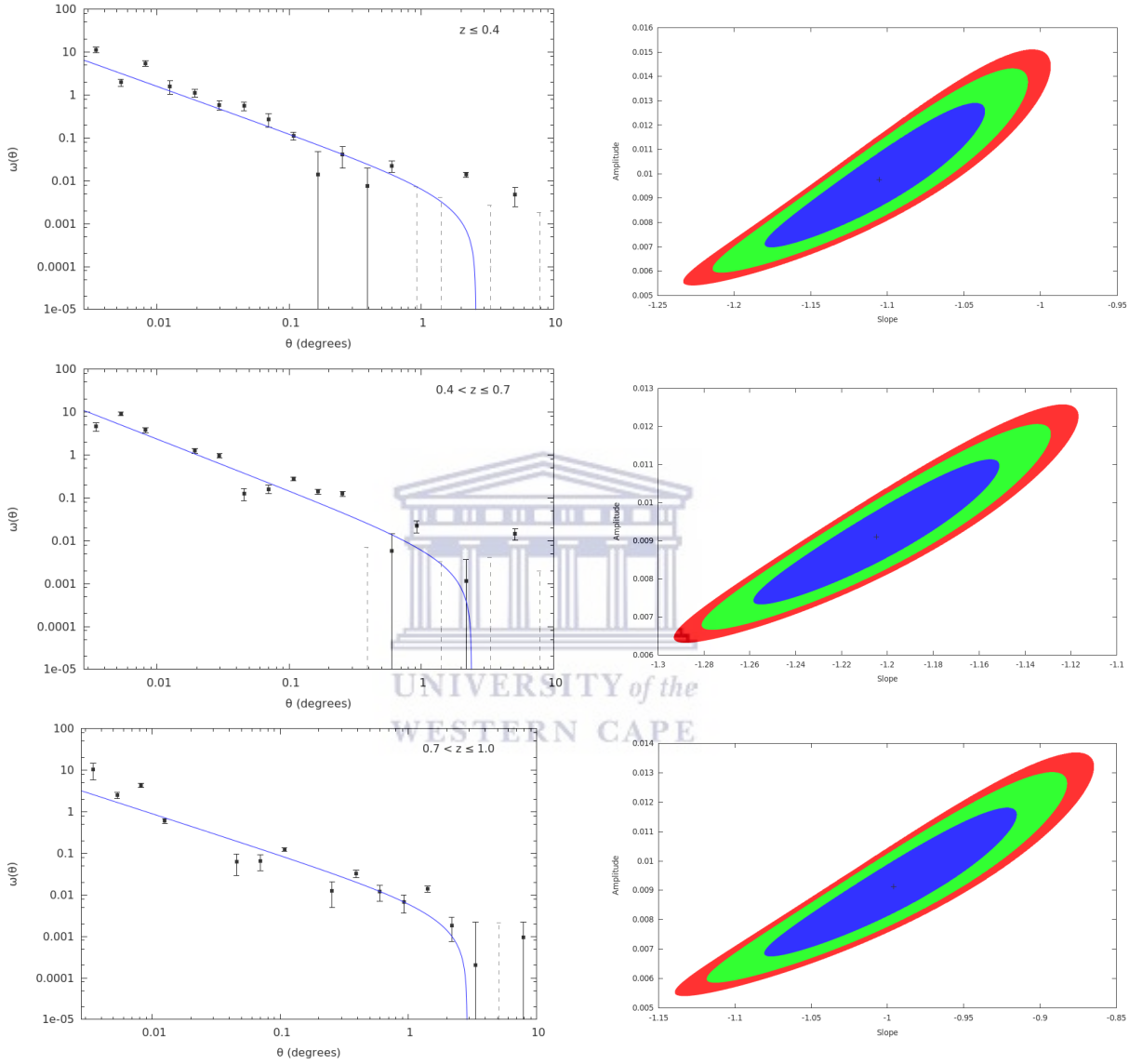


Figure 2.12: Left: The angular correlation function of eastern field JVLA radio sources cross-matched with SDSS optical sources for three redshift bins. Right: The 68%, 90% and 95% confidence levels for the large-angle single power law fits.

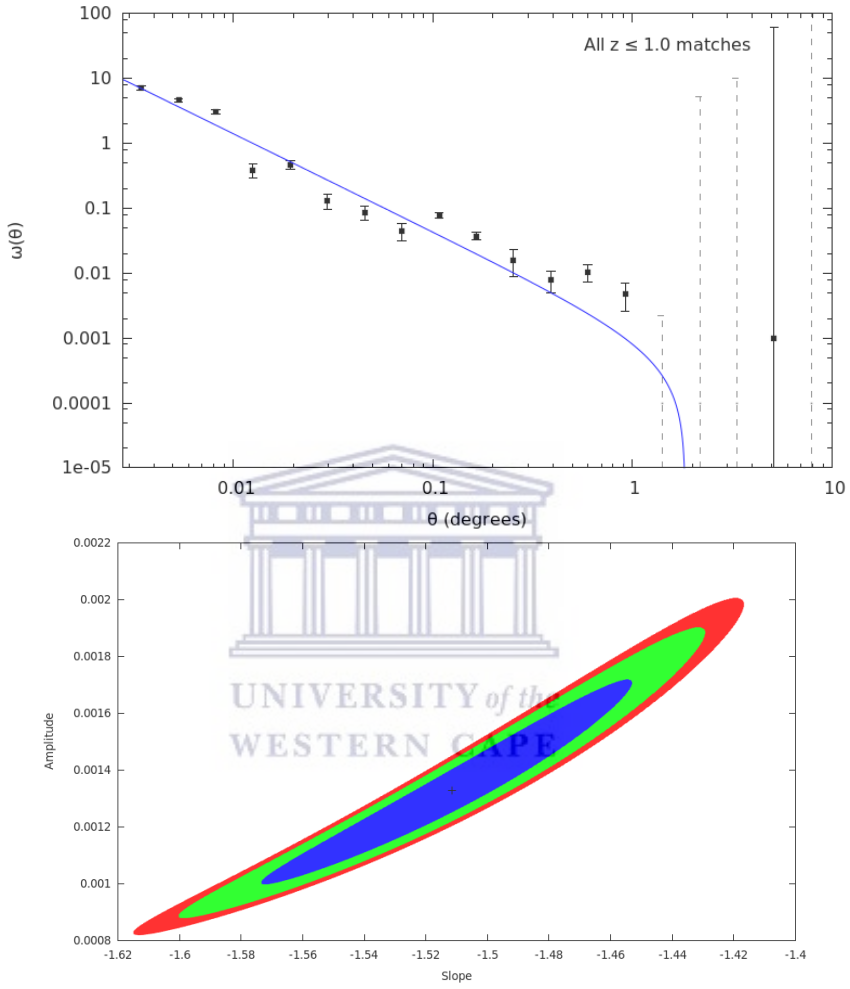


Figure 2.13: Top: The angular correlation function of all eastern field JVLA sources with SDSS optical matches. Bottom: The 68%, 90% and 95% confidence levels for the large-angle single power law fits.

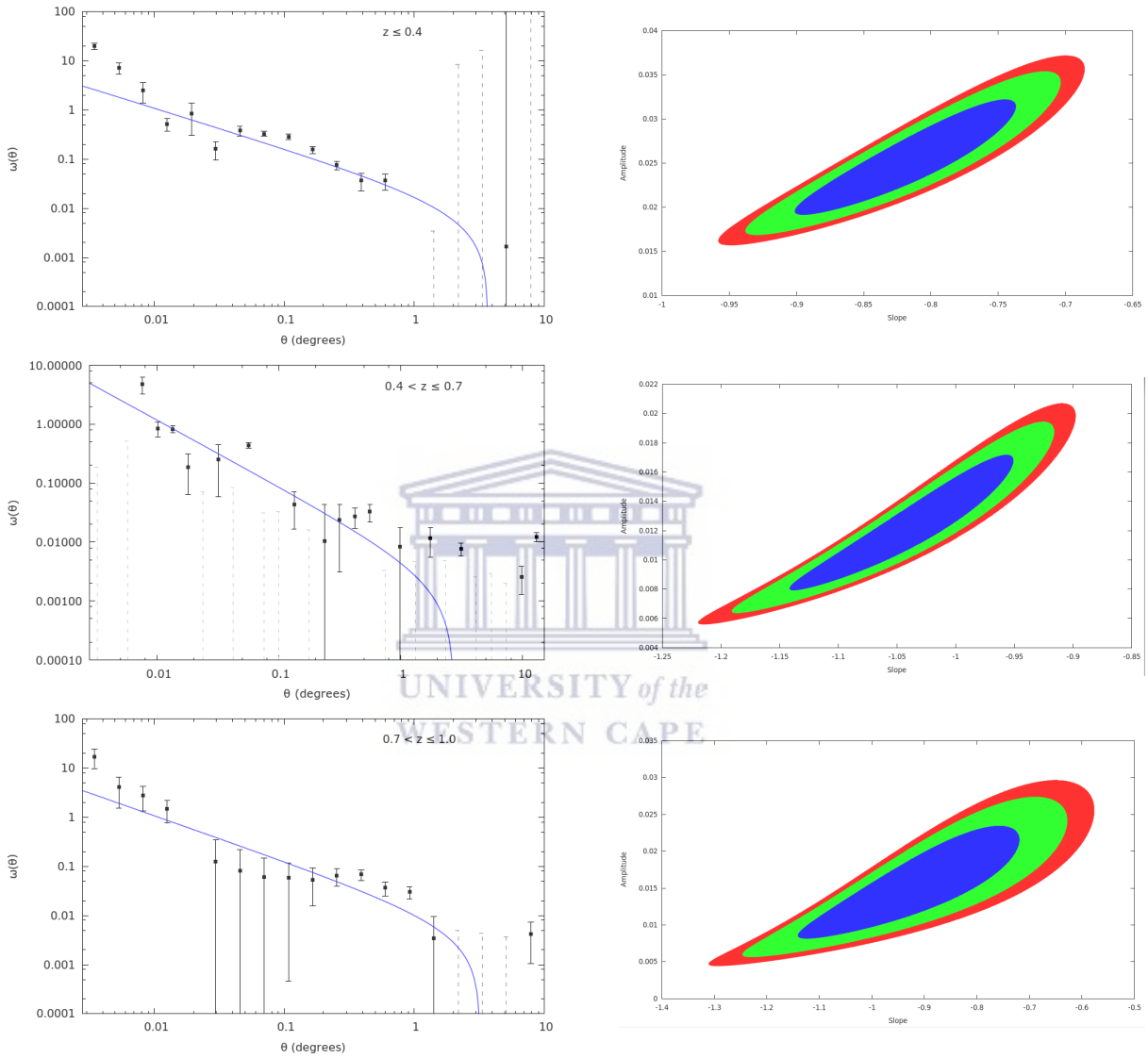


Figure 2.14: Top: The angular correlation function of western field JVLA radio sources cross-matched with SDSS optical sources for three redshift bins. Bottom: The 68%, 90% and 95% confidence levels for the large-angle single power law fits.

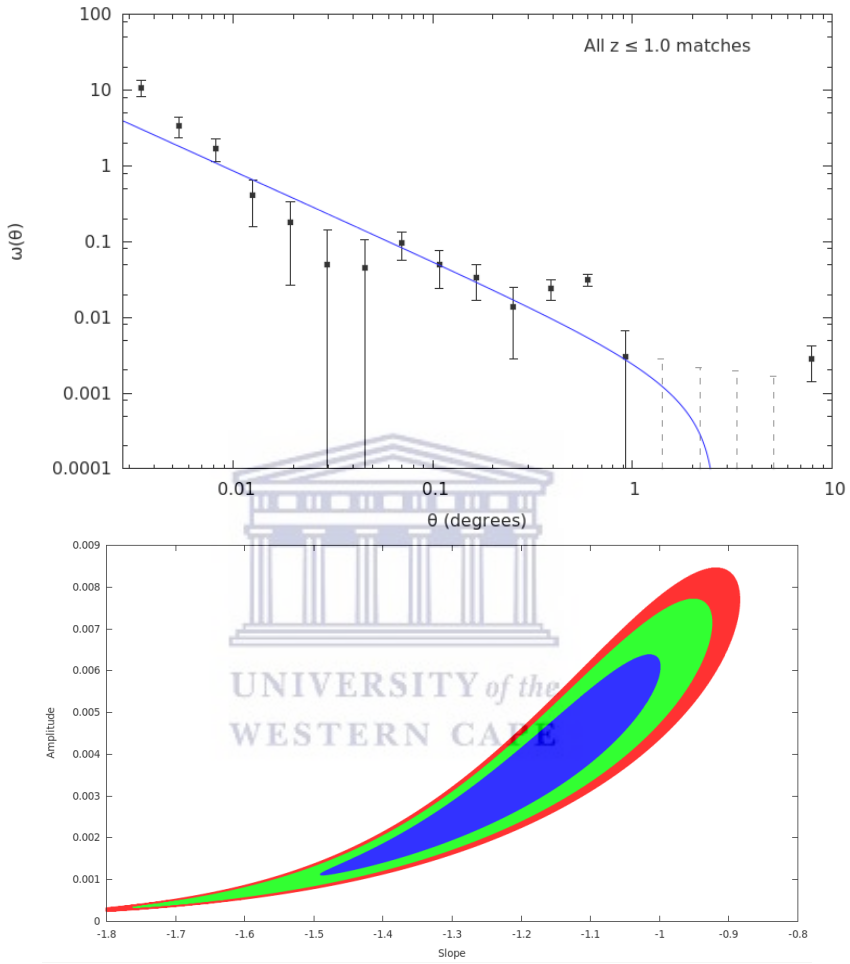


Figure 2.15: Left: The angular correlation function of all western field JVLA sources with SDSS optical matches. Right: The 68%, 90% and 95% confidence levels for the large-angle single power law fits.

Table 2.2: Best-fit parameters for the two-point angular correlation functions calculated for matched sources in the eastern field, in three separate redshift bins and the complete matched (z < 1) sample. Also, the corresponding correlation length r_0 and galaxy bias b(z) values. All uncertainties quoted are 1σ, obtained from χ^2_{min} fitting.

Eastern Sample	N_{obj}	z_{med}	A ($\times 10^{-2}$)	γ	r_0 (h^{-1}Mpc)	b(z_{med})
zbin 1 ($z \leq 0.4$)	634	0.22	$0.98^{+0.32}_{-0.28}$	$2.11^{+0.07}_{-0.07}$	$4.37^{+0.70}_{-0.75}$	$1.01^{+0.20}_{-0.20}$
zbin 2 ($0.4 < z \leq 0.7$)	687	0.57	$0.91^{+0.20}_{-0.18}$	$2.21^{+0.05}_{-0.06}$	$7.77^{+0.79}_{-0.83}$	$1.90^{+0.30}_{-0.28}$
zbin 3 ($0.7 < z \leq 1.0$)	873	0.83	$0.91^{+0.27}_{-0.24}$	$2.00^{+0.08}_{-0.09}$	$10.11^{+1.75}_{-1.84}$	$2.18^{+0.54}_{-0.47}$
All Matched	2,194	0.62	$0.13^{+0.04}_{-0.03}$	$2.51^{+0.06}_{-0.06}$	$5.16^{+0.63}_{-0.65}$	$1.48^{+0.35}_{-0.29}$

Table 2.3: Best-fit parameters for the two-point angular correlation functions calculated for matched sources in the western field, in three separate redshift bins and the complete matched (z < 1) sample. Also, the corresponding correlation length r_0 and galaxy bias b(z) values. All uncertainties quoted are 1σ, obtained from χ^2_{min} fitting.

Western Sample	N_{obj}	z_{med}	A ($\times 10^{-2}$)	γ	r_0 (h^{-1}Mpc)	b(z_{med})
zbin 1 ($z \leq 0.4$)	681	0.21	$2.52^{+0.71}_{-0.61}$	$1.82^{+0.08}_{-0.08}$	$6.34^{+1.06}_{-1.11}$	$1.33^{+0.26}_{-0.23}$
zbin 2 ($0.4 < z \leq 0.7$)	612	0.57	$1.22^{+0.50}_{-0.43}$	$2.04^{+0.10}_{-0.09}$	$7.96^{+1.53}_{-1.74}$	$1.76^{+0.50}_{-0.44}$
zbin 3 ($0.7 < z \leq 1.0$)	1,024	0.84	$1.52^{+0.82}_{-0.71}$	$1.93^{+0.21}_{-0.21}$	$12.37^{+3.76}_{-4.81}$	$2.53^{+1.44}_{-1.08}$
All Matched	2,317	0.65	$0.35^{+0.29}_{-0.24}$	$2.20^{+0.29}_{-0.20}$	$6.79^{+2.44}_{-3.38}$	$1.62^{+1.38}_{-0.88}$

2.3.5 Clustering Evolution Results

Figures 2.17, 2.19 and Figures 2.18, 2.20 show the most significant findings of this chapter. They show the evolution in redshift of r_0 and b(z), respectively, in comparison with the clustering models for various radio sub-populations, calculated by the SKADS semi-empirical simulation (Wilman et al. 2008). In these figures, our r_0 and b(z) values are also compared to those obtained by Lindsay et al. (2014a), who perform a clustering analysis on a $S_{1.4GHz} > 1$mJy sample from the JVLA FIRST survey, with spectroscopic (r > 19.8) and photometric information from the Galaxy And Mass Assembly (GAMA) survey (Driver et al. 2011). The results presented from their analysis, are for z < 0.5 and z > 0.5 redshift bin samples, a sample of all sources containing optical matches, as well as their full radio sample. For both the eastern and western fields,

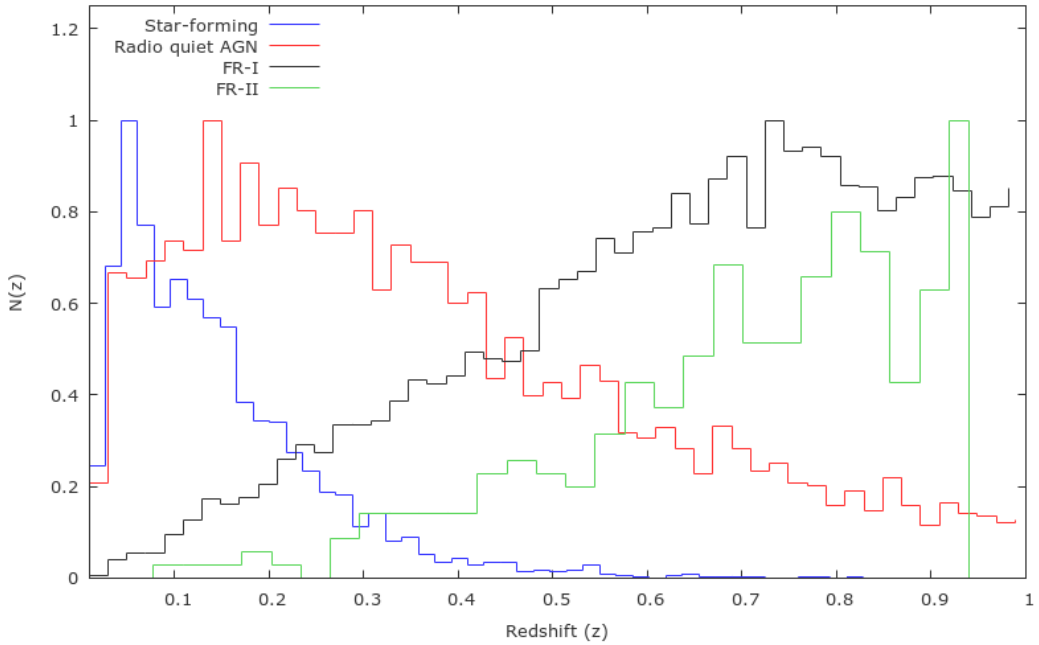


Figure 2.16: A comparison of the normalised redshift distribution of different radio source types in a $S_{1.4} > 440\mu$Jy S^3 extragalactic sample. These distributions aid our understanding of which source types are dominant in the various redshift bins used.

the correlation length r_0 and the galaxy bias b(z) are found to increase with redshift. This is to be expected for radio observations, as higher redshifts are dominated by FRI and FRII type galaxies, which trace large halo masses in the Universe, while sources at low redshifts trace radio quiet and star-forming galaxies which are expected to be less strongly clustered. A closer look at Figures 2.18 and 2.20 show that the $z < 0.4$ source population is closest to the SKADS evolution models for star-forming and radio quiet type galaxies, as expected. Whereas the results for the two higher redshift bins closely follow the SKADS models for the bias evolution of FR-I radio galaxies, which dominate the population at higher redshifts (see Figure 2.16), and the bias results for the full JVLA sample agree with this finding. The kinks and truncation in the SKADS models at $z = 1.5$ for the r_0 and bias evolution models, respectively, were imposed by Wilman et al. (2008) in order to avoid an unphysical increase in clustering. The correlation length results for the third redshift bin ($0.7 < z \leq 1.0$), as well as the total JVLA radio sample, shown in Figures 2.17 and 2.19, show that this high-redshift sample closely follows the properties of a more strongly clustered population comprised mostly of FRII type galaxies. The results for the redshift evolution of the bias also compare well to the bias evolution found by Passmoor et al. (2013), who also measure

the bias in three different redshift bins, for a $S_{1.4GHz} > 2$mJy FIRST sample with SDSS matches.

Tables 2.2 and 2.3 also present the best fit parameters and resulting r_0 and b(z) values for the measured ACFs of the full sample of JVLA sources having an optical match, for both the eastern and western fields, respectively. The resulting ACFs for the eastern and western fields are seen in Figures 2.13 and 2.15, respectively, and are quite intriguing. The parameters resulting from the ACF of all JVLA sources having an optical SDSS match, show that when the ACF is calculated over a large redshift range, the clustering signal is somewhat diluted, as seen by the low clustering amplitude. The r_0 and b(z) values are also slightly lower than the average of that of the three redshift bins, adding further justification to this point.

2.4 Conclusion

A clustering analysis of a $S_{1.4GHz} > 440\mu$Jy, $10'' \times 16''$ resolution sample of radio sources observed with the JVLA telescope in CnB configuration was performed. The clustering analysis was performed for the entire sample, as well as samples resulting from the collapsing of sources within a radius of 0.02° into a single source. A cross-matched sample of radio galaxies with optical SDSS sources was then created, providing host galaxy redshifts, of which $\sim$32% are spectroscopic. A clustering analysis of sources split into three bins in redshift was then performed, in order to study the evolution of the correlation length r_0 and bias b(z) with redshift. I compare these findings to evolution models from the S^3 extragalactic simulation of different types of radio sources.

Although the clustering results of the full sample are difficult to interpret, as they represent the clustering properties of different source types over a large redshift range, they are given for completeness and are useful in making comparisons with previous clustering measurements in literature. The analysis of the full and collapsed eastern and western samples yield similar results, with r_0 values of $r_0 = 15.67^{+10.30}_{-8.93} h^{-1}$Mpc and $r_0 = 12.54^{+1.70}_{-1.89}\ h^{-1}$Mpc for the eastern field and $r_0 = 12.68^{+3.30}_{-4.12}\ h^{-1}$Mpc and

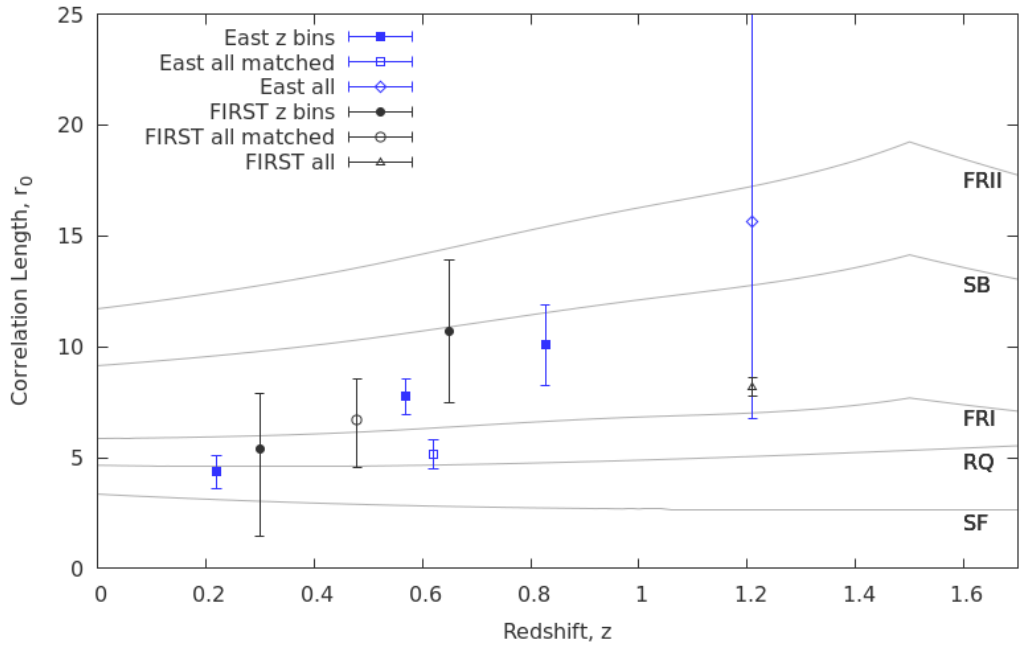


Figure 2.17: Measured correlation length (r_0) values for the eastern field, compared to SKADS prediction models for the evolution of r_0 with redshift, for various radio source types, denoted by the grey lines. The measured r_0 values for this analysis are shown in blue, where the open square corresponds to all sources with an SDSS match, filled squares correspond to the matched sources split into three redshift bins, and the open diamond corresponds to the full sample. The black symbols show the results obtained by Lindsay et al. (2014a) for a clustering analysis of a $>$ 1mJy sample of FIRST sources. The open triangle corresponds to the full sample, whilst the open circle corresponds to all FIRST sources with matches from GAMA and the filled circles correspond to matched sources separated into two redshift bins. In the case of the high-redshift bin, as in the Lindsay et al. (2014a) FIRST analysis, the redshift is estimated by using the median redshift of the S^3 $S_{1.4} > 440\mu$Jy sample (see Section 2.2.2).

$r_0 = 13.56^{+1.50}_{-1.59}\ h^{-1}$Mpc for the western field. The same is observed for the b(z) values for the full and collapsed samples of the eastern and western fields, with values of $b = 2.45^{+2.10}_{-1.23}$ and $b = 2.86^{+0.67}_{-0.57}$, and $b = 2.74^{+1.51}_{-1.05}$ and $b = 3.18^{+0.62}_{-0.53}$, respectively. These results are indicative of a strongly clustered, highly biased sample, dominated by massive sources such as AGN type galaxies. According to favoured models of galaxy formation, massive galaxies are thought to attain their mass through a biased hierarchical build-up. From previous studies, we expect that powerful radio galaxies would reside in the most overdense regions of the early Universe and therefore provide powerful test-cases of the biased growth of cosmic structures during the formative stages of the most massive galaxies.

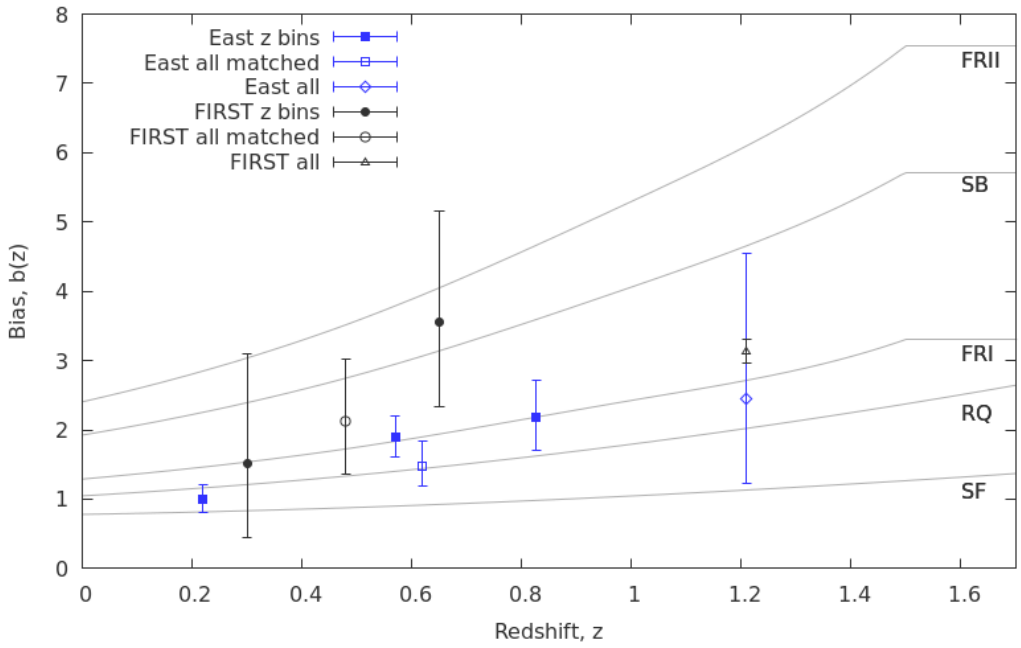


Figure 2.18: Measured galaxy bias (b(z)) values for the eastern field, compared to SKADS prediction models for the evolution of b(z) with redshift, for various radio source types, denoted by the grey lines. The symbols corresponding to the measured b(z) values are as defined in Figure 2.17.

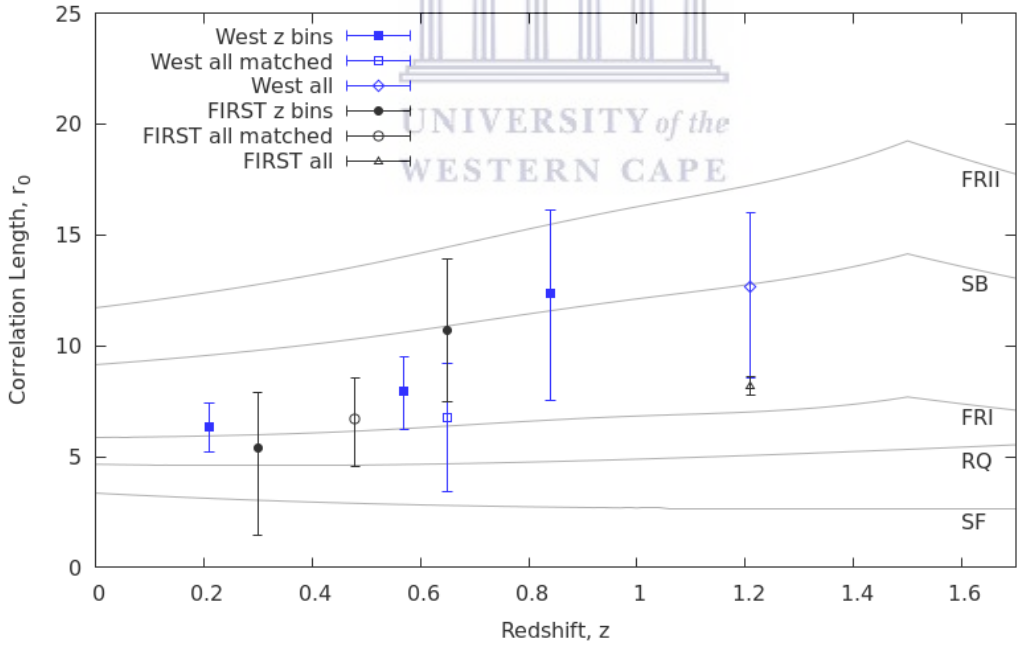


Figure 2.19: Measured correlation length (r_0) values for the western field, compared to SKADS prediction models for the evolution of r_0 with redshift, for various radio source types, denoted by the grey lines. The symbols corresponding to the measured r_0 values are as defined in Figure 2.17

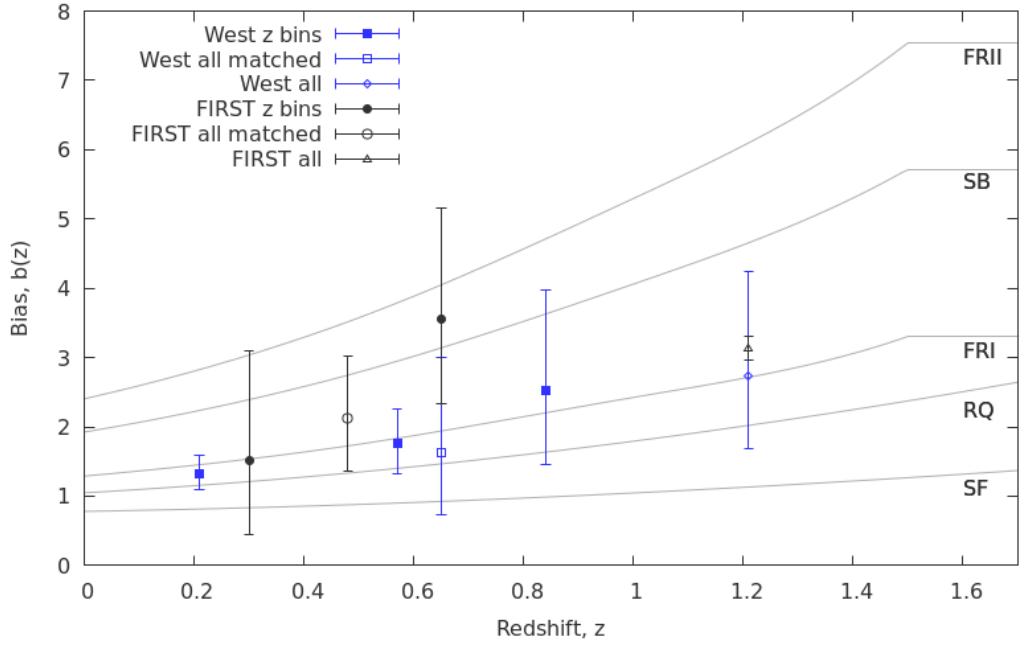


Figure 2.20: Measured galaxy bias (b(z)) values for the western field, compared to SKADS prediction models for the evolution of b(z) with redshift, for various radio source types, denoted by the grey lines. The symbols corresponding to the measured b(z) values are as defined in Figure 2.17)

The ACF results of the separate redshift bins show an increase in both correlation length and bias values with redshift, at a rate consistent with the simulated r_0 and b(z) evolution models from SKADS. Clustering measurements of the low-redshift bin ($z \leq 0.4$) yield correlation length and bias values of $r_0 = 4.37^{+0.70}_{-0.75}$ h^{-1}Mpc and $b = 1.01^{+0.20}_{-0.20}$, and $r_0 = 6.34^{+1.06}_{-1.11}$ h^{-1}Mpc and $b = 1.33^{+0.26}_{-0.23}$ for the eastern and western fields, respectively. Similarly, for the mid-redshift bin ($0.4 < z \leq 0.7$) I find r_0 and $b(z)$ values of $r_0 = 7.77^{+0.79}_{-0.83}$ h^{-1}Mpc and $b = 1.90^{+0.30}_{-0.28}$, and $r_0 = 7.96^{+1.53}_{-1.74}$ h^{-1}Mpc and $b = 1.76^{+0.50}_{-0.44}$ for the eastern and western fields, respectively. Finally, in the highest redshift bin ($0.7 < z \leq 1.0$) I find r_0 and b(z) values of $r_0 = 10.11^{+1.75}_{-1.84}$ h^{-1}Mpc and $b = 2.18^{+0.54}_{-0.47}$, and $r_0 = 12.37^{+3.76}_{-4.81}$ h^{-1}Mpc and $b = 2.53^{+1.44}_{-1.08}$ for the eastern and western fields, respectively. When compared to clustering evolution models from the S^3 simulation, the first redshift bin follows the clustering properties of star-forming and radio quiet galaxies, as expected. The results for the higher redshift bins agree very well with the clustering predictions of more massive, highly clustered FR-I, FR-II and star-burst type galaxies.

The results for the ACF analyses of the full eastern and western samples containing optical matches (z < 1) produced r_0 and b(z) values of $r_0 = 5.16^{+0.63}_{-0.65}$ h^{-1}Mpc and $b = 1.48^{+0.35}_{-0.29}$, and $r_0 = 6.79^{+2.44}_{-3.38}$ h^{-1}Mpc and $b = 1.62^{+1.38}_{-0.88}$ for the eastern and western fields, respectively, showing the dilution of the clustering signal caused by the superposition of sources from unrelated redshift slices during the measurement of the ACF. In the chapters to follow we will more closely investigate the clustering properties of different radio source sub-populations.

Chapter 3

Clustering properties of radio galaxies based on radio luminosity

In the previous chapter I could, to some degree, infer the clustering of various radio sub-populations. This was achieved by splitting the sources into bins in redshift and calculating the angular correlation function for the source sample in each redshift slice. With input from the redshift distribution of different types of radio galaxies from the S^3 simulation, it was possible to infer the radio source type expected to dominate each redshift bin and thereby estimate their clustering properties, i.e. the correlation length r_0 and bias $b(z)$. The results were also compared to the r_0 and bias evolution models from the S^3 simulations for different radio sub-populations. In this chapter, in order to gain further understanding of the clustering properties of different types of radio sources and how these evolve over cosmic time, the JVLA sources having $z < 1$ optical matches from Chapter 2, are split based on radio luminosity, as each luminosity range would probe a different radio source population. A clustering analysis is then performed for each luminosity bin sample. Throughout this chapter, when referring to luminosity, it will be the radio luminosity of the source at 1.4 GHz.

3.1 Radio Luminosity

The radio luminosity L_ν of a source is defined as the total power per unit bandwidth radiated by the source at frequency ν. Thus, using the area of a sphere and the inverse-square law, the relation between luminosity and flux density is given by:

$$L_\nu = 4\pi \times d_L^2 S_\nu \tag{3.1}$$

where S_ν is the total flux density at frequency ν and d_L is the luminosity distance in the adopted cosmology and is calculated as follows,

$$d_L = (1+z) \times d_M(z),, \tag{3.2}$$

and

$$d_M = d_H \int_0^z \frac{1}{E(z')} dz', \tag{3.3}$$

where E(z) is as in Equation 1.12 and $d_H = \frac{c}{H_0}$. We therefore calculate the radio luminosity using the formula

$$L_\nu = 4\pi \times d_L^2 \times S_\nu \times (1+z)^{\alpha-1} \times 10^{-26}, \tag{3.4}$$

where L_ν is now given in W Hz^{-1}, S_ν is the measured radio flux density in Jy and $(1+z)^\alpha$ is the standard k-correction used for radio astronomy, where α is the radio spectral index which measures the dependence of the flux density of a source on frequency, i.e. $S_\nu \propto \nu^\alpha$. A value of $\alpha = -0.8$ is usually adopted for radio galaxies (Intema et al. 2011; Hardcastle et al. 2016).

3.1.1 The Luminosity of Radio Galaxy Populations

The radio galaxy population at low luminosity, i.e. $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, is likely to consist of mostly of star-forming galaxies, as well as less powerful, lower redshift AGN and radio-quiet sources (Condon et al. 2002). The radio source population above $L_{1.4\mathrm{GHz}}$=10^{23} W Hz^{-1} is likely to comprise mostly of AGN and few SFGs. Less powerful, low-luminosity AGN are most likely to be Fanaroff-Riley type I galaxies whilst the

more powerful, edge-bright Fanaroff-Riley type II galaxies are high-luminosity sources (Fanaroff & Riley 1974). The division in luminosity of FRI and FRII galaxies occurs at roughly $L_{1.4\mathrm{GHz}} \sim 10^{25}$ W Hz^{-1}, buy may vary depending on the optical luminosity of the host galaxies of AGN (Ledlow & Owen 1996), although some recent studies suggest that the divide in power between FRI and FRII type AGN may not exist (Mingo et al. 2019). With regard to the high luminosity bin, we would expect the radio population at $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1} to consist only of powerful, high redshift AGN (Best & Heckman 2012), as the highest redshift sources must be extremely luminous in order to be above the radio detection limit.

3.2 Luminosity Limited Samples

In this chapter, I use the sample of JVLA sources for which $z < 1$ optical (SDSS) matches were obtained in Chapter 2, as well as the sample of unmatched JVLA sources. As in Chapter 2, the matched sample is limited to $z < 1$, as we can expect redshifts below this limit to be reliable. In order to study the clustering properties, as well as the evolution in clustering of different types of radio galaxy populations, we calculate the angular correlation function for samples split by radio luminosity. We have chosen the following splits: (i) $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, (ii) 10^{23} W Hz^{-1} $< L_{1.4\mathrm{GHz}} < 10^{26}$ W Hz^{-1} and (iii) $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}. The high luminosity bin, however, is made up of all JVLA sources lacking an SDSS match, as it is reasonable to assume, given the depth of the SDSS coadded data, that sources without an optical match are most likely to be high redshift (z > 1.0), high luminosity sources. The redshift distribution, as well as the luminosity distribution of the low and mid-luminosity samples are presented in Figures 3.1 and 3.2, respectively. The low luminosity bin is comprised mostly of low-redshift sources, with a tail to higher redshift ($z \sim 0.7$), whereas the mid-luminosity bin contains sources which dominate at higher redshifts and peak at $z \sim 0.8$. In this chapter, all ACFs, corresponding correlation length and bias values, as well as their respective uncertainties, are calculated using the same methods as those described in Chapter 2.

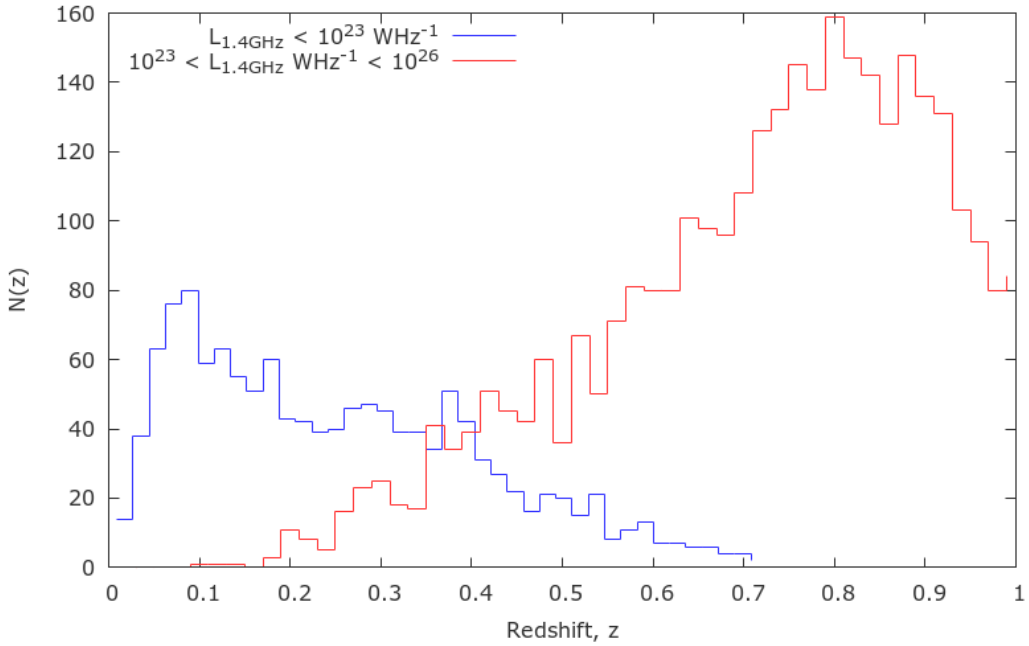


Figure 3.1: The redshift distribution of the matched ($z < 1$) JVLA sources, split into two bins in luminosity, i.e. $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1} and 10^{23} W Hz^{-1} $< L_{1.4\mathrm{GHz}} < 10^{26}$ W Hz^{-1}.

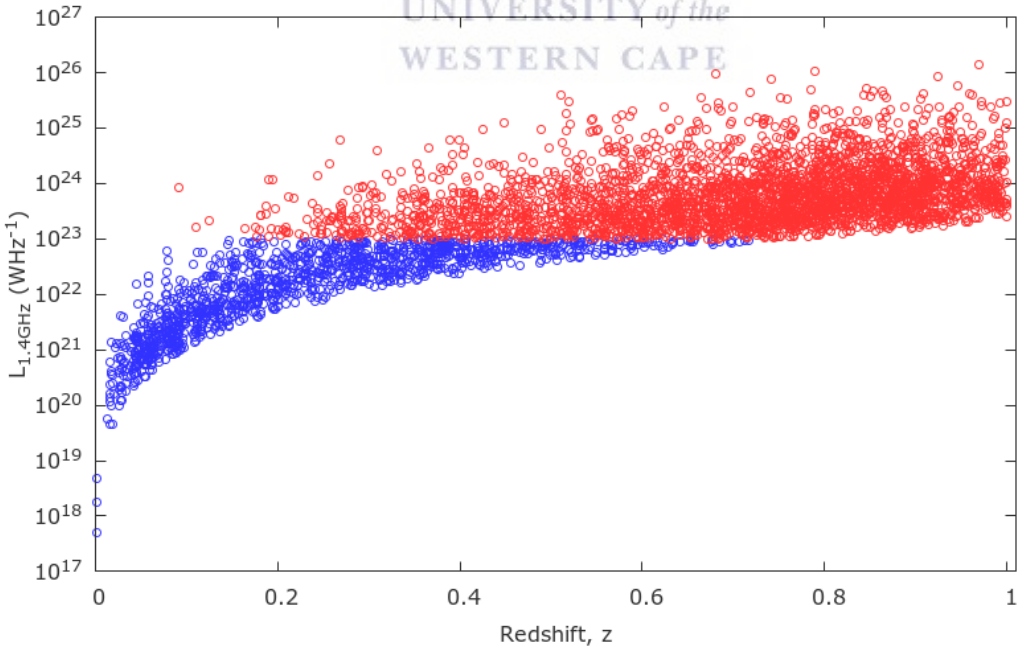


Figure 3.2: Radio luminosity distribution of the matched ($z < 1$) JVLA sources, where sources are split at $L_{1.4\mathrm{GHz}}$=10^{23} W Hz^{-1} into a low and mid radio luminosity bin.

3.3 Results & Discussion

3.3.1 Measured ACFs

The ACFs calculated for these three samples are shown in Figures 3.3 and 3.4, for the eastern and western fields, respectively. The corresponding best fit parameters to the ACFs, for the eastern and western fields are presented in Table 3.1 and Table 3.2, respectively. Table 3.3 shows the percentage of each radio source type that would be expected to make up each of our chosen luminosity bins. These estimates are obtained using sources from a S^3 $S_{1.4\mathrm{GHz}} > 440\mu$Jy, $z < 1$ sample for the low and mid-luminosity bins, to match the $z < 1$ limited JVLA sample, and a $z < 3$ sample for the high luminosity bin, as this is the limit we impose on the SKADS sample when estimating the redshift distribution of the unmatched JVLA sample. To obtain information regarding the relative source fractions in each luminosity bin, the luminosity of each SKADS source is calculated using its $S_{1.4\mathrm{GHz}}$ flux and redshift, and the sources are then split into the three bins in luminosity. The SKADS sources are each already assigned a source type from the S^3 simulations.

The ACFs calculated for the low luminosity bin, for both the eastern and western fields, follow the form of a single power-law, however some slight upturn is seen at $\theta \sim 0.02$ degrees in the ACF of the eastern sample. A more distinct upturn is observed, at angular scales similar to that in the low luminosity bin, in the ACFs of the mid-luminosity bin ($10^{23} < L_{1.4\mathrm{GHz}}$ W Hz$^{-1} < 10^{26}$) for both the eastern and western samples, though the upturn is steeper for the western field. There is now a distinct break in the ACF towards small angular scales, i.e. the ACF takes the form of a double power-law. This radio sample has been cross-matched with host galaxies from SDSS, so we would not expect this upturn to be caused by the components of resolved extended sources. It is possible that false identifications with optical sources may have occurred, however, an upturn was not observed in the ACFs of the matched JVLA sources split by redshift in Chapter 2. The upturn in the ACFs at small angular scales may also be due to the clustering of galaxies which share the same dark matter halo, i.e. the contribution from the 1-halo term of the halo occupation model

(HOD) (Zheng et al. 2005). Lower mass sources such as star-forming galaxies or low mass AGN, which could share the same dark matter haloes, are galaxies which may contribute to the 1-halo term. This deviation from the single power-law, marking the transition between the 1- and 2-halo terms, has previously been found to occur at $r \sim 1 - 2h^{-1}$Mpc ($\theta \sim 1$ arcmin at $z \sim 0.5$) (Zehavi et al. 2004; Zehavi et al. 2005; Sawangwit et al. 2011). This corresponds to the scale at which the upturn in observed for the low ($z_{\rm med}$= 0.25) and mid-luminosity ($z_{\rm med}$= 0.75) ACFs in this chapter, i.e. $\theta \sim 0.01 - 0.03$ deg. Due to this upturn, a double power-law of $\omega(\theta) = A\theta^{\eta} + B\theta^{\kappa}$ is fit to the ACFs of the mid-luminosity bins, where the large- ($A\theta^{\eta}$) and small-angle ($B\theta^{\kappa}$) power laws are fit separately over different scales. The clustering amplitude B and slope κ values measured for the small-angle power law (fitted for $\theta < 0.02$) are found to be $B = 5.58^{+0.48}_{-0.68} \times 10^{-5}$; $\kappa = -2.19^{+0.31}_{-0.32}$ and $B = 8.09^{+0.08}_{-0.10} \times 10^{-5}$; $\kappa = -2.14^{+0.04}_{-0.06}$ for the eastern and western samples, respectively. These parameters differ slightly to those measured by Blake & Wall (2002a) for a $S_{1.4\rm GHz} > 10$ mJy sample of NVSS sources, who find values $B = 0.95 \pm 0.02 \times 10^{-5}$ and $\kappa = -3.18 \pm 0.01$. Blake & Wall (2002a) attributed their measured small-angle power-law solely to the affect of multiple-component radio sources, as the NVSS sources were not matched with host galaxy positions. Our JVLA data are significantly deeper and more likely to pick up sources within the same dark matter halo.

The small-scale clustering measured in the low and mid-luminosity bin is however much less pronounced than that observed in the ACF of the high luminosity bin, for both the eastern and western fields. These samples consist of all unmatched JVLA sources, which we have assumed to have high luminosities of $L_{1.4\rm GHz} > 10^{26}$ W Hz^{-1} and redshifts $z < 3$. This assumption is supported by the resulting ACF, in both the eastern and western fields. For these ACFs, the small angle power-law is far more pronounced than all ACFs presented thus far in this Thesis, indicating that there are likely a high number of large, extended galaxies which have been resolved into multiple source components and which are most likely powerful, radio-loud AGN. This is supported by the source fractions presented in Table 3.3, which suggest that the high luminosity sample would comprise solely of AGN, and would be the only luminosity

sample containing FRII galaxies. In the case of the mid- and high luminosity bins, we have fitted the large-angle power-law ($A\theta^{\eta}$) over the range $\theta > 0.02$ degrees, as this is roughly the angular scale below which an up-turn/split in the ACF is observed (see Figures 3.3 and 3.4). We can therefore be fairly confident that at scales larger than this, the clustering measured is that of individual galaxies and not of radio source components.

3.3.2 Correlation Length & Bias Measurements

The results presented in Tables 3.1 and 3.2 show that in both the eastern and western fields, the correlation length and bias values of the samples increase with an increase in radio luminosity. The median redshift of the samples also increases with luminosity, showing that objects which have a higher radio luminosity occupy higher redshift ranges, as expected for a flux limited sample. The $\eta = 1 - \gamma$) values measured for all luminosity bins, in both eastern and western fields, are found to be ~ -0.8, consistent with those commonly measured for radio sources (Roche & Eales 1999; Norberg et al. 2002; Blake & Wall 2002a). The low luminosity samples of the eastern and western fields have measured clustering length and bias values of $r_0 = 8.02^{+1.88}_{-1.73}\ h^{-1}$Mpc; $b(z) = 1.62^{+0.45}_{-0.38}$ and $r_0 = 7.98^{+1.40}_{-1.53}\ h^{-1}$Mpc; $b(z) = 1.68^{+0.40}_{-0.35}$, respectively. These values are higher than that of local galaxies with masses comparable to the Milky Way ($r_0 \sim 5h^{-1}$Mpc - Davis & Peebles 1983) and are in line with clustering values measured for more massive galaxies (Brand et al. 2005; Wake et al. 2008; Lindsay et al. 2014b). The mid-luminosity samples of the eastern and western fields have clustering length and bias values of $r_0 = 9.53^{+2.93}_{-2.52}\ h^{-1}$Mpc, $b(z) = 2.01^{+0.85}_{-0.59}$ and $r_0 = 9.98^{+3.10}_{-2.98}\ h^{-1}$Mpc, $b(z) = 1.80^{+0.60}_{-0.51}$, respectively. These results show an increase in clustering compared to that of the low luminosity bin and are in keeping with the clustering values found for AGN (Magliocchetti et al. 2017; Hale et al. 2018). These measurements are also in line with the clustering values found by Zehavi et al. (2005) of $r_0 = 9.8 - 11.2 \pm 0.2h^{-1}$Mpc for different luminosity ranges, for a sample of luminous red galaxies (LRGs). Accord-

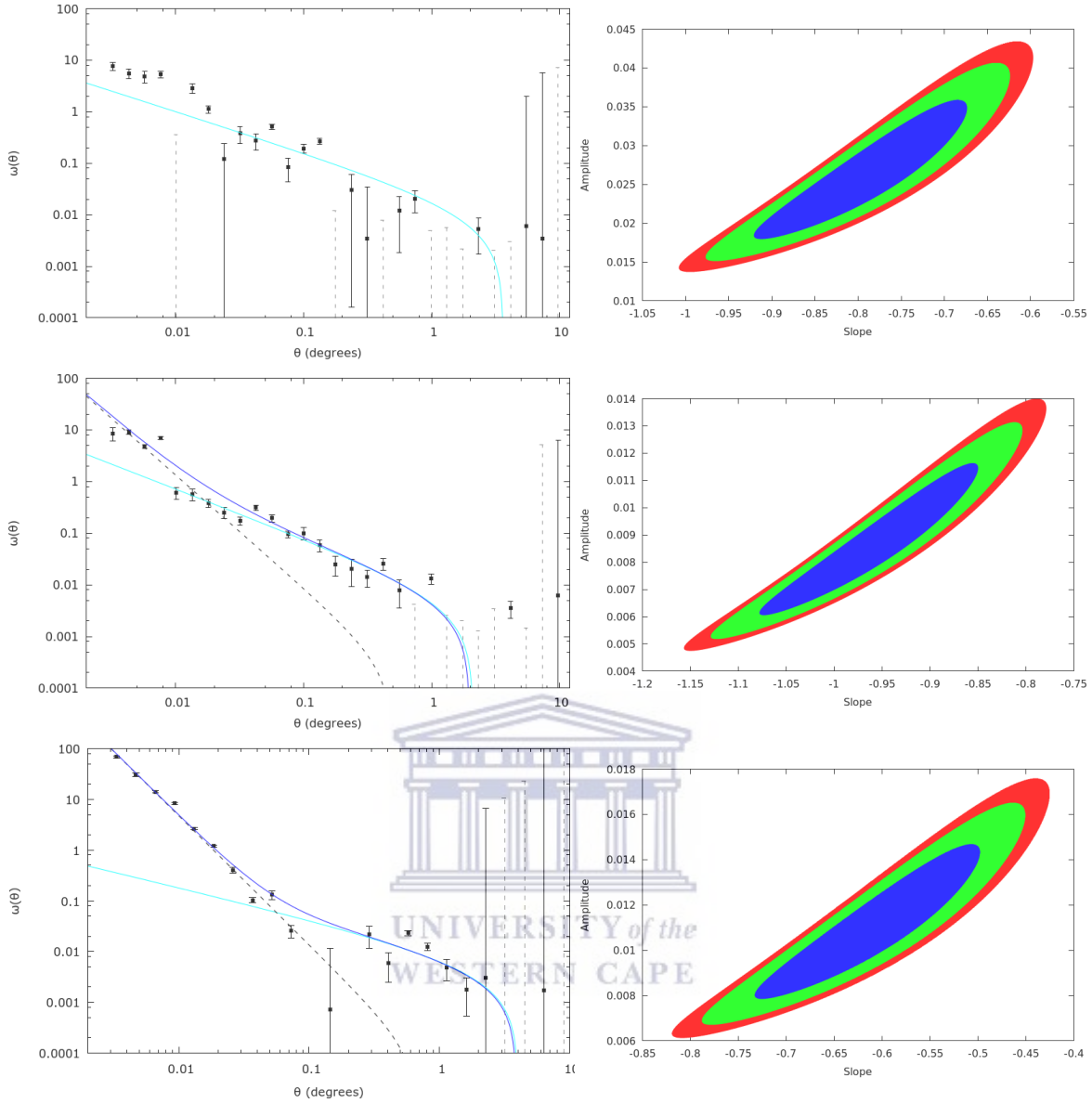


Figure 3.3: Left: The angular correlation function of eastern field JVLA sources with $z < 1$ SDSS cross-matches, which have been split into two bins in luminosity, i.e. $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1} and $10^{23} < L_{1.4\mathrm{GHz}}$ W Hz$^{-1} < 10^{26}$, with a third high luminosity bin ($L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}) occupied by all unmatched JVLA sources (presented in this order from top to bottom). The cyan coloured line and dashed black line indicate the best fit power-laws to the data, which dominate over large ($\theta > 0.1°$) and small angular scales, respectively. The blue line depicts the best fit double power-law, i.e. the sum of the small and large angle power laws minus an integral constraint σ^2 and the grey dashed lines show the negative points of the ACF. Right: The 68%, 90% and 95% confidence levels for the large-angle single power law χ^2_{min} fits.

ing to the source fractions presented in Table 3.3, it is suggested that ∼80 % of sources in this radio luminosity bin are FRI galaxies, which are hosted by massive galaxies, thus supporting the high clustering values measured. As shown in Figure 3.5, sources

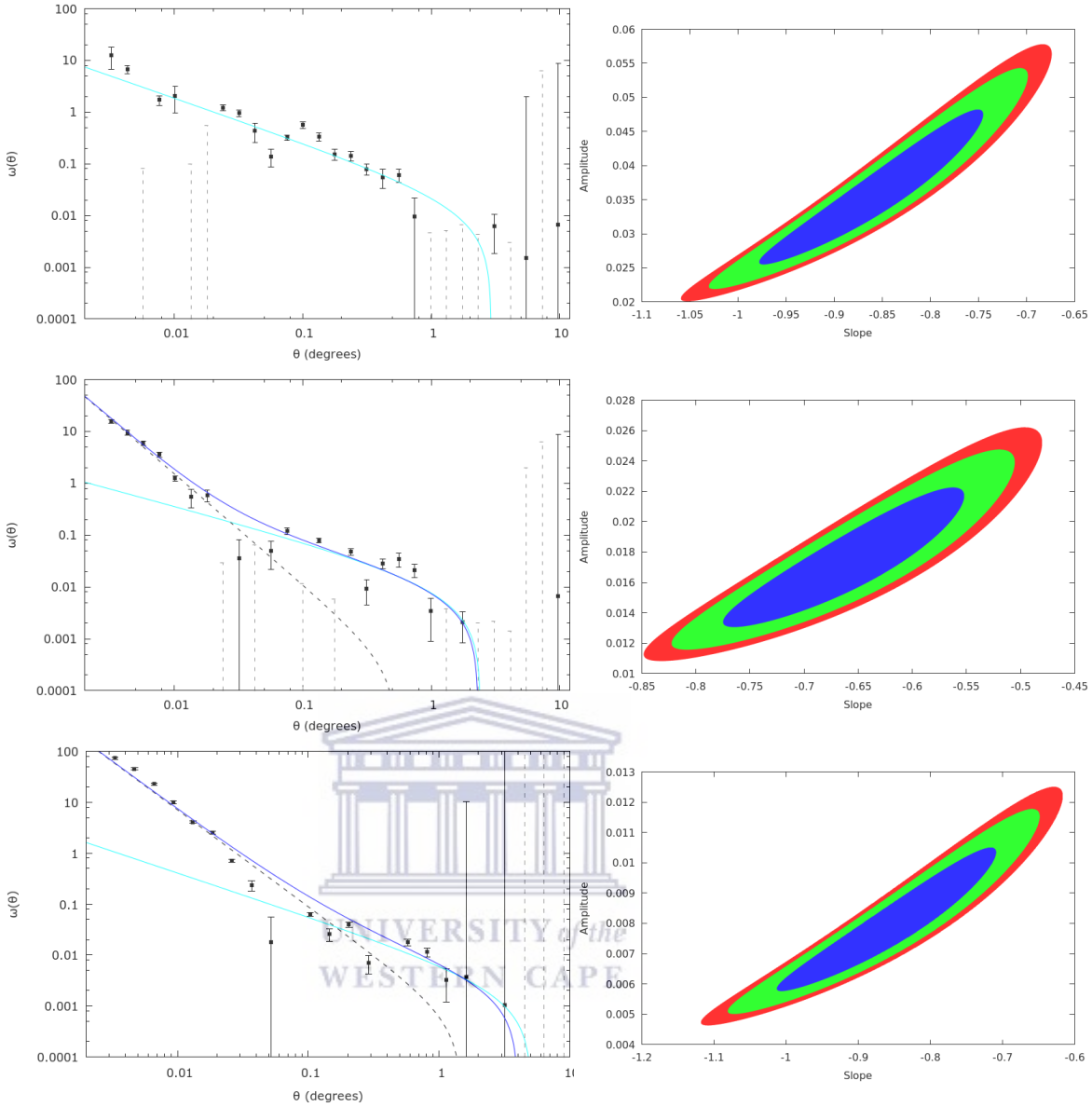


Figure 3.4: Left: The angular correlation function of western field JVLA sources with $z < 1$ SDSS cross-matches, which have been split into two bins in luminosity, i.e. $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1} and $10^{23} < L_{1.4\mathrm{GHz}}$ W Hz^{-1} $< 10^{26}$, with a third high luminosity bin ($L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}) occupied by all unmatched JVLA sources (presented in this order from top to bottom). The cyan coloured line and dashed black line indicate the best fit power-laws to the data, which dominate over large (θ >0.1°) and small angular scales, respectively. The blue line depicts the best fit double power-law, i.e. the sum of the small and large angle power laws minus an integral constraint σ^2 and the grey dashed lines show the negative points of the ACF. Right: The 68%, 90% and 95% confidence levels for the large-angle single power law χ^2_{min} fits.

in the $10^{23} < L_{1.4\mathrm{GHz}}$ WHz^{-1} $< 10^{26}$ luminosity bin occupy higher redshifts than that of the low luminosity bin.

The clustering analysis of the high luminosity bin, which consists of all unmatched JVLA sources, yields very large correlation length and bias values of $r_0 = 17.84^{+7.13}_{-5.86}$ h^{-1}Mpc, $b(z) = 3.21^{+1.49}_{-1.01}$ and $r_0 = 19.30^{+3.88}_{-4.48}$ h^{-1}Mpc, $b(z) = 3.72^{+1.45}_{-0.38}$ for the eastern and western samples, respectively. The redshift distribution used to model this sample, in order to determine the correlation length and bias values, is that of the S^3 $S_{1.4} > 440\mu$Jy sample at z < 3, which has a median redshift of z_{med}=2.03. These higher clustering values are to be expected, as the sources in this luminosity bin are likely to dominate the higher redshift, higher luminosity range of radio sources, which we would expect to be strongly clustered. As shown in Table 3.3, the radio population at $L_{1.4GHz} > 10^{26}$ W Hz^{-1} should comprise only of AGN. The high measured bias values suggest that these high luminosity, high-redshift AGN reside in more massive dark matter halos and are thus a highly clustered population which strongly trace out the dark matter distribution. The correlation length and bias values measured for all luminosity bins agree well between the eastern and western samples.

Table 3.1: Best-fit parameters for the two-point angular correlation functions calculated for matched sources in the eastern field, split by radio luminosity into three separate bins: $L_{1.4GHz} < 10^{23}$ W Hz^{-1}, $10^{23} < L_{1.4GHz}$ W Hz$^{-1} < 10^{26}$ and the unmatched JVLA sources, which are assumed to occupy the highest luminosity bin of $L_{1.4GHz} > 10^{26}$ W Hz^{-1}. Also, the corresponding correlation length r_0 and galaxy bias $b(z)$ values. All uncertainties quoted are 1σ, obtained from χ^2_{min} fitting.

Eastern Sample	N_{obj}	z_{med}	A($\times10^{-2}$)	γ	χ^2_{red}	$r_0(h^{-1}$Mpc)	$b(z_{med})$
$L_{1.4GHz} < 10^{23}$ WHz^{-1}	692	0.25	$2.60^{+1.00}_{-0.80}$	$1.79^{+0.13}_{-0.12}$	3.50	$8.02^{+1.73}_{-1.88}$	$1.62^{+0.45}_{-0.38}$
$10^{23} < L_{1.4GHz}$ WHz$^{-1} < 10^{26}$	1,501	0.74	$0.85^{+0.35}_{-0.25}$	$1.96^{+0.12}_{-0.11}$	3.11	$9.53^{+2.93}_{-2.52}$	$2.01^{+0.85}_{-0.59}$
Unmatched	3,166	2.03	$1.09^{+0.41}_{-0.29}$	$1.61^{+0.12}_{-0.11}$	2.86	$17.84^{+7.13}_{-5.86}$	$3.21^{+1.49}_{-1.01}$

3.3.3 Clustering Properties of SKADS Samples

To allow for direct comparison with the clustering results measured for the JVLA samples binned in radio luminosity, clustering analyses are performed for S^3 sources at $S_{1.4GHz} > 440$ μJy and having redshifts of z < 1 for the low and mid-luminosity bins and z < 3 for the high luminosity bin, yielding a sample of 33,274 simulated extragalactic sources covering 20×20 deg^2. The sample is split into the same luminosity

Table 3.2: Best-fit parameters for the two-point angular correlation functions calculated for matched sources in the western field, split by radio luminosity into three separate bins: $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, $10^{23} < L_{1.4\mathrm{GHz}}$ W Hz^{-1} $< 10^{26}$ and the unmatched JVLA sources, which are assumed to occupy the highest luminosity bin of $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}. Also, the corresponding correlation length r_0 and galaxy bias $b(z)$ values. All uncertainties quoted are 1σ, obtained from χ^2_{min} fitting.

Western Sample	N_{obj}	z_{med}	A($\times 10^{-2}$)	γ	χ^2_{red}	$r_0(h^{-1}\mathrm{Mpc})$	$b(z_{\mathrm{med}})$
$L_{1.4\mathrm{GHz}} < 10^{23}$ WHz^{-1}	616	0.19	$3.51^{+1.30}_{-1.01}$	$1.86^{+0.12}_{-0.12}$	4.18	$7.98^{+1.40}_{-1.53}$	$1.68^{+0.40}_{-0.35}$
$10^{23} < L_{1.4\mathrm{GHz}}$ WHz^{-1} $< 10^{26}$	1,700	0.75	$1.72^{+0.48}_{-0.42}$	$1.66^{+0.11}_{-0.11}$	3.08	$9.98^{+3.10}_{-2.98}$	$1.80^{+0.60}_{-0.51}$
Unmatched	3,588	2.03	$0.79^{+0.26}_{-0.21}$	$1.86^{+0.16}_{-0.15}$	2.60	$19.30^{+3.88}_{-4.48}$	$3.72^{+1.45}_{-0.38}$

Table 3.3: Percentage of source types in each of the three luminosity bins, based on luminosity cuts applied to a SKADS simulated $S_{1.4} > 440\mu$Jy data sample at z < 1 for the low and mid-luminosity bins and z < 3 for the high luminosity bin.

Luminosity Bin	$L < 10^{23}$ WHz^{-1}	$10^{23} < L$ (WHz^{-1}) $< 10^{26}$	$L > 10^{26}$ WHz^{-1}
Star-Forming	59%	1.5%	0%
Starburst	10%	11%	0%
Radio-Quiet	22%	7%	0%
FRI	8%	80%	70%
FRII	0%	0%	30%

bins as for the analysis of our JVLA sample, i.e. $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, 10^{23} W Hz^{-1} $< L_{1.4\mathrm{GHz}} < 10^{26}$ W Hz^{-1} and $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}.

The measured ACFs for the SKADS luminosity binned samples are presented in Figure 3.6 with the corresponding best-fit parameters, correlation length and bias values presented in Table 3.4. All ACFs exhibit a single power-law form, and are well constrained due to the large number of simulated sources in each sample. A double power-law form of the ACFs is not expected due to the fact that the exact positions and total luminosities are known, for all source components in the SKADS simulation. The clustering measurements of the SKADS samples are also shown to increase with luminosity as expected. The correlation length and bias values measured for the

low luminosity sample are $r_0 = 3.47^{+0.39}_{-0.37}$ h^{-1}Mpc and $b(z) = 0.79^{+0.07}_{-0.07}$, respectively. These values are much lower than the clustering values of $r_0 = 8.02^{+1.73}_{-1.88}$ h^{-1}Mpc; $b(z) = 1.62^{+0.45}_{-0.38}$ and $r_0 = 7.98^{+1.40}_{-1.53}$ h^{-1}Mpc; $b(z) = 1.68^{+0.40}_{-0.35}$ measured for the eastern and western JVLA low luminosity samples, respectively. The SKADS $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1} sample occupies a lower redshift range than that of the JVLA sample (see Figures 3.1 and 3.5) and has a lower median redshift of $z = 0.11$, compared to the median redshift of $z = 0.22$ of the JVLA low-luminosity sample. These results suggest that the SKADS simulations underestimate the number of low luminosity AGN, such as lower luminosity FRI galaxies.

The clustering properties measured for the SKADS mid-luminosity bin are found to be $r_0 = 7.19^{+0.95}_{-0.88}$ h^{-1}Mpc and $b(z) = 1.66^{+0.30}_{-0.25}$, which are slightly lower (but consistent within uncertainties) than those measured for the mid-luminosity JVLA sample, i.e. $r_0 = 9.53^{+2.93}_{-2.52}$ h^{-1}Mpc; $b(z) = 2.01^{+0.85}_{-0.59}$ and $r_0 = 9.98^{+3.10}_{-2.98}$ h^{-1}Mpc; $b(z) = 1.80^{+0.60}_{-0.51}$ for the eastern and western samples, respectively. The SKADS values are also slightly lower than those generally found for an AGN dominated sample ($r_0 \sim 10$ h^{-1}Mpc; Magliocchetti et al. 2004; Magliocchetti et al. 2017; Hale et al. 2018). The redshift distribution of the $10^{23} < L_{1.4\mathrm{GHz}}$ WHz$^{-1} < 10^{26}$ JVLA sample (see Figure 3.1) is shown to peak at a higher redshift of $z \sim 0.8$, where as that of the SKADS sample at the same luminosity limit is shown to have a broader redshift distribution with the number of high-redshift sources being fairly constant from $z \sim 0.6$ onwards. This also suggests that many of the JVLA sources in this luminosity bin have not been matched to optical galaxies. It is also possible that there may be some FRII galaxies present in this luminosity bin sample, while the SKADS simulation does not account for any FRII galaxies in this luminosity range.

The measured clustering values for the SKADS high luminosity sample ($L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}) are indicative of a very highly clustered sample, with correlation length and bias values of $r_0 = 20.21^{+1.83}_{-1.80}$ h^{-1}Mpc and $b(z) = 24.69^{+10.02}_{-6.81}$, respectively. The bias, although poorly constrained, is found to be extremely large compared to results from this chapter as well as measurements found in other work, even at the lower limit.

This is attributed to the steep slope of the measured ACF ($\eta = 2.88^{+0.04}_{-0.04}$), as well as the higher redshift distribution of the sample. Given that, according to the SKADS source fraction estimates for each luminosity bin (see Table 3.3), the high luminosity bin is the only SKADS sample in which FRII sources are found, the steep slope (and thus high bias) measured may be due to an overestimate by Wilman et al. (2008) of the halo masses which host FRII galaxies. The correlation length measured for the SKADS high luminosity sample is in agreement with that found for the JVLA high luminosity samples (within the uncertainties) with $r_0 = 17.84^{+7.13}_{-5.86}$ h^{-1}Mpc and $r_0 = 19.30^{+3.88}_{-4.48}$ h^{-1}Mpc measured for the eastern and western samples, respectively. This supports our assumption that the unmatched JVLA sources occupy the $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1} luminosity bin.

Table 3.4: Best-fit parameters for the two-point angular correlation functions calculated for SKADS $S_{1.4\mathrm{GHz}}$ sources, split by radio luminosity into three separate bins: $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, $10^{23} < L_{1.4\mathrm{GHz}}$ W Hz^{-1} $< 10^{26}$ and $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}. Also, the corresponding correlation length r_0 and galaxy bias $b(z)$ values. All uncertainties quoted are 1σ, obtained from χ^2_{min} fitting.

SKADS Sample	N_{obj}	z_{med}	A($\times 10^{-2}$)	γ	χ^2_{red}	$r_0(h^{-1}\mathrm{Mpc})$	$b(z_{\mathrm{med}})$
$L_{1.4\mathrm{GHz}} < 10^{23}$ WHz^{-1}	5,955	0.11	$2.27^{+0.33}_{-0.30}$	$1.70^{+0.05}_{-0.06}$	3.26	$3.47^{+0.39}_{-0.37}$	$0.79^{+0.07}_{-0.07}$
$10^{23} < L_{1.4\mathrm{GHz}}$ WHz^{-1} $< 10^{26}$	22,484	0.63	$0.30^{+0.06}_{-0.05}$	$2.13^{+0.05}_{-0.05}$	2.09	$7.19^{+0.95}_{-0.88}$	$1.66^{+0.30}_{-0.25}$
$L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}	4,835	2.03	$0.25^{+0.05}_{-0.05}$	$2.88^{+0.04}_{-0.04}$	4.86	$20.21^{+1.83}_{-1.80}$	$24.69^{+10.02}_{-6.81}$

3.3.4 Evolution of Clustering with Redshift

In Figures 3.7 and 3.8 we compare the correlation length and galaxy bias values measured for the JVLA samples at various luminosity cuts, to the r_0 and $b(z)$ evolution models calculated by SKADS, based on prescribed halo masses for the different types of radio galaxies. The trends in both the eastern and western fields are found to be very similar. The evolution of the r_0 values measured for the JVLA samples follows a steep trend with redshift, and does not follow any of the SKADS evolution models.

The r_0 evolution of the JVLA luminosity samples does not follow the same evolution pattern as that measured for the SKADS luminosity samples. The comparison of the correlation length values measured for the luminosity bin samples with that of the full cross-matched samples, show lowered r_0 values for the full cross-matched sample, i.e. these values do not follow the same trend in redshift laid out by the luminosity samples. This demonstrates how the clustering signal is diluted in a sample in which the sources span a large range in redshift, i.e. sample a large volume. The results for the measured bias of all JVLA luminosity samples, closely follow the bias evolution model of FRI type AGN, which according to Table 3.3, should make up majority of the source population at this radio flux-density limit ($S_{1.4} > 440\mu$Jy). Evidence for the flattening of the bias at high redshift ($z \sim 2$) is shown, in agreement with trends in the evolution of the bias found by Hale et al. (2018) for a sample of powerful AGN. These results suggest that the truncation of the AGN bias models (at $z = 1.5$) by Wilman et al. (2008) may be justified. The observed evolution of both the correlation length and bias is driven by the dominance of different radio source populations at different luminosity cuts. That is, the low luminosity bin is expected to be dominated by SFGs while the mid-luminosity bin is thought to comprise of a greater number of AGN, and we would expect the high luminosity to comprise solely of high-redshift AGN.

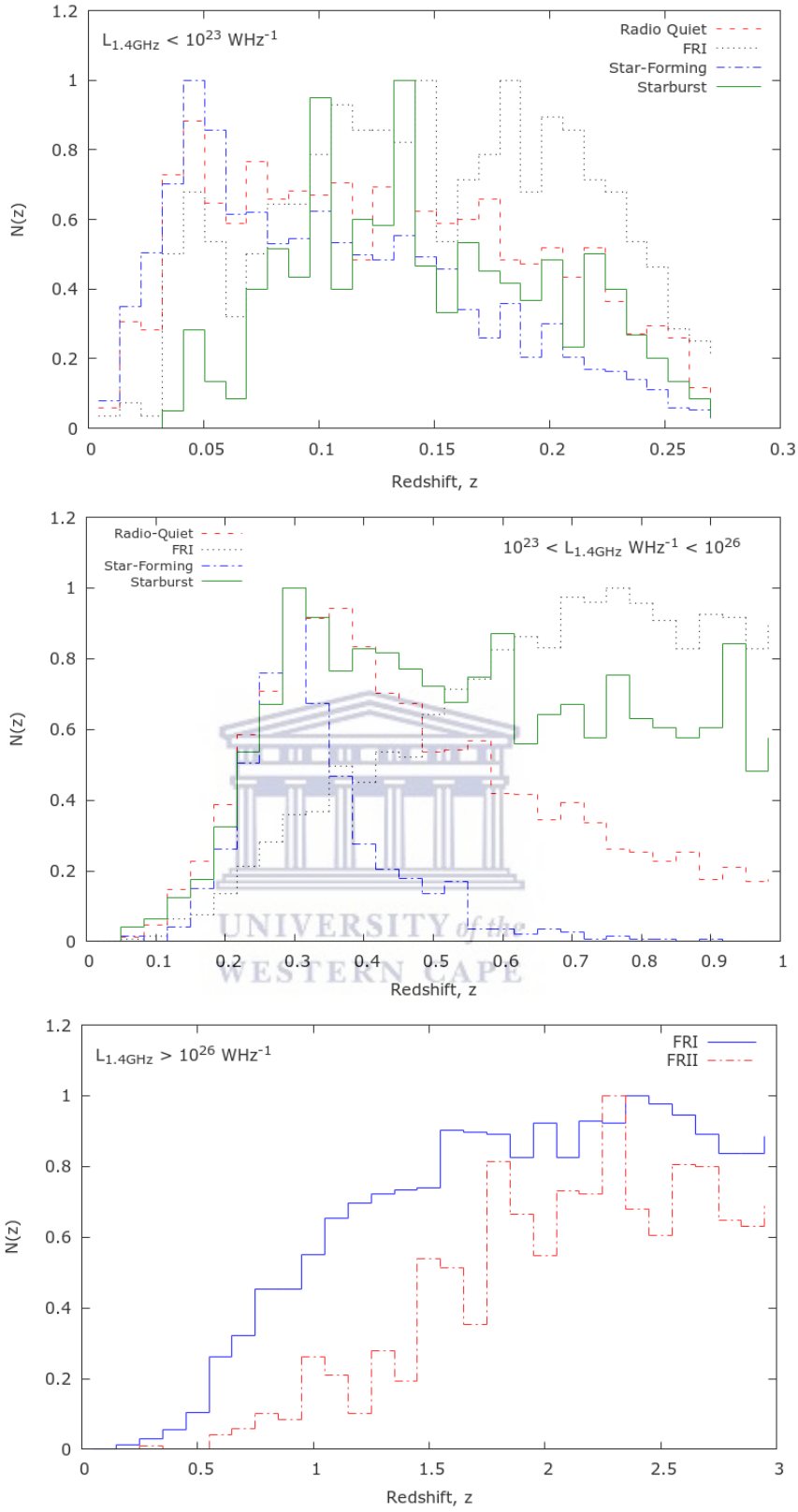


Figure 3.5: The normalised redshift distributions of the S^3 radio sub-populations in the three radio luminosity bins, i.e. from top to bottom: (i) $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, (ii) 10^{23} W Hz^{-1} $< L_{1.4\mathrm{GHz}} < 10^{26}$ W Hz^{-1} and (iii) $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}.

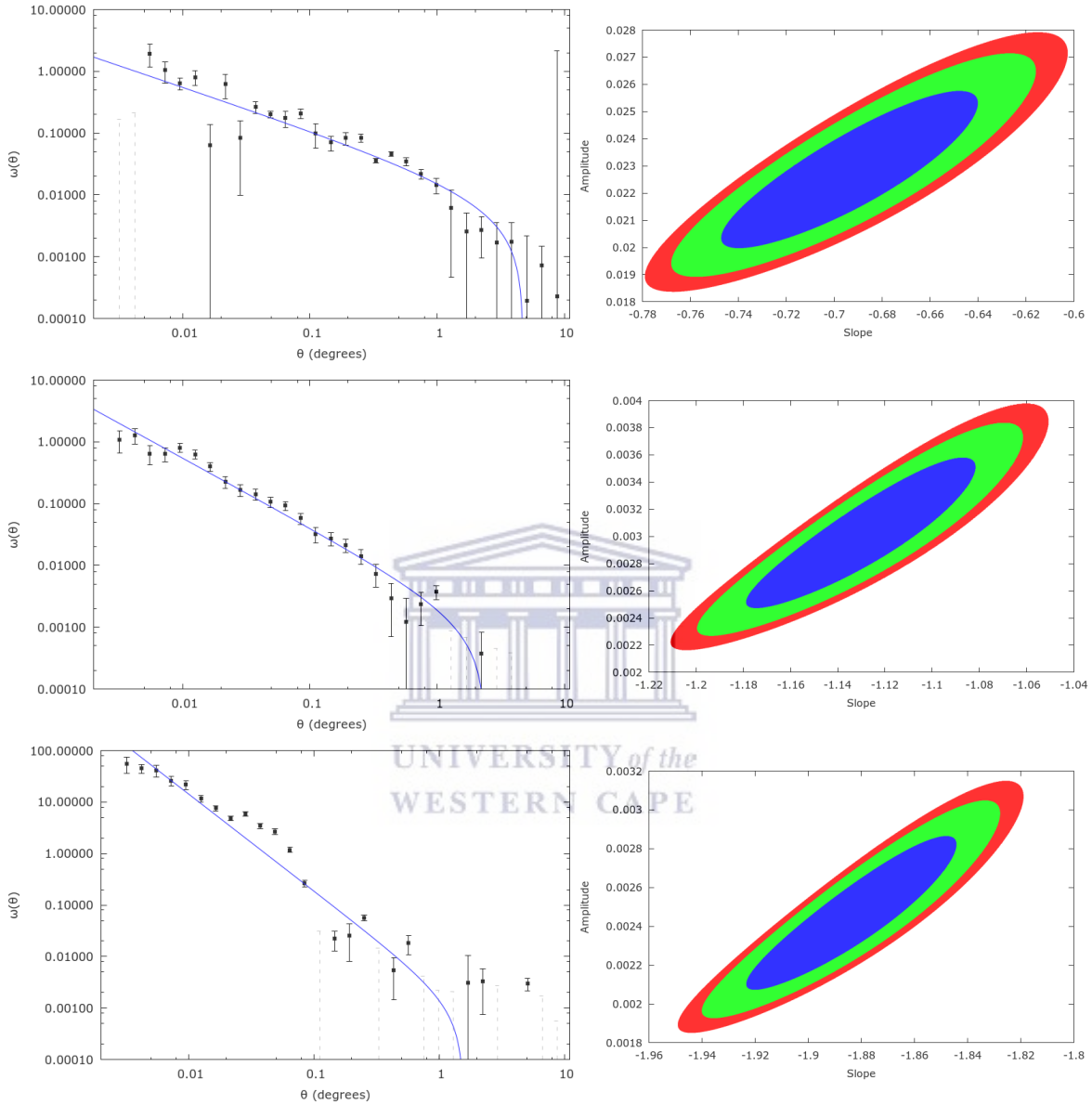


Figure 3.6: Left: The angular correlation function of SKADS $S_{1.4\mathrm{GHz}}$ sources, which have been split into three bins in luminosity. From Top to Bottom these are: $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}, $10^{23} < L_{1.4\mathrm{GHz}}$ W Hz^{-1} $< 10^{26}$ and $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Right: The 68%, 90% and 95% confidence levels for the large-angle single power law χ^2_{min} fits.

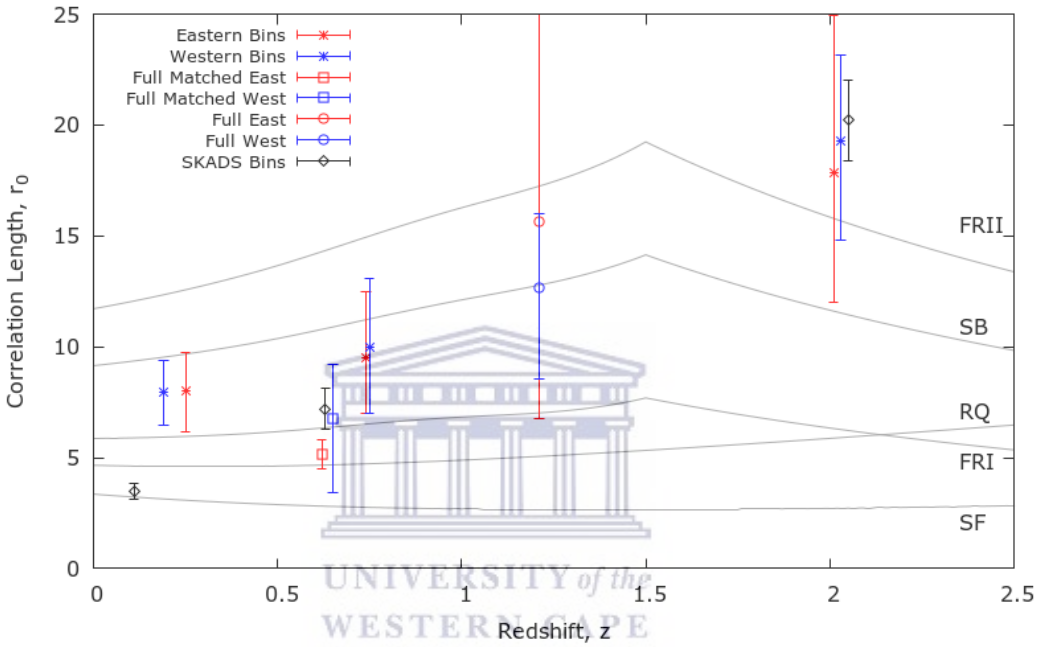


Figure 3.7: Measured correlation length (r_0) values for both the eastern and western fields of the JVLA Stripe 82 Snapshot Survey, compared to SKADS models for the evolution of r_0 with redshift, for various radio source types, denoted by the grey lines. The correlation length values for the eastern field are shown in red and those for the western field shown in blue. The asterisks correspond to the values for the samples split by luminosity, the open squares correspond to the full samples of matched sources, the open circles correspond to the full data samples and the black diamonds correspond to the SKADS luminosity bins.

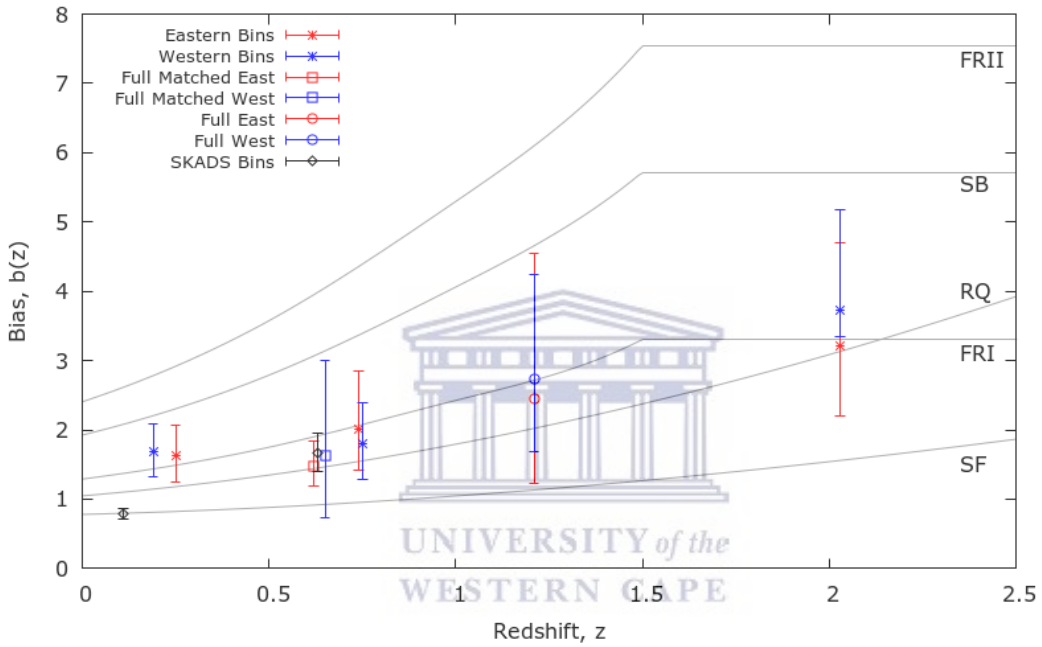


Figure 3.8: Measured galaxy bias ($b(z)$) values for both the eastern and western fields of the JVLA Stripe 82 Snapshot Survey, split into three bins by luminosity, compared to SKADS models for the evolution of $b(z)$ with redshift, for various radio source types, denoted by the grey lines. The symbol representations are the same as those described in Figure 3.7, however the high bias measured for the SKADS high luminosity sample is omitted.

3.4 Conclusions

In this chapter I measure the clustering properties of a sample of JVLA sources having z < 1 SDSS matches, split into a low luminosity ($L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}) and mid-luminosity ($10^{23} < L_{1.4\mathrm{GHz}}$ WHz$^{-1} < 10^{26}$) bin. The high luminosity bin ($L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1}) is comprised of all JVLA sources lacking optical matches, as these sources are assumed to comprise of very high radio luminosity, high redshift sources, due to the fact that these sources would be unlikely to have an optical match. For the low luminosity bin we find evidence of sample which is more clustered than the general galaxy population, with values of $r_0 = 8.02^{+1.73}_{-1.88}$ h^{-1}Mpc; $b(z) = 1.62^{+0.45}_{-0.38}$, and $r_0 = 7.98^{+1.40}_{-1.53}$ h^{-1}Mpc; $b(z) = 1.68^{+0.40}_{-0.35}$ for both the eastern and western field samples respectively. The measured clustering properties of the $L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1} SKADS sample are found to be $r_0 = 3.47^{+0.39}_{-0.37}$ h^{-1}Mpc and $b(z) = 0.79^{+0.07}_{-0.07}$. These low clustering values, as well as the redshift distribution plots from Figure 3.5, suggest that the SKADS simulations underestimate the redshift range of low luminosity sources at this flux limit, as there are no low luminosity sources above $z \sim 0.3$ in the SKADS data. The redshift distribution of the JVLA low luminosity sample (see Figure 3.1) shows sources which occupy a higher redshift range up to $z \sim 0.7$. These results suggest that even though the low luminosity sample is dominated by star-forming galaxies, there are likely to be a greater number of lower luminosity AGN present than modeled by SKADS, with stronger clustering at higher redshift. A similar conclusion is reached in work by Lindsay et al. (2014a) for the clustering analysis of a low luminosity ($10^{23} < L_{1.4\mathrm{GHz}}$ WHz$^{-1} < 10^{23.5}$) sample of $S_{1.4\mathrm{GHz}} > 90$ μJy JVLA sources.

In the mid-luminosity bin, I find evidence for an upturn at small ($\theta \sim 0.02$ deg) angular scales. This is possibly due to the 1-halo term of the HOD model, i.e. from the clustering of galaxies which occupy the same dark matter haloes. Some of the upturn may also be due to false identifications of host galaxies with components of extended radio sources during the cross matching process, leading to increased clustering at small scales. Blake & Wall (2002a) find that even a tiny fraction of resolved extended radio

galaxies can greatly affect the small angle clustering measurements. The large-scale clustering properties measured for the $10^{23} < L_{1.4\mathrm{GHz}}$ WHz$^{-1} < 10^{26}$ luminosity bin are found to be $r_0 = 9.53^{+2.93}_{-2.52}$ h^{-1}Mpc; $b(z) = 2.01^{+0.85}_{-0.59}$ and $r_0 = 9.98^{+3.10}_{-2.98}$ h^{-1}Mpc; $b(z) = 1.80^{+0.60}_{-0.51}$ for the eastern and western samples, respectively. These values are in line with those previously found for samples of AGN and LRGs (Zehavi et al. 2004; Magliocchetti et al. 2004; Magliocchetti et al. 2017; Hale et al. 2018), which trace a more massive, and therefore more biased, galaxy population. The results are similar to those found for a SKADS sample at this luminosity cut. The fact that the clustering properties measured for the JVLA mid-luminosity samples correspond with those measured for AGN, supports the estimate from SKADS that a sample at this luminosity cut would be FRI dominated ($\sim 80\%$).

The sensitivity of the JVLA Stripe 82 Snapshot survey has allowed for the study of a high luminosity, higher-redshift population of AGN, which are found to have high clustering values of $r_0 = 17.84^{+7.13}_{-5.86}$ h^{-1}Mpc, $b(z) = 3.21^{+1.49}_{-1.01}$ and $r_0 = 19.30^{+3.88}_{-4.48}$ h^{-1}Mpc, $b(z) = 3.72^{+1.45}_{-0.38}$ for the eastern and western samples, respectively. These clustering measurements are indicative of a highly clustered sample, which supports the notion that the radio population at $L_{1.4\mathrm{GHz}} > 10^{26}$ W Hz^{-1} is comprised solely of AGN sources, including powerful FRII galaxies. These high luminosity, high redshift galaxies would inhabit the most massive dark matter haloes (Mandelbaum et al. 2009; Donoso et al. 2014). The clustering analysis of the SKADS high luminosity sample yields a very high bias value of $b(z) = 24.69^{+10.02}_{-6.81}$, due to the steep slope of the measured ACF. It is therefore suggested that the estimate by Wilman et al. (2008) of the halo mass required to host FRII galaxies ($10^{14}\mathcal{M}_{\odot}$) is too high or that there exists some other error in the clustering models of SKADS at high redshift. The results of the evolution in the bias measured for the JVLA luminosity binned samples are found to closely follow the SKADS bias model for FRI galaxies. We conclude that sources with higher radio luminosity occupy higher redshifts and are found to be highly clustered and highly biased, suggesting that they are excellent tracers of the dark matter distribution in the Universe.

Chapter 4

Clustering properties of various radio sub-populations

In previous chapters we could, from the measured clustering properties of samples split by either redshift or radio luminosity, extract information about the clustering properties of different types of radio populations. This was guided by comparisons with the estimated clustering evolution models from the S^3 simulations. In this chapter, we are able to analyse samples of radio galaxies which have been cross-matched with optical host galaxies and classified into various radio source types, in order determine the clustering properties and evolution of various radio source populations. We continue to use data from the JVLA 1.4 GHz Snapshot Survey described in Section 2.1. A clustering analysis is performed for a sample of 1420 sources which have been classified as either star-forming galaxies or AGN, with a portion of the AGN further classified as either high or low excitation radio galaxies, i.e. HERGs and LERGs.

4.1 HERGs and LERGs

Radio AGN have been found to have different properties, some of which can be explained by the effect of the galaxy's orientation (Barthel 1989; Antonucci 1993). However other observations, such as galaxies which are radio-loud, but lack the high-excitation lines in optical spectra that are usually characteristic of AGN (Hine & Longair 1979b), require further explanation. Some AGN observations do not show evidence of accretion related emission in the X-ray (Hardcastle et al. 2006), or infrared emission from a torus (Ogle et al. 2006). Based on these observations, it is now widely accepted that there exist two fundamentally different types of AGN; Those which have an accretion disk, dusty torus and high-excitation lines, i.e. high-excitation radio galaxies (HERGs), and those which do not display such characteristics, i.e. low-excitation radio galaxies (LERGs). These differences are attributed to the fact that there must exist two different types of AGN accretion (Best et al. 2005b, Hardcastle et al. 2006). HERGs efficiently accrete cold gas, via an accretion disk, at a relatively high rate (1-10 % of the Eddington rate) (Elvis et al. 1994; Narayan 2005; Ho 2008; Heckman & Best 2014). HERGs also radiate efficiently across a broad range of the EM spectrum, where most of the energy comes from photons emitted from in and around the accretion disk. HERGs are often also associated with star-formation activity, due to the supply of cold gas (Kauffmann et al. 2003; Herbert et al. 2011; Hardcastle et al. 2013; Gürkan et al. 2015). HERGs are therefore thought to be more prevalent at earlier epochs, where the higher rate of mergers and interaction between galaxies may be a plausible explanation for the supply of cold gas. On the other hand, LERGs accrete hot gas relatively slowly, via advection-dominant flow which occurs when the accretion rate is typically below 1% of the Eddington rate (Best & Heckman 2012; Fernandes et al. 2015). In these cases, the accretion of material onto the black hole does not produce a large amount of radiated energy, but often leads to radio jets being produced (Merloni & Heinz 2007). The accretion process and stellar properties of the host galaxy is then self-regulated by mechanical feedback from the radio jets, hence the term radio-mode feedback. Due to the observed characteristics of these AGN types, HERGs and LERGs are also often referred to as cold mode or radiative mode and hot mode or jet mode, respectively.

It is understood that the two different modes of accretion are strongly influenced by environment (Hardcastle 2004; Janssen et al. 2012; Gendre et al. 2013). LERGs are thought to be present in high-mass passive halos (Best & Heckman 2012; Hale et al. 2018) in hotter, denser environments, whereas HERGs are thought to reside in lower mass (possibly younger) haloes in less dense environments. The high temperature shock-heated gas in massive galaxies would cool slowly, thus leading to a slow accretion of gas onto the black hole. It had also been suggested that most FRI sources tend to be LERGs and that most FRII sources are found to be HERGs (Lin et al. 2010), although discrepancies have been found (Laing et al. 1994). This is likely, as FRI host galaxies are found in richer clusters than FRII galaxies, where as FRII host galaxies are usually found in smaller groups (Zirbel 1997). This is further supported in work by Hine & Longair (1979b), who found that radio galaxies in clusters, which are commonly FRI type galaxies, tend to exhibit only absorption lines or weak [OII] emission in their spectra, where as most radio galaxies with strong emission lines were not members of clusters. There is also a distinction in terms of the host galaxies of these sources in that HERGs are typically bluer with lower stellar masses (Smolčić et al. 2009b; Janssen et al. 2012), and have 4000 Å breaks[†] which are less pronounced than that of LERGs of the same mass and radio luminosity, indicating a younger stellar population (Kauffmann et al. 2003; Best & Heckman 2012). LERGs however, tend to be hosted by massive galaxies which have an older stellar population (Best et al. 2005a; Herbert et al. 2011; Pracy et al. 2013). Work by Best & Heckman (2012) also find that although HERGs and LERGs are both found over the full range of radio luminosities, LERGs seem to be dominant at lower radio luminosities and HERGs at higher radio luminosities, with HERGs becoming the dominant population above $L_{1.4\mathrm{GHz}} \sim 10^{26}$W Hz^{-1}.

4.2 Data

In this Chapter I use samples from the Stripe 82 1-2 GHz Jansky Very Large Array Snapshot Survey data, presented in Section 2.1. These are samples in which sources

[†]The 4000 Å break in a galaxy spectrum is caused by the blanket absorption of high energy radiation from metals found in stellar atmospheres, as well as by a lack of hot, blue stars.

were cross-matched with SDSS Stripe 82 coadded imaging data and subsequently classified as either SGFs or AGN (Prescott et al. 2018), with a portion of the AGN further classified as HERGs and LERGs (Whittam et al. 2018). The cross-matching process also makes use of a catalogue created by Hodge et al. (2011), from a survey which was carried out over the Stripe 82 field using the JVLA in A configuration, resulting in a sample with a mean rms noise of 52 μJy per beam and higher angular resolution of 1.8 arcsec. These data therefore allow for greater positional accuracy of the host galaxy, making it easier to identify the multiwavelngth counterparts. However, there are sources which appear as single point sources in the Hodge et al. (2011) catalogue, but are actually part of large, extended sources, which can only be seen when compared with the JVLA CnB data from Heywood et al. (2016). The cross-matched sources were then checked for optical spectra, allowing those with detections of the required emission lines to be classified as different radio source types. The optical spectra used are those from the SDSS I - IV surveys, as well as the Baryon Oscillation Spectroscopic Survey observations. Spectroscopic data from the SDSS 14th Data Release (SDSS DR14 - Abolfathi et al. 2018).

4.2.1 Cross-matching

The cross-matching of sources was carried out by the visual inspection of all 11 768 sources from the Heywood et al. (2016) catalogue, by a team of ∼10 people, including myself, using the script XMATCHIT described in Prescott et al. (2018). The images used for inspection were as follows: For each component in the JVLA CnB catalogue, two cutouts from the co-added SDSS r-band optical images, each with a different zoom level, were produced. Radio overlay contours of both the Heywood et al. (2016) data and Hodge et al. (2011) data were then added. The positions of the radio components from Heywood et al. (2016) and the SDSS sources were also shown, allowing the inspector to identify the most likely optical counterpart for each radio galaxy. Examples of these overlay images are shown in Figure 4.1. The cutout image for each radio component was examined by two people, and if a discrepancy arose, it would then be reviewed by a team of three after which a final classification would be given. Out of the 11,768 sources, a matched catalogue containing 4,795 unique sources was produced.

Out of these 4,795 sources, 2,799 had photometric optical matches with 1,997 having spectroscopic matches. Further details of the cross-matching process and the reliability and completeness of the matched catalogue can be found in Prescott et al. (2018).

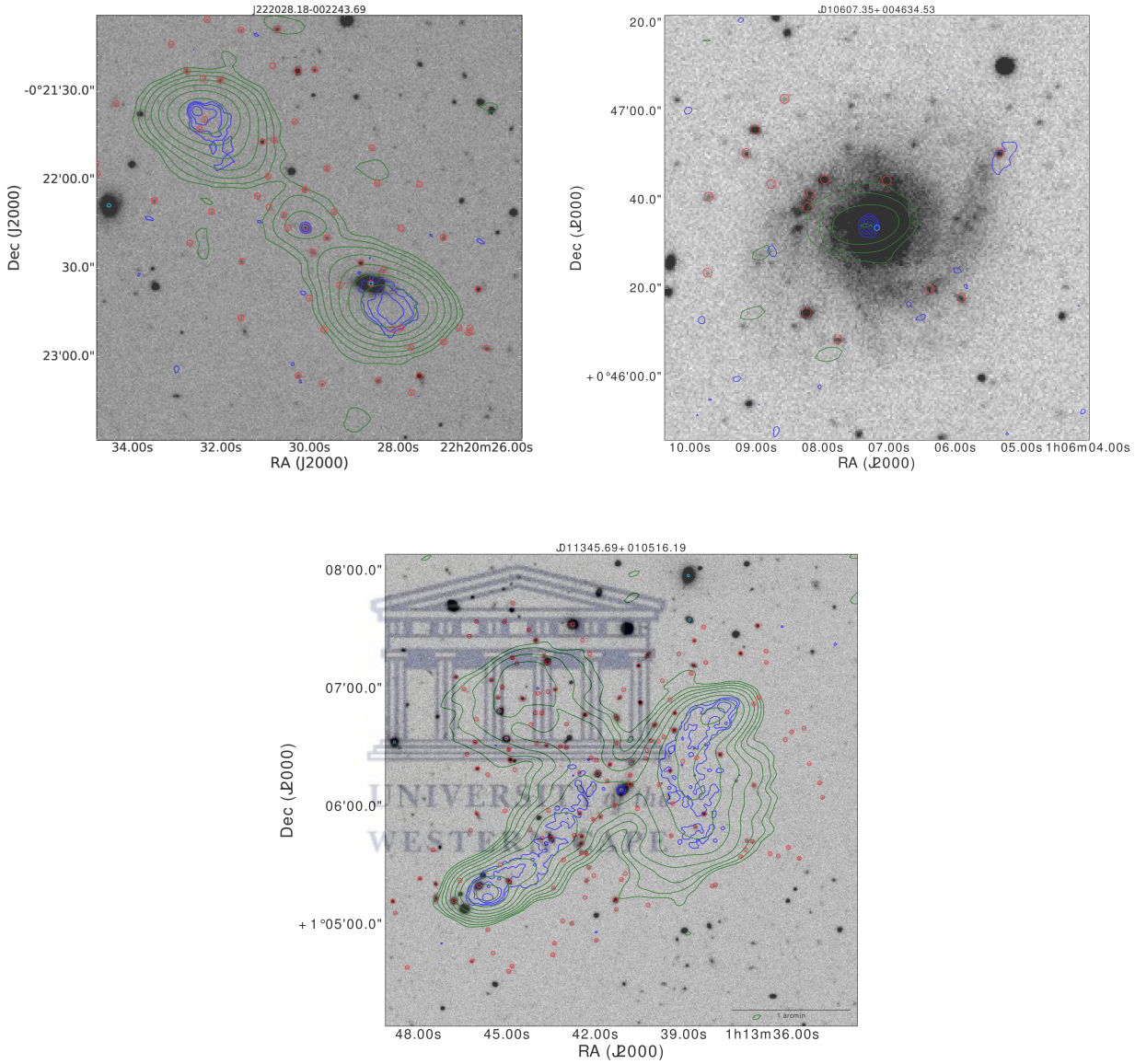


Figure 4.1: Examples of overlay images made by Prescott et al. (2018) to enable the cross-matching of the radio components from the JVLA Stripe 82 Snapshot Survey radio components and optical SDSS sources by visual inspection. The images are that of the co-added SDSS r-band data. Green contours represent the 1.4 GHz data from the Stripe 82 Snapshot Survey (Heywood et al. 2016), the blue contours represent radio data from the Hodge et al. (2011) survey and the red circles denote the positions of sources from SDSS. The light blue circles denote observed QSOs from SDSS. Top Left: Example of an FRII cross-match and classification. Top Right: Example of the cross-matching and classification of a nearby SFG. Bottom: Example of a rare, x-shaped, giant 'Green Bean' radio galaxy.

4.2.2 Classification

The classification of the JVLA CnB sources was carried out by Prescott et al. (2018). In order to classify HERG and LERG sources, the AGN were first separated from the star-forming galaxies in the sample using three diagnostics, following the method used by Best & Heckman (2012). The SDSS sample used for this task was limited to spectroscopic sources which have been classified as galaxies and which have redshifts of $z < 0.85$, to ensure that the required emission lines were available, this resulted in a sample of 1,518 sources. The diagnostics make use of star formation indicators such as the Balmer lines and the 4000 Å break, as well as stellar mass and radio luminosities. The resulting diagnostic plots from Prescott et al. (2018) are shown in Figure 4.2

• The first diagnostic used was the BPT diagram (Baldwin, Phillips & Terlevich - Baldwin et al. 1981), which uses the ratio of forbidden emission lines and the Balmer series in order to distinguish between AGN and SFGs (see top left panel of Figure 4.2). The dividing line, below which sources are classified as SFGs, is given by Equation 4.1 (Kewley et al. 2001). To ensure a thorough classification, objects were only included when each of the four emission lines had an S/N > 3.0. As a result, 660 sources were classified (456 as AGN and 204 as SFGs) using this diagnostic and the remaining 858 sources were left unclassified.

• The second diagnostic used was the Hα luminosity ($L_{\mathrm{H}\alpha}$) versus radio luminosity (L_{Radio}) method (Kauffmann et al. 2008). In the case of SFGs, the Hα and radio luminosities should correlate, as both are tracers of the star formation rate of a galaxy, where as AGN would be expected to have higher radio luminosities, as they produce an excess of radio emission. The two galaxy types were separated along the line given by Equation 4.2 (see bottom panel of Figure 4.2). For this diagnostic, only galaxies having an Hα flux detected with S/N > 3.0 are used. This allowed for the classification of 942 sources (587 as AGN and 355 as SFGs), thus the remaining 576 galaxies were unclassified using this method.

• The third diagnostic uses the D_{4000} versus $\mathrm{L}_{\mathrm{Radio}}/M_*$ method, i.e. the 4000 Å break

and the ratio of radio luminosity and stellar mass M_* of a galaxy, developed by Best et al. (2005b). In general, AGN would be expected to have a stronger 4000 Å break and large $L_{\rm Radio}/M_*$ value due to their excess radio emission. This diagnostic was used following the method of Best et al. (2005b). D_{4000} measurements were available for 885 galaxies, allowing for a classification of 577 AGN and 308 SFGs, leaving 633 as unclassified (see top right panel of Figure 4.2).

Due to the fact that an object may be classified as an AGN from one diagnostic and may not be classified as such by the others, there are 27 possible classification combinations for each object. From these combinations, an overall classification is assigned using the same approach as Best & Heckman (2012). The final classifications yielded a sample of 340 SFGs and 1149 AGN sources.

$$\log_{10}([\text{OIII}]/\text{H}\beta) = 1.31 + 0.61 \times \log_{10}([\text{NII}]/\text{H}\alpha) + 0.05 \tag{4.1}$$

$$\log_{10}(L_{\text{H}\alpha}) = 1.12 \times \log_{10}(L_{\rm Radio}) - 16.5 \tag{4.2}$$

The classification scheme used by Whittam et al. (2018) to classify the AGN as either HERGs and LERGs also follows the same criteria as that of Best & Heckman (2012), which uses a combination of line ratios and [OIII] equivalent width. The criterion for lines to be considered a detection was that they have an Amplitude-over-Noise (AoN) > 1.5. The following five steps were used for the classification:

• The first was to calculate the Excitation Index (EI- Buttiglione et al. 2010), which is the relative intensity of low and high excitation lines, for the emission lines: [OIII]; Hα; Hβ; [NII]; [SII] and [OI]. Sources for which all six lines were detected and which had an EI value at least 1σ greater than 0.95, were classified as HERGs and those with an EI value at least 1σ below 0.95, were classified as LERGs.

• The second made use of the Kewley et al. (2006) versions of the BPT diagram (if the required lines were detected) and the Seyfert-LINER division. Sources $\geq 1\sigma$ from

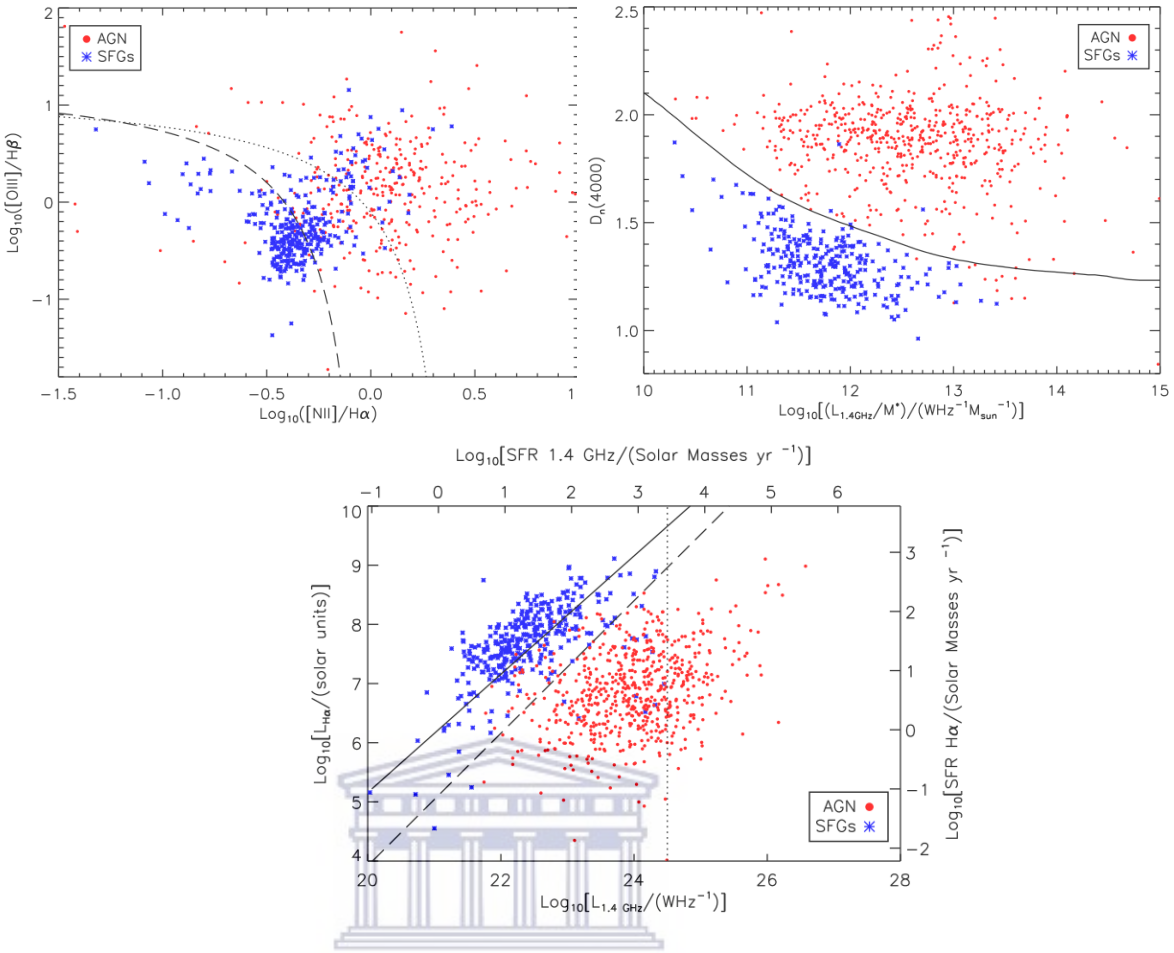


Figure 4.2: Diagnostic plots from Prescott et al. (2018) used to separate the cross-matched sample into AGN and SFGs. Top Left: BPT diagram for spectroscopic objects within $z < 0.85$, the dashed line from Kewley et al. (2001) separates AGN and SFGs. Here, the overall classifications from the three diagnostics are plotted. Top Right: Dn_{4000} against the L_{Radio}/M_* ratio for the spectroscopic sample. The solid line shows the track of SFGs by Bruzual & Charlot (2003) used to separate AGN and SFGs. Bottom: Plot of Hα luminosity vs radio luminosity where the dashed line shows the dividing line between AGN and SFGs. The vertical dotted line is at $\log_{10}(L_{Radio}/\mathrm{WHz}^{-1})= 24.5$, above which all objects should be AGN (Best & Heckman 2012). Once again the overall classifications from the three diagnostics are plotted. The black solid line indicates where estimates of the SFRs, determined from both the radio and Hα luminosities, are equal.

the Seyfert-LINER dividing line and in the Seyfert region were classified as HERGs, and those in the LINER region as LERGs.

• The third step made use of the [OIII] equivalent width, where if the width was more than 1σ greater than 5 Å (the dividing value commonly used in literature), the source

was classified as a HERG.

• This fourth step involved the repetition of steps 1 through 3, without the 1σ restrictions.

• Finally, [NII]/Hα against [OIII]/Hα diagnostic (Cid Fernandes et al. 2010) is used (if the required lines are detected). If the AoN value for the [OIII] line was < 1.5, the flux value was then considered an upper limit, allowing additional sources to be classified as LERGs. Sources unable to be classified using the five outlined steps were classified as 'probable LERGs' if they had an [OIII] equivalent width < 5 Å and an [OIII] measurement where the S/N > 3.

The resulting classification of the 1,161 AGN were as follows: 60 HERGs, 149 LERGs, 600 probable LERGs and 81 QSOs (these were identified as such by the SDSS pipeline). 271 sources lacked the line measurements required to be classified and thus remained unclassified. Including the probable LERG classifications, LERGs are the dominant population, making up 61 percent of the full classified sample (including the star-forming galaxies) and 84 percent of the AGN sample. In terms of the redshift distribution, the HERGs are found to occupy higher redshifts than the LERGs, with the samples having a median redshift of 0.40 and 0.23 for the HERGs and LERGs, respectively. The probable LERG sample has a median redshift of 0.39, higher than that of the definite LERG sample. This is most likely due to the fact that in the case of the higher-redshift LERGs, the emission lines required for definitive classification are more difficult to detect. For our analysis, we use the samples of sources classified as HERGs, LERGs, probable LERGs and SFGs. The total classified sample spans a redshift range of $0.01 < z < 0.7$ and luminosity range of 10^{21} W Hz^{-1} $< L_{1.4\mathrm{GHz}} <$ 10^{27} W Hz^{-1}.

4.3 ACF Results and Discussion

In this section I present the results of the clustering analysis of JVLA sources which have been cross-matched with SDSS host galaxies and thereafter classified as LERGs, probable LERGs, HERGs and SFGs. The angular correlation function is measured for the following three samples: A combined sample of LERGs and probable LERGs, a combined sample of LERGs, probable LERGs and HERGs (hereafter referred to as the LERG and HERG sample) and a sample of SFGs. For each sample, the angular correlation function is measured for the full sample, as well as the sample split into bins in redshift, so that the evolution of the clustering of these various radio populations can be studied. Due to the very low source count relative to other correlation function analyses in this Thesis, combined samples are used to allow for better ACF measurements and thus better constraints on the resulting clustering. The redshift distribution of the sources classified as LERGs, HERGs and SFGs are shown in Figure 4.3, where the LERG distribution shown is inclusive of the sources classified as probable LERGs. The large majority of star-forming galaxies occupy lower redshifts as expected, while the AGN (LERGs and HERGs) occupy a broad range in redshift with a peak at around z$\sim$0.4. A comparison between the redshift distributions of sources classified as LERGs and those classified as probable LERGs is shown in Figure 4.4. The distribution of sources which are confidently classified as LERGs is skewed to lower redshifts, due to the fact that the number of AGN for which the spectroscopy required for a thorough classification is available, is larger at lower redshifts. The LERG and HERG samples were split into three bins in redshift, where as the clustering evolution of the SFGs is investigated using only two lower bins in redshift, due to 80% of the sources lying below z$\sim$0.15.

Figures 4.5 - 4.10 which present all measured ACFs, show that all ACFs are in the form of a single power law, as expected, as our samples consist only of reliably cross-matched and classified sources, thus there is no measure of small-scale clustering from the components of extended radio sources.

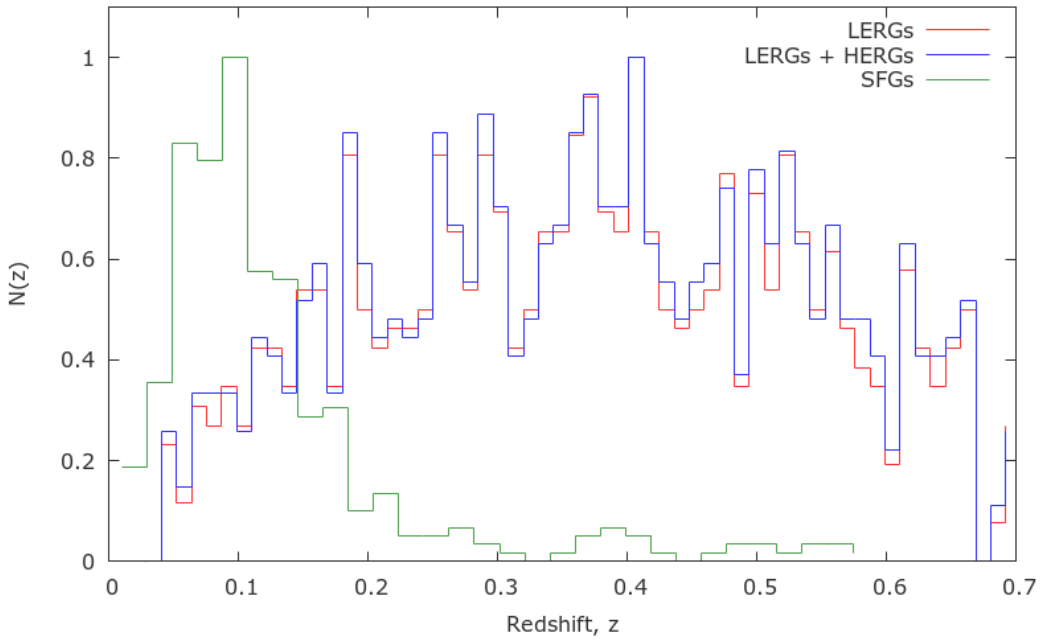


Figure 4.3: The normalised redshift distribution of the full sample of JVLA Stripe 82 sources classified as either LERGs, HERGs or star-forming galaxies, where LERGs include sources classified as probable LERGs.

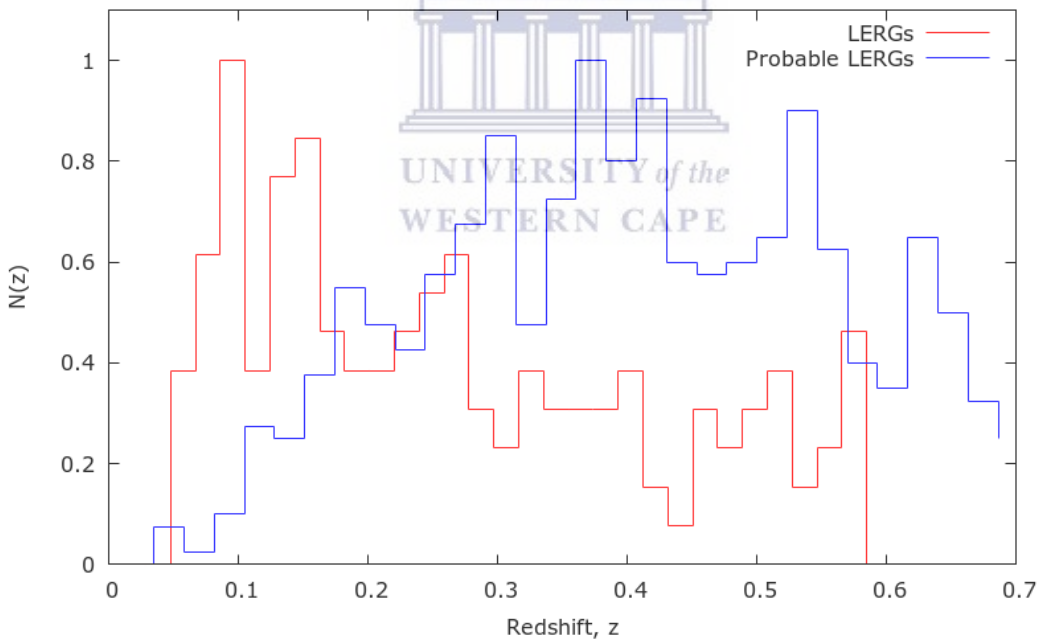


Figure 4.4: A comparison of the normalised redshift distributions of JVLA Stripe 82 sources classified as LERGs and those classified as probable LERGs.

4.3.1 Combined LERG and probable LERG Sample

The best-fit parameters to the ACF measured for the combined sample of LERGs and probable LERGs are presented in Table 4.1. The clustering amplitudes have values of

$A = 3.40^{+2.03}_{-1.47} \times 10^{-2}$, $A = 4.31^{+1.74}_{-3.76} \times 10^{-2}$ and $A = 7.36^{+4.14}_{-3.33} \times 10^{-2}$ for redshift bins $z < 0.3$, $0.3 \leq z < 0.45$ and $0.45 \leq z \leq 0.7$, respectively. These measured clustering amplitudes, especially those measured for the ACFs of the second and third redshift bins, are large compared to those measured in previous chapters, which have mostly been of the order $\sim 10^{-3}$. These clustering amplitudes are also larger than those measured for the ACFs presented by Lindsay et al. (2014a), Lindsay et al. (2014b) and Hale et al. (2018), which are also of the order $\sim 10^{-3}$. These results are indicative of a strongly clustered sample. The measured γ values are ~ 2.0 for the full sample as well as the redshift bin samples, which are in agreement with those measured for radio populations in other work (Lindsay et al. 2014b; Magliocchetti et al. 2017). It is noted however, that the larger error bars due to the much lower source counts available for the analysis, relative to those used in previous chapters, make it harder to constrain the parameters of the measured ACFs.

The full LERG sample yields measured correlation length and bias values of $r_0 = 9.37^{+4.96}_{-5.10}$ and $b = 2.07^{+2.13}_{-1.17}$, respectively at $z_{med} = 0.37$, which agree well with the clustering measurements found by Hale et al. (2018) for LERG type galaxies at $z < 1$, i.e. $r_0 = 9.7^{+1.2}_{-1.3}$ and $b = 2.9^{+0.3}_{-0.3}$. These results are indicative of a strongly clustered sample with measured correlation lengths and bias values in keeping with that of AGN. This is also in keeping with previous findings that LERGs reside in large, red, late-type galaxies which reside in hot, dense environments (Hardcastle 2004; Gendre et al. 2013) and would thus be expected to occupy higher mass haloes. The correlation length and bias values measured for the first and second LERG redshift bins samples are very similar to those measured for the low ($L_{1.4\mathrm{GHz}} < 10^{23}$ WHz^{-1}) and high ($10^{23} < L_{1.4\mathrm{GHz}}$ W Hz$^{-1} < 10^{26}$) luminosity bins, in Chapter 3.

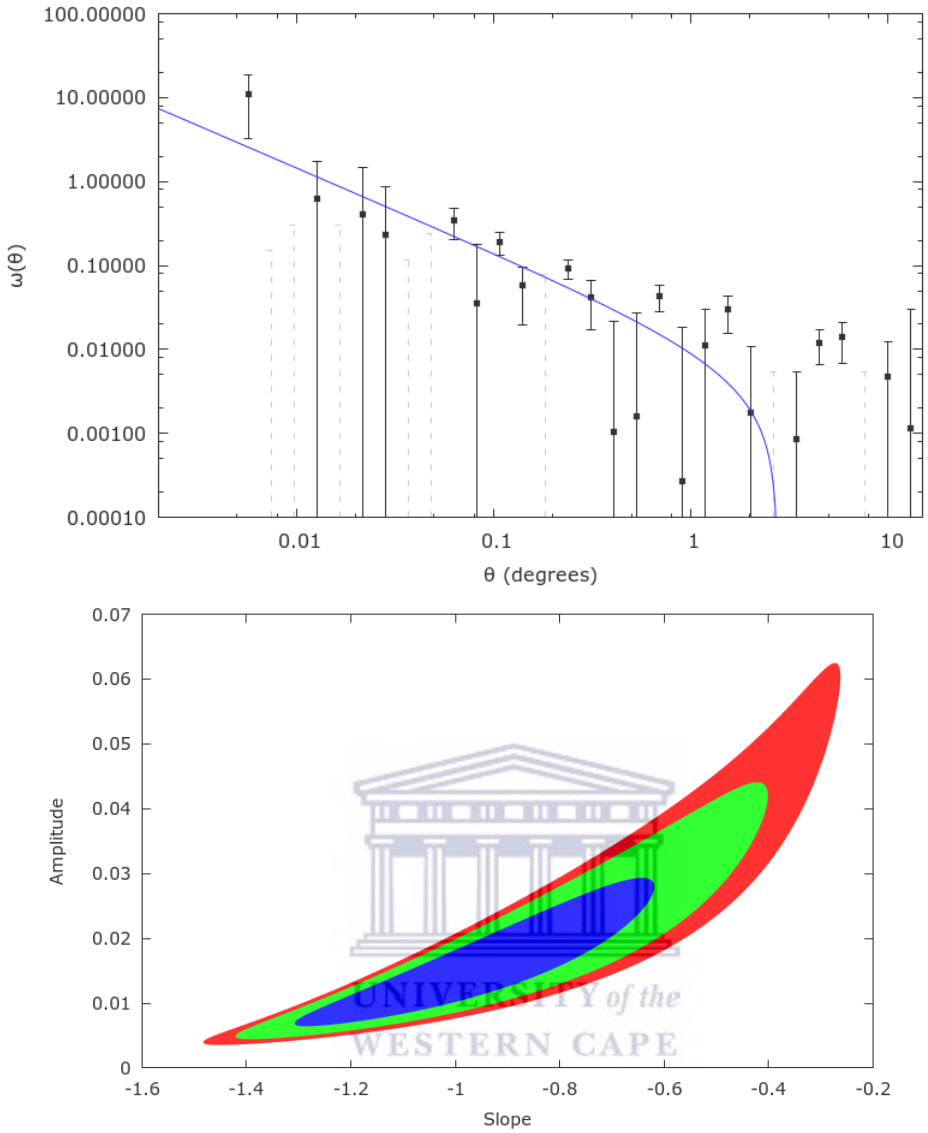


Figure 4.5: Top: The angular correlation function of a sample of sources which have been classified as either LERGs or probable LERGs. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the χ^2_{min} fits.

4.3.2 Combined LERG and HERG Sample

The measured ACFs for the combined LERG and HERG samples are presented in Figures 4.7 and 4.8, with the corresponding best-fit parameters and clustering properties presented in Table 4.2. Similar to those found for the combined LERG plus probable LERG sample, large clustering amplitude values of $A = 5.71^{+3.59}_{-2.43} \times 10^{-2}$

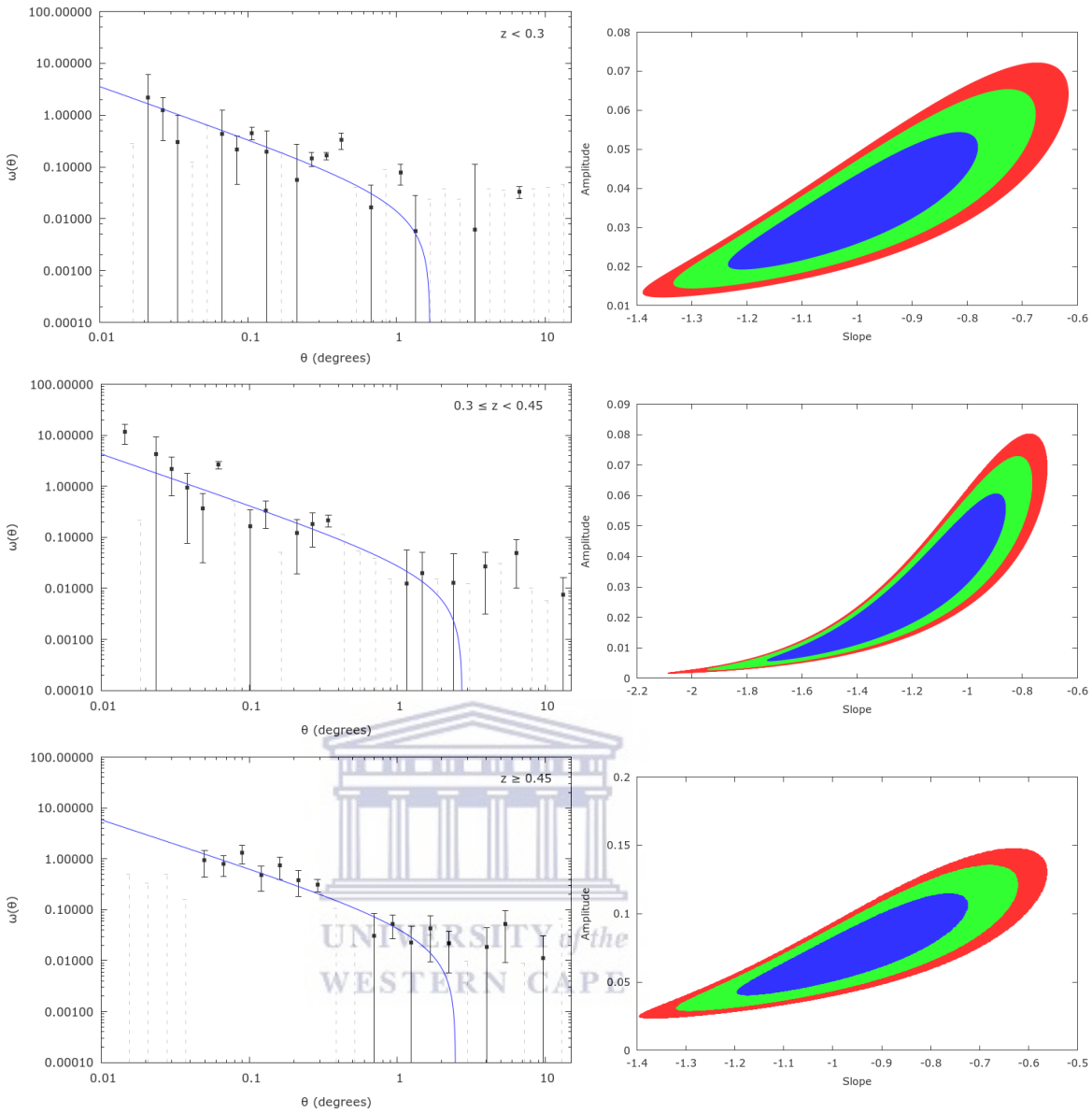


Figure 4.6: Left: The angular correlation functions of a sample of LERGs and probable LERGs, split into three bins in redshift. From Top to Bottom, the redshift bins are as follows: z $<$ 0.3, 0.3 $\leq$ z $<$ 0.45 and 0.45 $\leq$ z $<$ 0.7. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Right: The 68%, 90% and 95% confidence levels for the χ^2_{min} fits.

and $A = 6.02^{+5.38}_{-4.90} \times 10^{-2}$ are found for the second (0.3 $\leq$ z $<$ 0.45) and third (0.45 $\leq$ z $\leq$ 0.7) redshift bins, respectively, once again indicating that these are strongly clustered samples. The measured γ values are also similar to those found for the measured ACFs of the LERG samples with values of $\gamma = 1.94 - 2.22$ and are thus consistent with values expected for a radio source population. The full LERG plus

Table 4.1: Best-fit parameters of the fit to the two-point angular correlation functions calculated for sources classified as either LERGs or probable LERGs. The ACFs are that of the full sample, as well as the sample split into three bins in redshift. Also, the corresponding correlation length r_0 and galaxy bias b(z) values. All uncertainties quoted are 1σ uncertainties, obtained from χ^2_{min} fitting.

LERGs	N_{obj}	z_{med}	χ^2_{red}	A($\times 10^{-2}$)	γ	$r_0(h^{-1}\mathrm{Mpc})$	b(z_{med})
Full Sample	749	0.37	1.67	$1.39^{+1.53}_{-0.76}$	$2.01^{+0.30}_{-0.39}$	$9.37^{+4.96}_{-5.10}$	$2.07^{+2.13}_{-1.17}$
z < 0.3	280	0.20	3.65	$3.40^{+2.03}_{-1.47}$	$2.01^{+0.22}_{-0.23}$	$7.48^{+2.33}_{-2.61}$	$1.69^{+0.91}_{-0.65}$
$0.3 \leq$ z < 0.45	217	0.37	2.69	$4.31^{+1.74}_{-3.76}$	$2.00^{+0.72}_{-0.14}$	$9.24^{+0.12}_{-6.24}$	$2.03^{+2.39}_{-1.37}$
$0.45 \leq$ z ≤ 0.7	252	0.54	1.18	$7.36^{+4.14}_{-3.33}$	$1.95^{+0.25}_{-0.23}$	$11.93^{+1.04}_{-2.96}$	$2.50^{+0.83}_{-0.81}$

HERG sample yields measured correlation length and bias values of $r_0 = 10.88^{+6.58}_{-7.44}$ and $b = 2.30^{+3.00}_{-1.54}$, respectively at $z_{med}= 0.37$. These values measured are in agreement with those by Magliocchetti et al. (2017), who find clustering values of $r_0 = 10.9 \pm 0.5$ and $b = 3.9^{+1.0}_{-1.3}$ for their sample of AGN. The addition of HERGs to the LERGs plus probable LERGs sample sees a decrease in the measured correlation length and bias values for the first redshift bin, a slight increase in the second redshift bin, and very similar clustering values for the third redshift bin, however all values are consistent within the uncertainties. As shown from clustering analyses of HERG type galaxies by Hale et al. (2018), HERGs are expected to be less clustered than LERGs, to allow for efficient accretion in cooler, less dense environments (Ineson et al. 2015). This is supported by previous findings which show that HERGs tend to reside in host galaxies with lower mass than that of LERGs (Tasse et al. 2008; Smolčić et al. 2009b; Best & Heckman 2012), and would thus be expected to be less clustered. The differences in the LERG and LERG plus HERG clustering properties are indistinguishable within the uncertainties due to the very low number of HERGs (60) in the sample. In the case of both the LERG and LERG plus HERG samples, evidence of strong clustering is seen even at low median redshifts of $z_{med}= 0.2$ and $z_{med}= 0.37$ due to the sample being comprised only of reliably selected AGN.

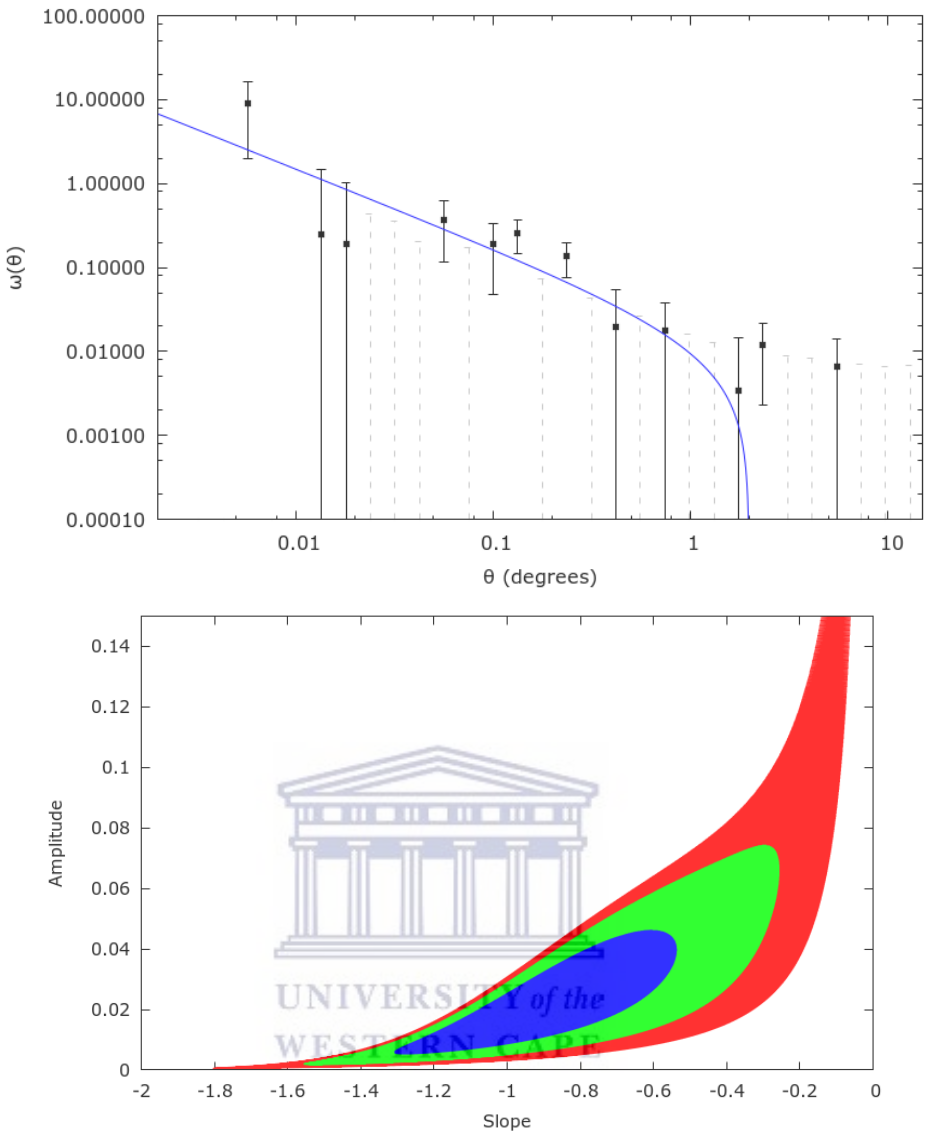


Figure 4.7: Top: The angular correlation function of a sample of sources which have been classified as either LERGs or HERGs. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the χ^2_{min} fits.

4.3.3 Star-Forming Galaxies

The measured ACFs for the star-forming galaxy samples are presented in Figures 4.9 and 4.10, with the best fit-parameters presented in 4.3. The values measured for the clustering amplitude of the ACFs for the two redshift bins are poorly constrained due to the low source counts in these bins, while the γ values measured are $\gamma \sim 2.2$ which is

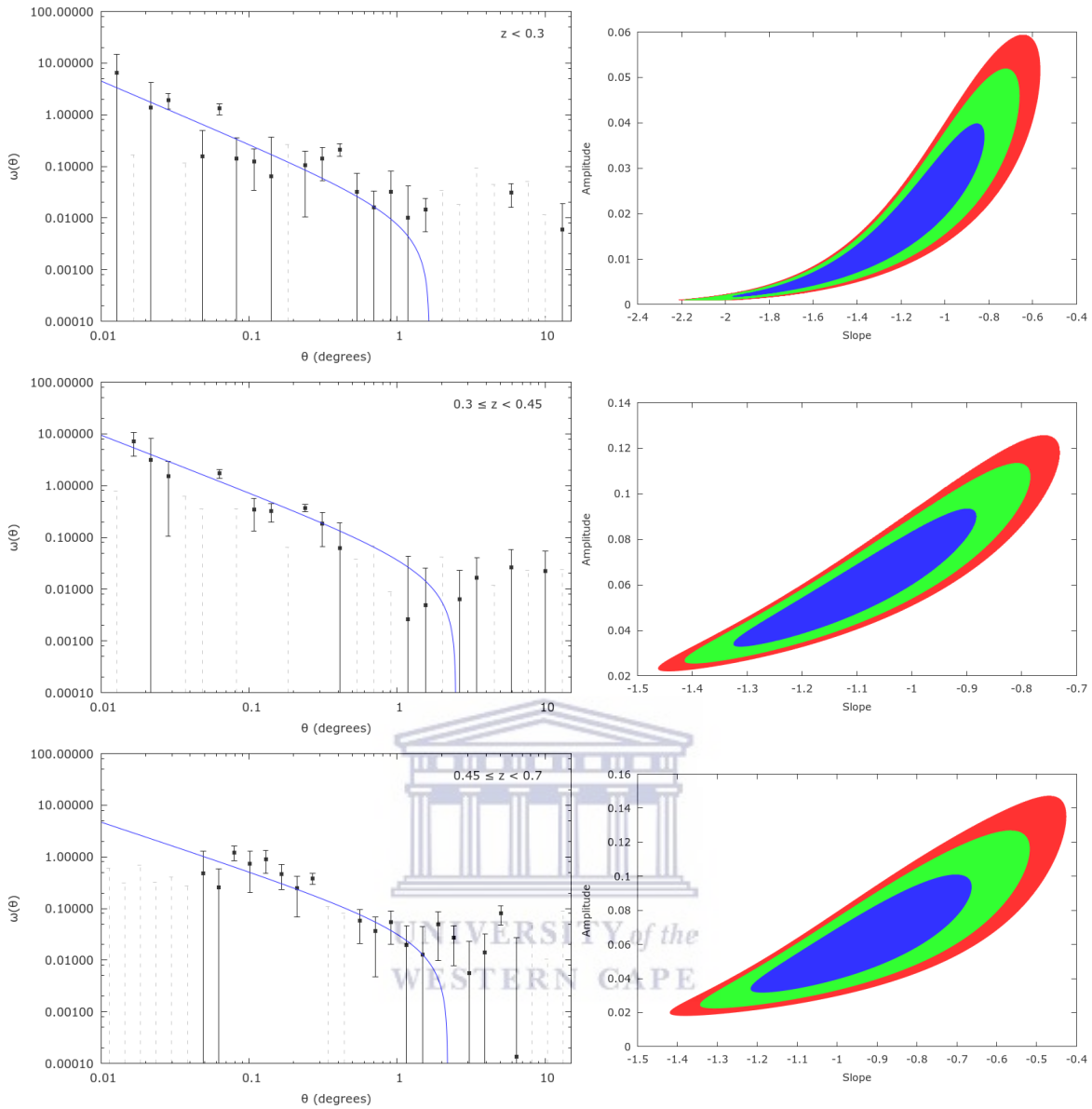


Figure 4.8: Left: The angular correlation functions of a sample of LERGs and HERGs, split into three bins in redshift. From Top to Bottom, the redshift bins are as follows: z $<$ 0.3, 0.3 $\leq$ z $<$ 0.45 and 0.45 $\leq$ z $<$ 0.7. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Right: The 68%, 90% and 95% confidence levels for the χ^2_{min} fits.

still consistent with those commonly found for radio sources. The full sample of star-forming galaxies yields measured correlation length and bias values of $r_0 = 3.81^{+1.25}_{-2.10}$ and $b = 0.91^{+0.70}_{-0.50}$, respectively at z_{med}= 0.09. These values are in line with a sample that is less strongly clustered and therefore likely to be less massive, as clustering strength has been shown to increase with stellar mass (Cochrane et al. 2018), which

Table 4.2: Best-fit parameters of the fit to the two-point angular correlation functions calculated for sources classified as either LERGs or HERGs. The ACFs are that of the full sample, as well as the sample split into three bins in redshift. Also, the corresponding correlation length r_0 and galaxy bias b(z) values All uncertainties quoted are 1σ uncertainties, obtained from χ^2_{min} fitting.

LERGs + HERGs	N_{obj}	z_{med}	χ^2_{red}	A($\times 10^{-2}$)	γ	$r_0(h^{-1}\mathrm{Mpc})$	b(z_{med})
Full Sample	809	0.37	1.13	$1.98^{+2.61}_{-1.43}$	$1.94^{+0.37}_{-0.40}$	$10.88^{+6.58}_{-7.44}$	$2.30^{+3.00}_{-1.54}$
z < 0.3	303	0.20	0.72	$1.63^{+2.34}_{-1.47}$	$2.22^{+0.74}_{-0.40}$	$5.67^{+3.06}_{-4.40}$	$1.40^{+8.28}_{-1.09}$
$0.3 \leq$ z < 0.45	228	0.37	1.33	$5.71^{+3.59}_{-2.43}$	$2.11^{+0.21}_{-0.23}$	$11.28^{+2.44}_{-3.00}$	$2.68^{+1.37}_{-0.97}$
$0.45 \leq$ z ≤ 0.7	278	0.55	1.82	$6.02^{+5.38}_{-4.90}$	$1.95^{+0.68}_{-0.45}$	$12.09^{+0.22}_{-8.36}$	$2.52^{+2.69}_{-1.72}$

in turn traces the halo mass (McCracken et al. 2015; Hatfield et al. 2016). The low correlation length is in agreement with that found in work by Dolley et al. (2014) for a sample of Spitzer star-forming galaxies at $z < 1$, making use of Spitzer mid-infrared data. The clustering values measured are also in line with those measured by Hale et al. (2018) for a sample of SFGs at z < 1. Hale et al. (2018) measure correlation length and bias values of $r_0 = 5.0^{+0.5}_{-0.6}$ and $b = 1.5^{+0.1}_{-0.2}$, respectively at $z_{\mathrm{med}} = 0.62$. These results show that SFGs are likely to reside in lower mass haloes than the general AGN population.

Table 4.3: Best-fit parameters of the fit to the two-point angular correlation functions calculated for sources classified as star-forming galaxies. The ACFs are that of the full sample, as well as the sample split into three bins in redshift. Also, the corresponding correlation length r_0 and galaxy bias b(z) values All uncertainties quoted are 1σ uncertainties, obtained from χ^2_{min} fitting.

Star Forming	N_{obj}	z_{med}	χ^2_{red}	A($\times 10^{-2}$)	γ	$r_0(h^{-1}\mathrm{Mpc})$	b(z_{med})
Full Sample	340	0.09	4.29	$2.52^{+2.61}_{-1.95}$	$2.19^{+0.37}_{-0.25}$	$3.81^{+1.25}_{-2.10}$	$0.91^{+0.70}_{-0.51}$
z < 0.07	115	0.05	0.87	$4.62^{+10.5}_{-3.85}$	$2.39^{+0.64}_{-0.53}$	$2.14^{+1.03}_{-1.42}$	$0.50^{+0.62}_{-0.32}$
$0.07 \leq$ z < 0.12	114	0.09	2.39	$11.3^{+8.70}_{-7.17}$	$2.25^{+0.33}_{-0.27}$	$4.07^{+0.54}_{-1.47}$	$1.01^{+0.45}_{-0.41}$

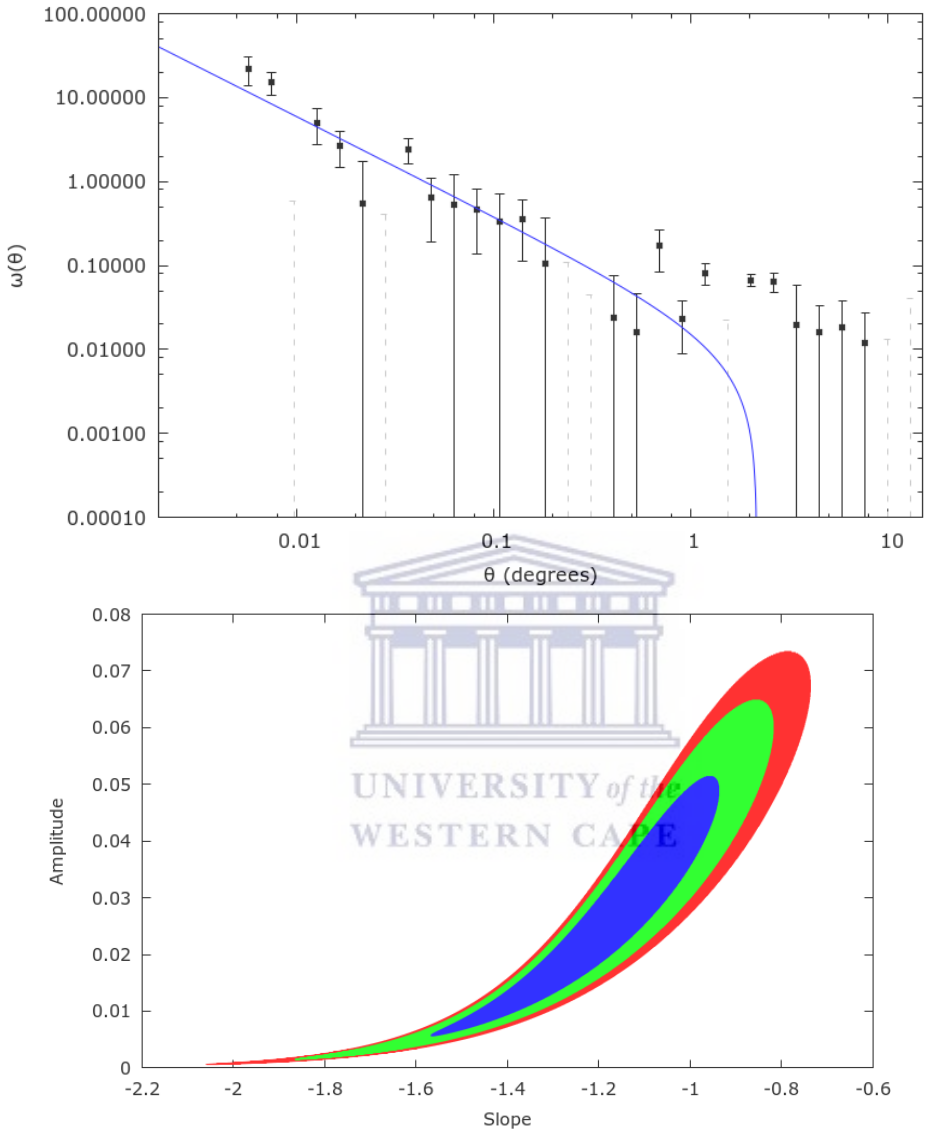


Figure 4.9: Top: The angular correlation function of a sample of sources which have been classified as star-forming galaxies. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Bottom: The 68%, 90% and 95% confidence levels for the χ^2_{min} fits.

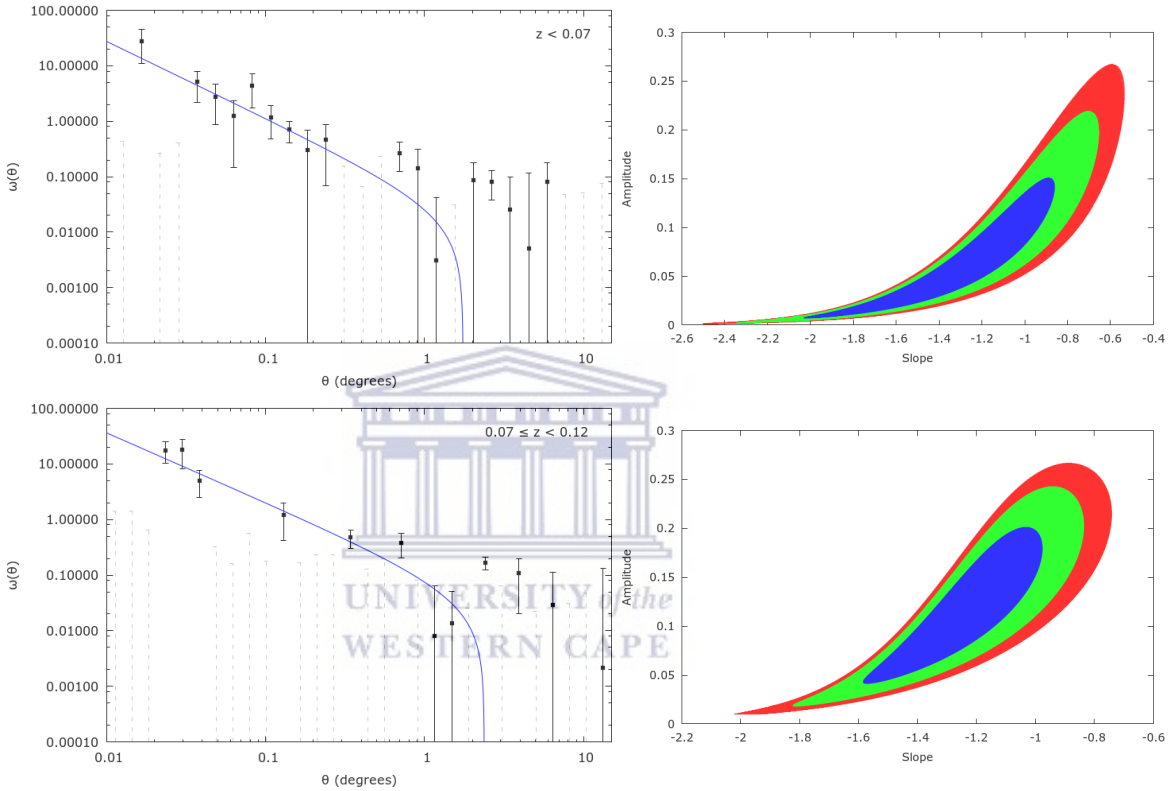


Figure 4.10: Left: The angular correlation functions of a sample of star-forming galaxies, split into two bins in redshift. From Top to Bottom, the redshift bins are as follows: z $<$ 0.07 and 0.07 $\leq$ z $<$ 0.12. The blue line represents the best fit power law to $\omega(\theta)$ minus the integral constraint, σ^2, and the grey dashed lines show the negative points of the ACF. Right: The 68%, 90% and 95% confidence levels for the χ^2_{min} fits.

4.3.4 Evolution of Clustering with Redshift

The r_0 and b(z) evolution of the LERG, HERG and SFG samples, which have been split into bins in redshift are shown in Figures 4.11 and 4.12 and are compared to clustering evolution models from the S^3 simulations. The results from clustering analyses of various radio source types by Hale et al. (2018), as well as Magliocchetti et al. (2017) are also included for comparison. The work by Hale et al. (2018) uses data from a 2-4 GHz JVLA survey over the COSMOS field, with a flux-density limit of $\sim$13μJy, in which the sources are classified as SFGs, radiatively inefficient (Moderate to Low Luminosity AGN; MLAGN) and radiatively efficient (High to moderate Luminosity AGN; HLAGN). MLAGN and HLAGN are thought to be roughly equivalent to LERGs and HERGs, respectively. Magliocchetti et al. (2017) perform a clustering analysis of sources from a $S_{1.4\mathrm{GHz}} > 0.15$ mJy JVLA sample, also over the COSMOS field, in which sources have been classified as AGN and SFGs.

An evolution of clustering with redshift is observed for our sample of SFGs, with measured r_0 and b(z) values of $r_0 = 2.14^{+1.03}_{-1.42}$ h^{-1}Mpc and $b = 0.50^{+0.62}_{-0.32}$, and $r_0 = 4.07^{+0.54}_{-1.47}$ h^{-1}Mpc and $b = 1.01^{+0.45}_{-0.41}$, respectively, for redshift bins z < 0.07 and $0.07 \leq$ z < 0.12. The correlation length and bias evolution of the SFGs in our sample and that of Hale et al. (2018) show a slow evolution in clustering and bias with redshift. This is in agreement with the S^3 evolution model, which shows a flatter evolution of the bias and r_0, as well as trends in evolution seen in work by Dolley et al. (2014). Results from Magliocchetti et al. (2017) show a steeper evolution of the clustering of SFGs from their work, as well as their comparisons with other work. Results by Dolley et al. (2014) suggest that the halo mass of SFGs can vary by 1-2 orders of magnitude and may be higher than those assumed by Wilman et al. (2008). Dolley et al. (2014) also deduce that majority of star forming galaxies at z < 1.0 are star-forming progenitors of blue galaxies in the local Universe.

For the LERG samples, an evolution of the clustering with redshift is observed, with measured r_0 and b(z) values of $7.48^{+2.33}_{-2.61}$ h^{-1}Mpc, $9.24^{+0.12}_{-6.24}$ h^{-1}Mpc, $11.93^{+1.04}_{-2.96}$ h^{-1}Mpc and $1.69^{+0.91}_{-0.65}$, $2.07^{+2.39}_{-1.37}$, $2.50^{+0.83}_{-0.81}$, respectively, for the three redshift bins, z

< 0.3, $0.3 \leq$ z < 0.45 and $0.45 \leq$ z < 0.7. A similar trend is seen for the combined sample of LERGs and HERGs, with clustering values of $5.67^{+3.06}_{-4.40}$ h^{-1}Mpc, $11.28^{+2.44}_{-3.00}$ h^{-1}Mpc, $12.09^{+0.22}_{-8.36}$ h^{-1}Mpc and $1.40^{+8.28}_{-1.09}$, $2.68^{+1.37}_{-0.97}$, $2.52^{+2.69}_{-1.72}$, respectively, for the same bins in redshift as the LERG samples. The bias values of LERGs and HERGs are shown to evolve with redshift at a similar rate to the S^3 models for starburst and FRI sources. However, our measurements, as well as those by Hale et al. (2018), suggest that the simulated bias evolution by Wilman et al. (2008) for FRI galaxies may be too low, as they do not align with that measured for HERGs or LERGs. The results from Hale et al. (2018) show that LERGs are found to be more highly clustered than HERGs, and the bias value for the Hale et al. (2018) LERG sample is seen to follow the same curve in evolution as our work, completing the trend at higher redshift. The measured correlation length and bias values in this Chapter, as well as those from Hale et al. (2018) suggest that the S^3 model for the bias evolution of FRII type AGN predicts larger r_0 and b(z) values for this radio subpopulation than those measured for HERGs. This is likely due to the fixed halo mass assigned by Wilman et al. (2008) to FRII type galaxies being larger than that which is deduced from observations. Higher redshift clustering measurements of HERG and LERG type sources carried out by Hale et al. (2018), which have been omitted from Figures 4.11 and 4.12 for the sake of clarity of the low redshift points, find evidence for the flattening in the bias evolution of HERGs at high redshift (z > 1.5), which suggests that at higher redshift, the halo mass required for hosting HERGs may be lower. This is possible due to the higher density of cold gas at higher redshift which is available for accretion. In contrast, the higher redshift clustering measurements of the LERG type sources from Hale et al. (2018) are shown to continue to increase.

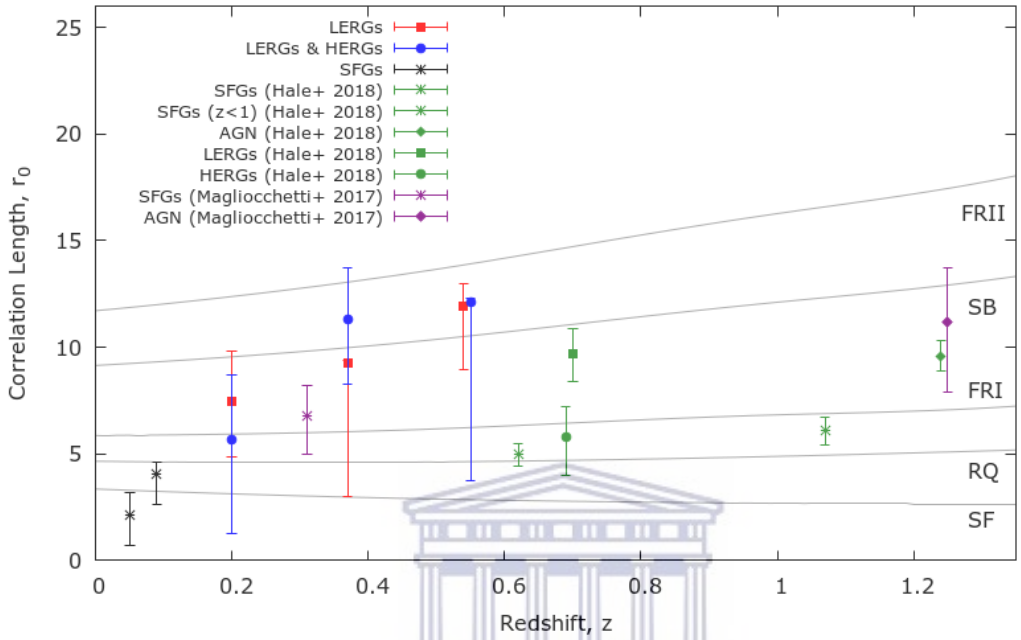


Figure 4.11: Evolution of the measured correlation length (r_0) values with redshift, for the combined samples of sources classified as LERGs and probable LERGs, LERGs and HERGs and star-forming galaxies. These are compared to S^3 models for the evolution of r_0 with redshift, for various radio subpopulations, denoted by the grey lines. The symbols corresponding to the measured r_0 values are as follows: The squares correspond to the bins in redshift of the combined LERG and probable LERG sample, the circles correspond to the bins in redshift of the combined LERG and HERG sample and the diamonds correspond to the bins in redshift of the sample of star-forming galaxies. Correlation length values measured by Hale et al. (2018) are shown in green and are represented as follows: the star symbol corresponds to SFGs at z < 1, the open triangle represents the full sample of SFGs, the open upside down triangle represents the full AGN sample, the solid triangle corresponds to LERG type galaxies and the solid upside down triangle corresponds to HERG type galaxies. The purple symbols represent r_0 values measured by Magliocchetti et al. (2017) and are defined as follows: the open diamond corresponds to a sample of SFGs and the cross corresponds to a sample of AGN.

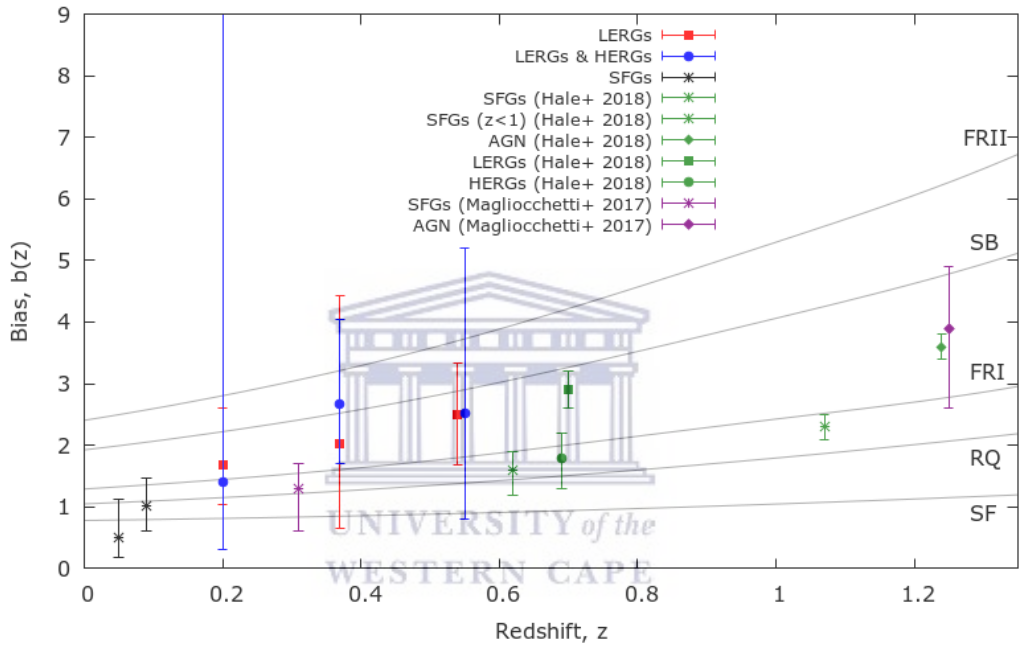


Figure 4.12: Evolution of the measured galaxy bias values with redshift, for the combined samples of sources classified as LERGs and probable LERGs, LERGs and HERGs and star-forming galaxies. These are compared to to S^3 models for the evolution of b(z) with redshift, for various radio subpopulations, denoted by the grey lines. The symbols corresponding to the measured b(z) values are as defined in Figure 4.11.

4.4 Conclusion

A clustering analysis is performed for a sample of 1149 radio sources from the Heywood et al. (2016) JVLA survey which were cross-matched with optical sources from co-added SDSS data and subsequently classified into different radio sub-populations. Of the 1149 sources, I measure the ACF for a combined sample of LERG (149) and probable LERG (600) sources and find correlation length and bias values of $r_0 = 9.37^{+4.96}_{-5.10}$ and $b = 2.07^{+2.13}_{-1.17}$, respectively at z_{med}= 0.37, for the full sample. This is indicative of a highly clustered sample which strongly traces the underlying dark matter distribution and is in agreement with previous clustering analyses of LERGs and AGN populations (Magliocchetti et al. 2004; Hale et al. 2018). Measuring the ACF of a combined sample of LERGs (LERGs plus probable LERGs - 749) and HERGs (60), we find clustering results of $r_0 = 10.88^{+6.58}_{-7.44}$ and $b = 2.30^{+3.00}_{-1.54}$ at z_{med}= 0.37, for the full sample. This shows an increase in the clustering properties from the LERG sample, however the addition of HERGs to the LERG sample yields clustering results which are indistinguishable from the LERG sample within the large uncertainties, thus we are unable to extract meaningful clustering information from the small HERG sample. An ACF analysis of a sample of star-forming galaxies (340) yields clustering measurements of $r_0 = 3.81^{+1.25}_{-2.10}$ and $b = 0.91^{+0.70}_{-0.50}$ at z_{med}= 0.09, for the full sample, indicating that SFGs are more weakly clustered than the AGN population, as found in previous work (Zehavi et al. 2011; Magliocchetti et al. 2013; Hale et al. 2018).

It is clear that there is great value in performing a clustering analysis on samples of reliably classified sources, as we are able to investigate the true clustering properties of these massive AGN type radio galaxies, which have been shown to be predominantly hosted by massive elliptical galaxies (McLure et al. 1999; Jarvis et al. 2001; Seymour et al. 2007). The clustering measurements of these samples conclusively show that AGN are much more biased than SFGs, confirming that AGN must reside in significantly more massive halos than SFGs. It is also noted that these samples of classified sources are expected to be more strongly clustered than that of unmatched samples, due to the fact that the sources must be a more massive subset in order for them

to have optical detections. The bias is seen to evolve with redshift for LERGs and HERGs, as well as for SFGs, with the evolution of the latter being less steep. The halo masses assumed in the S^3 simulations for FRI and FRII galaxies are likely incorrect, in that they seem to be underestimated for FRIs and overestimated for FRIIs. The high clustering values measured for the combined sample of LERGs and probable LERGs further support previous findings that there is a link between halo mass and the observed efficiency of AGN, i.e. the accretion mode. That is, LERGs are found to agree with the high clustering values previously measured, which imply that they exist in more massive dark matter halos and are strongly clustered. This is in keeping with the theory that inefficient accreters reside in denser environments, as in hot, dense environments, the material in the galaxy has more kinetic energy, making it harder to accrete onto the supermassive black hole.

By comparing our bias results with the bias models from S^3, we are able to estimate the typical dark matter halo mass of the different radio source types we've studied, based on the redshift independent halo masses adopted by Wilman et al. (2008). The bias values measured for our sample of SFGs suggest a halo mass of $M_{\rm halo}\sim 10^{11} - 3\times10^{12}h^{-1}\mathcal{M}_{\odot}$ for these sources. This estimate coincides with estimates from the literature that star formation peaks at halo masses $10^{11.8}-10^{12.4}\mathcal{M}_{\odot}$. This is estimated by calculating the ratio of stellar mass to halo mass ($M_*/M_{\rm halo}$), although McCracken et al. (2015) highlight the dependence of this estimate on redshift. Halo mass estimates are important for star formation and galaxy evolution studies, as the halo mass is thought to affect the removal of cold gas from galaxies (quenching). The halo mass for the AGN (HERGs + LERGs) sources is estimated at $M_{\rm halo}\sim 1-5\times10^{13}h^{-1}\mathcal{M}_{\odot}$ for redshifts $z<1.5$, in agreement with halo mass estimates of AGN and HERG sources ($M_{\rm halo}\sim 1-2\times10^{13}h^{-1}\mathcal{M}_{\odot}$), as well as LERG sources ($M_{\rm halo}\sim 3-4\times10^{13}h^{-1}\mathcal{M}_{\odot}$) from Hale et al. (2018). These halo mass estimates are larger than those of SFGs as expected, as AGN are found to be more massive, more biased tracers of the dark matter distribution.

Chapter 5

Conclusions

5.1 Summary

The topic of this Thesis is the use of radio sources to trace large-scale structure in the Universe, with the goal of determining the clustering properties of radio populations, how they relate to the dark matter distribution and the evolution of these properties with cosmic time. A further objective is to study the clustering of different radio source types, as well as their bias, and thereby gain information about the dark matter halo masses corresponding to different radio source types. This is achieved by performing a two-point angular correlation function analysis of various radio samples. The best-fit parameters of the measured ACFs, along with the redshift distribution of the sample, are used to determine the correlation length r_0 and bias $b(z)$, which is calculated at the median redshift of the respective sample. The clustering measurements in this Thesis fill the gap between the clustering analyses of wide, shallow ($S_{1.4\mathrm{GHz}} > 1$ mJy) surveys (Lindsay et al. 2014a; Lindsay et al. 2014b) and those of narrow, deeper ($S_{1.4\mathrm{GHz}} > 20$ μJy) surveys (Hale et al. 2018; Magliocchetti et al. 2017). Below I describe the analysis carried out in each chapter, followed by a summary of the main findings and conclusions.

In Chapter 2, the angular correlation function is measured for a sample of 11,431 radio sources from a $S_{1.4\mathrm{GHz}} > 440$ μJy JVLA Stripe 82 Snapshot Survey (Heywood et al. 2016) over ~ 100 deg^2 and having a resolution of 10×16 arcseconds. The survey

is conducted over an eastern and a western field, each having 5,490 and 5,941 sources, respectively. Data from the S^3 simulations (Wilman et al. 2008) are used to provide an estimate of the redshift distribution for samples which lack redshift information. I also make use of SKADS bias evolution models for comparison to our measured bias evolution trends. SKADS produces bias models of five different radio source types, i.e. normal star-forming galaxies, starbursts, radio-quiet AGN, FRI and FRII type AGN, where each of these source types are attributed a fixed halo mass, and a bias based on the halo masses and bias model from Mo & White (1996). I calculate the ACF for the full samples of radio sources, the results of which, even though hard to interpret, are given for completeness and to allow for comparison with clustering measurements in literature. A clustering analysis is also performed for the full samples after a collapsing technique is applied, in order to mitigate the effect of multiple-component sources on the clustering at small angular scales.

The JVLA sample is then cross-matched with deep optical coadded SDSS data (g$\sim$24.5), in order to obtain redshift information. A $z < 1$ cut is applied, as the redshifts below this limit are likely to be more reliable. This yields a sample of 4,511 sources, having 1420 spectroscopic matches and 3091 photometric matches. The JVLA $z < 1$ matched samples for the eastern and western fields are split into three bins in redshift and a clustering analysis is performed for the samples of each redshift slice. The ACF is also measured for the full $z < 1$ matched samples.

In order to gain further understanding of the clustering properties of different radio source types and how these properties evolve over cosmic time, the matched JVLA $z < 1$ sample used in Chapter 2 is also split by radio luminosity. Thus, in Chapter 3 sources are split into a low luminosity ($L_{1.4\mathrm{GHz}} < 10^{23}$ W Hz^{-1}) and mid-luminosity sample ($10^{23} < L_{1.4\mathrm{GHz}}$ W Hz$^{-1} < 10^{26}$), while the high luminosity sample ($L_{1.4\mathrm{GHz}}$ W Hz$^{-1} < 10^{26}$) comprises of all JVLA unmatched sources, as it is reasonable to assume that sources lacking an optical match would be high luminosity, high redshift sources. To allow for a direct comparison, the same luminosity cuts are applied to a SKADS $S_{1.4\mathrm{GHz}} > 440$ μJy sample and the ACF measured for each population.

In order to obtain clearer, more reliable clustering measurements for different types of radio sources, in Chapter 4 I then perform a clustering analysis of samples which have been cross-matched via visual inspection, to deep optical SDSS data, and thereafter classified (for those sources with spectroscopic information) into different radio sub-populations, i.e. SFGs, AGN, LERGs, probable LERGs and HERGs. Unfortunately, the number of classified sources is low relative to the sample sizes in previous chapters, resulting in measured clustering properties which are less constrained. We therefore make use of combined samples in order to extract information regarding the clustering properties. The ACF is calculated for the sample of SFGs, a combined sample of LERGs and probable LERGs and a combined sample of LERGs, probable LERGs and HERGs.

The main findings of this Thesis are summarised as follows:

• The clustering measurements of the full radio samples yield correlation length and bias values in the region of $r_0 \sim 13 - 16h^{-1}$Mpc and $b = 2.4 - 3.0$. These results show that radio sources trace a massive (and therefore biased) galaxy population. The clustering measurements fall between the clustering strength of ordinary local galaxies ($r_0 \sim 5h^{-1}$Mpc; Davis & Peebles 1983; Loveday et al. 1995) and Abell clusters ($r_0 \sim 25h^{-1}$Mpc; Bahcall & Soneira 1983; Postman et al. 1992; Gonzalez et al. 2002). Similar correlation length values have been found for elliptical galaxies, with r_0 values ranging from $r_0 \sim 7 - 12h^{-1}$Mpc (Willmer et al. 1998; Norberg et al. 2002), as well as for Extremely Red Objects (EROs) at $z \sim 1$ ($r_0 = 12 \pm 3.0h^{-1}$Mpc - Daddi et al. 2001; $r_0 = 10 - 13h^{-1}$Mpc - Roche et al. 2002), Distant Red Galaxies (DRGs) at $z \sim 1 - 2$ ($r_0 \sim 11h^{-1}$Mpc - Foucaud et al. 2007) and Luminous Red Galaxies (LRGs) at $z \sim 0.55$ ($r_0 = 12.3 \pm 1.2h^{-1}$Mpc - Wake et al. 2008). This supports findings which suggest that LRGs and radio galaxies, i.e. the population which dominates the clustering signal in a relatively deep radio survey, trace the same underlying dark matter halo distribution.

• The agreement between the clustering measurements of the original samples and

the corresponding collapsed samples shows that the use of the collapsing technique to mitigate the effect of source substructure on the clustering measured at small angular scales, does not alter the cosmological clustering of the original sample and is likely successful in ensuring that the clustering measured is from individual sources.

• Low clustering values ($r_0 \sim 3{-}6h^{-1}$Mpc and $b \sim 1$) are measured for the $z \leq 0.4$ redshift samples ($r_0 = 4.37^{+0.70}_{-0.75}h^{-1}$Mpc; $b = 1.01^{+0.20}_{-0.20}$ (east) and $r_0 = 6.34^{+1.06}_{-1.11}h^{-1}$Mpc; $b = 1.33^{+0.26}_{-0.23}$ (west)), as well as for the full sample of SFGs ($r_0 = 3.81^{+1.25}_{-2.10}h^{-1}$Mpc; $b = 0.91^{+0.70}_{-0.51}$). These values are all indicative of a less clustered sample, in which the sources are less biased traces of the underlying dark matter distribution and occupy low mass dark matter haloes. By comparing the bias results with the bias evolution models from SKADS, we estimate the halo mass of star-forming galaxies to be $M_{\rm halo} \sim 10^{11} - 3 \times 10^{12}h^{-1}\mathcal{M}_{\odot}$, based on the constant halo mass adopted by Wilman et al. (2008). This estimate corresponds with findings from literature that the measure of star formation, given by the ratio of stellar mass to halo mass ($M_*/M_{\rm halo}$), is shown to peak in halo masses of $10^{11.8} - 10^{12.4}\mathcal{M}_{\odot}$, however this is dependent on the redshift range probed (McCracken et al. 2015).

• Clustering results in line with those of radio AGN samples ($r_0 \sim 7 - 11h^{-1}$Mpc and $b \sim 1.5 - 2.3$) are found for the $0.4 < z \leq 0.7$ redshift samples ($7.77^{+0.79}_{-0.83}h^{-1}$Mpc; $b = 1.90^{+0.30}_{-0.28}$ and $r_0 = 7.96^{+1.53}_{-1.74}$; $b = 1.76^{+0.50}_{-0.44}$), the low luminosity samples ($r_0 = 8.02^{+1.73}_{-1.88}h^{-1}$Mpc; $b = 1.62^{+0.45}_{-0.38}$ and $r_0 = 7.98^{+1.40}_{-1.53}h^{-1}$Mpc; $b = 1.68^{+0.40}_{-0.35}$), mid-luminosity samples ($r_0 = 9.53^{+2.93}_{-2.52}h^{-1}$Mpc; $b = 2.01^{+0.85}_{-0.59}$ and $r_0 = 9.98^{+3.10}_{-2.98}h^{-1}$Mpc; $b = 1.80^{+0.60}_{-0.51}$), as well as the LERG + probable LERG sample ($r_0 = 9.37^{+4.96}_{-5.10}h^{-1}$Mpc; $b = 2.07^{+2.13}_{-1.17}$) and the LERG + HERG sample ($r_0 = 10.88^{+6.58}_{-7.44}h^{-1}$Mpc; $b = 2.30^{+3.00}_{-1.54}$). These values are in line with previous clustering measurements of AGN (Magliocchetti et al. 2004; Brand et al. 2005; Magliocchetti et al. 2017; Hale et al. 2018). Comparing these bias results with the bias models from SKADS, we estimate the halo mass of AGN (HERGs + LERGs) to be in the region of $M_{\rm halo} \sim 1 - 5 \times 10^{13}h^{-1}\mathcal{M}_{\odot}$ for redshifts $z < 1.5$, in line with halo mass estimates of AGN, HERGs and LERGs by Hale et al. (2018). These halo mass estimates are larger than those of SFGs as expected, as

AGN are found to be more massive, occupy more massive dark matter haloes and as such, are more biased tracers of the dark matter distribution.

• High clustering values ($r_0 > 11h^{-1}$Mpc and $b \sim 2.5-4$) are found for a high redshift $0.7 < z \leq 1.0$ sample ($r_0 = 12.37^{+3.76}_{-4.81}h^{-1}$Mpc; $b = 2.53^{+1.44}_{-1.08}$), as well as the high luminosity samples ($r_0 = 17.84^{+7.13}_{-5.86}h^{-1}$Mpc; $b = 3.21^{+1.49}_{-1.01}$ and $r_0 = 19.30^{+3.88}_{-4.48}h^{-1}$Mpc; $b = 3.72^{+1.45}_{-0.38}$), which comprise of all JVLA sources lacking an optical match. This reflects that higher luminosity sources occupy higher redshifts. Similarly high clustering values are also measured for the high redshift ($0.45 \leq$ z < 0.7) samples of LERGs + probable LERGs ($r_0 = 11.93^{+1.04}_{-2.96}$; $b = 2.50^{+0.83}_{-0.81}$) and LERGs + HERGs ($r_0 = 12.09^{+0.22}_{-8.36}$; $b = 2.52^{+2.69}_{-1.72}$). These clustering values are higher than those measured in previous clustering analyses of radio galaxies, which suggests that we were able to probe a different radio source population than those previously studied. This is due to the $S_{1.4\text{GHz}}$= 440μJy flux density limit, which yields a radio source population that is still AGN dominated (unlike deeper surveys), yet sensitive enough to detect a significant number of high-redshift, powerful AGN.

• Although the low luminosity samples ($L_{1.4\text{GHz}} < 10^{23}$ W Hz^{-1}) are expected to be dominated by SFGs, the redshift distribution, as well as the higher correlation length and bias results measured for these samples (eastern and western fields), suggest that they contains a fair amount of low luminosity AGN. In comparison, the S^3 simulation underestimates the presence of low luminosity AGN in their sample.

• Correlation length and bias values are found to increase with redshift. These clustering properties also increase with an increase in radio luminosity, due to the dominance of different radio source populations at different luminosity cuts. In both Chapters 2 and 3 the bias evolution measured is found to follow the bias model for FRI sources from the SKADS simulation, which suggest that this $S_{1.4\text{GHz}}$ >440μJy JVLA sample is AGN dominated. This also supports the SKADS source composition estimate that, for the given flux limit, the JVLA sample is comprised 85% of AGN and 15% of SFGs.

• The high clustering values measured for the LERG + probable LERG sample supports the notion that LERGs are hosted by massive, late type galaxies and also suggests a link between accretion mode and halo mass, i.e. LERGs are found in highly clustered, dense environments, where the surplus of hot gas is harder to accrete onto the supermassive black hole.

The large-scale clustering studies in this Thesis have filled the gap between the clustering analyses of wide, shallow ($S_{1.4\mathrm{GHz}} > 1$ mJy) surveys (Lindsay et al. 2014a; Lindsay et al. 2014b) and those of narrow, deeper ($S_{1.4\mathrm{GHz}} > 20$ μJy) surveys (Hale et al. 2018; Magliocchetti et al. 2017), allowing for a more complete picture of the clustering and bias of the radio source population. They have also provided insight into the clustering properties of different radio source types, which thus far has not been well studied due to the unavailability of deep data at other wavelengths, which are necessary for source classification. The high-redshift and high luminosity populations have been shown to be highly biased and could therefore be valuable in future studies of the nature of dark matter and the relationship between baryons and dark matter. Measurements of the bias in this Thesis suggest that the S^3 halo masses assumed by Wilman et al. (2008) may need to be reviewed, as they are underestimated for SFGs and overestimated for FRII type AGN. In order to achieve more accurate constraints on the clustering characteristics, biases and halo masses of radio populations, we require surveys which probe larger volumes (i.e. do not trade areal coverage for depth) over parts of the sky where the deep photometry needed for cosmological clustering studies, or the deep spectroscopy needed for source classification, are available.

5.2 Future Work

5.2.1 A Clustering Analysis of JVLA S82 Data Using HOD Modelling

The work of this Thesis could be improved by the use of a more sophisticated treatment of the measured $\omega(\theta)$ and the way in which the bias in inferred. Halo occupation distribution (HOD) models (see Cooray & Sheth (2002) for a full description) describe how galaxies populate the clumps of dark matter in the Universe, i.e. the probability

distribution for the number of galaxies N hosted by a dark matter halo as a function of its mass M ($P(N|M)$). 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 (Blake et al. 2008; Cooray et al. 2010; Geach et al. 2012). The proposed model depends on both the halo occupation number, i.e. the number of galaxies in a halo of a given mass, for a galaxy sample brighter than some limit, and the location of the galaxies within these halos.

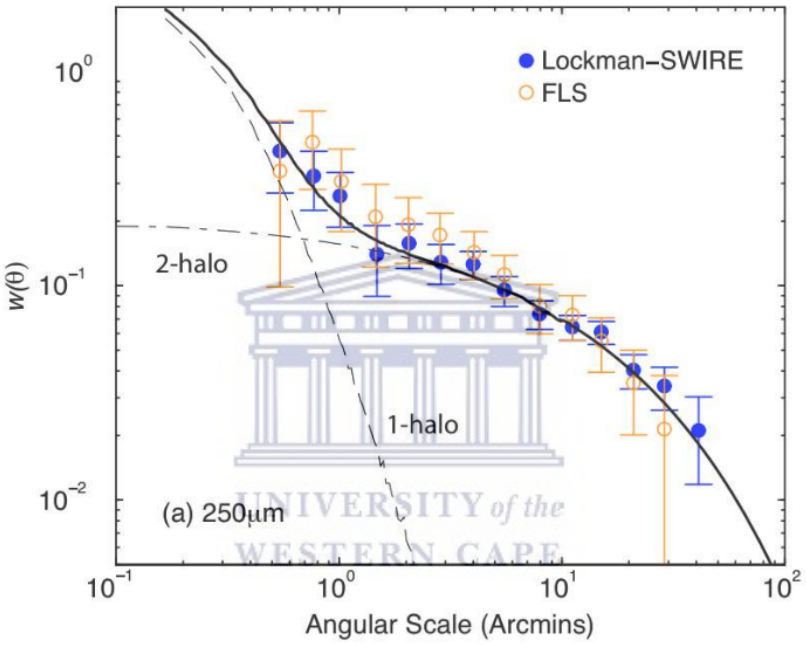


Figure 5.1: The angular correlation function of 250μm selected sources in two wide fields with HerMES (Cooray et al. 2010) overplotted with HOD model fitting. The lines illustrate the contribution from the 1-halo term (multiple sources in the same halo) and the 2-halo term (linear clustering).

In the HOD models, the clustering of galaxies on scales larger than a typical halo ($\sim 1 - 2h^{-1}$ Mpc) results from pairs of galaxies in separate halos, called the 'two-halo term', while the clustering on smaller scales is due to pairs of galaxies within the same parent halo, i.e. the 'one-halo term'. An example of the contributions of each term to the measured ACF are shown in Figure 5.1, in which an ACF of sub-mm sources from the Herschel Multi-tiered Extragalactic Survey (HERMES; Oliver et al. 2010) is fitted by an HOD model. The HOD model fitted is usually based on three parameters, i.e. (i)

the minimum mass required to host a single central galaxy, the minimum mass required to host a single satellite galaxy, and the slope (α) of $P(N|M)$, which determines the average number of satellite galaxies as a function of the host halo mass (Zehavi et al. 2004). The bias can then be derived from the HOD model fit to the ACF, given a model for the bias as a function of halo mass (Sheth et al. 2001). Performing such an analysis could allow for better accuracy when calculating the bias and eliminate the dependence of the bias on Limber Inversion. Successful HOD modelling using radio surveys would require some specific treatment of sources in order to mitigate the contribution of radio galaxy substructure (cores, lobes & hotspots) to the 1-halo term. We also require sufficient survey depth to allow for the detection of the less powerful, lower-mass satellite galaxies, as well as wide surveys to allow for accurate measurement of the 2-halo term.

5.2.2 An ACF Analysis of MeerKAT Sources

The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE - Jarvis et al. 2018) survey will use the 64 antenna MeerKAT array to carry out a wide, deep continuum observation down to a $\sim 1\mu$Jy sensitivity. The survey will be carried out over a 20 deg^2 region of the Southern sky which contains a wealth of extra-galactic information in fields such as COSMOS, ECDFS, XMM-LSS and ELAIS-S1. This would allow for a large-scale clustering analysis of a volume of radio galaxies which is yet to be probed. Almost all MIGHTEE sources will have photometric redshifts available from the VISTA Deep Extragalactic Observations survey (VIDEO - Jarvis et al. 2013) and some will have spectral classifications from DEVILS (Davies et al. 2018). Such a study of large-scale structure would thus allow for well constrained measurements of the ACF over various ranges in redshift, as well as for large samples of classified radio sub-populations. This would yield more accurate measurements of the galaxy bias and the bias evolution with redshift. In turn, significant progress could be made in determining the halo masses required to host various radio sources. Figure 5.2, taken from Jarvis et al. (2018), shows the expected improvement in constraining the ACF allowed by the analysis of $S_{1.4\mathrm{GHz}}= 5\mu$Jy MIGHTEE data over 7.5 deg^2, compared to the ACF measured by Hale et al. (2018) for a $S_{1.4\mathrm{GHz}} \sim 20\mu$Jy JVLA-

COSMOS sample over $\sim 1.5\text{deg}^2$. These survey characteristics of MIGHTEE would thus also allow for successful HOD modelling of the clustering of radio sources.

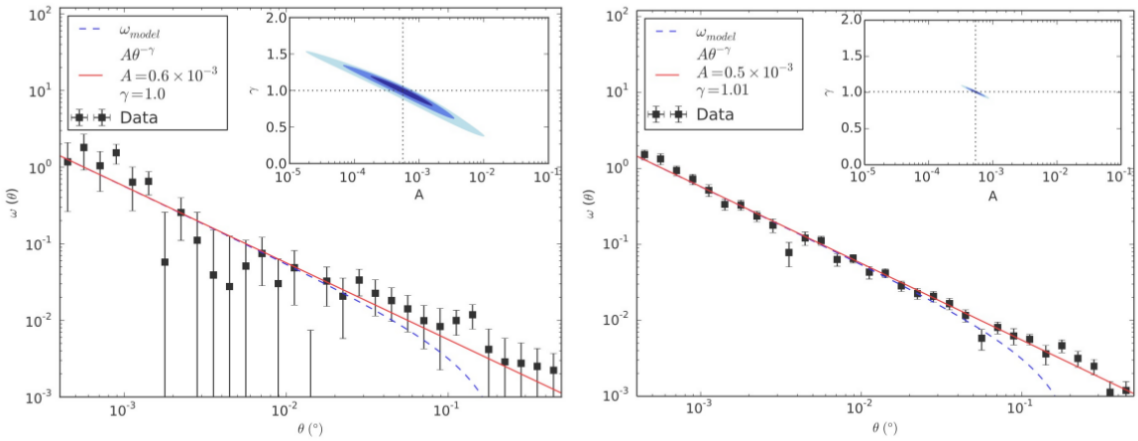


Figure 5.2: Left: The ACF measured by Hale et al. (2018) for a $S_{1.4\text{GHz}} \sim 20\mu$Jy JVLA-COSMOS sample over $\sim 1.5\text{deg}^2$. Right: The expected improvement in constraining the ACF allowed by 7.5 deg^2 of MeerKAT MIGHTEE data at $S_{1.4\text{GHz}}= 5\mu$Jy (Jarvis et al. 2018).

5.2.3 Clustering analyses with the SKA

The MeerKAT array will be incorporated into the Square Kilometer Array (SKA) telescope. The SKA phase 1 will consist of ∼200 dishes allowing for radio surveys of unprecedented volume, as shown in Figure 5.3 from Maartens et al. (2015). The radio continuum surveys from SKA1-MID will cover $\sim 30,000\text{deg}^2$ out to $z \sim 6$, detecting $\sim 10^8$ galaxies. Together with photometry from the Large Synoptic Survey Telescope (LSST; Ivezić et al. 2019), this would allow for studies of large-scale structure to CMB level precision. The sample resulting from a survey with a depth similar to MIGHTEE, but over the entire southern hemisphere, would allow for clustering studies which are able to accurately constrain the bias and its evolution over cosmic time, probing the largest scales in the Universe. It will also allow for more advanced studies of the effects of galaxy evolution processes on dark matter haloes and the resulting halo density profiles, providing an understanding of the effect of baryons on the underlying dark matter and how this affects the clustering signal. Different radio sources trace the underlying dark matter very differently and SKA will allow for the detection of star-forming galaxies, as well as the AGN to high redshifts. The majority of radio

sources detected with the sensitivity of SKA1 will be too faint at optical wavelengths in order for the availability of spectroscopic redshifts to be possible. However, work by Makhathini et al. (2015) show that morphological classifications of radio sources to high redshift is possible with SKA1-MID. Such classifications would then allow for the use of the multi-tracer technique (Seljak 2009). This technique makes use of the fact that some source populations are highly biased tracers of the dark matter, whilst others are less biased/unbiased tracers. Correlating the one against the other would allow us to overcome the effects of cosmic variance on large-scale clustering measurements.

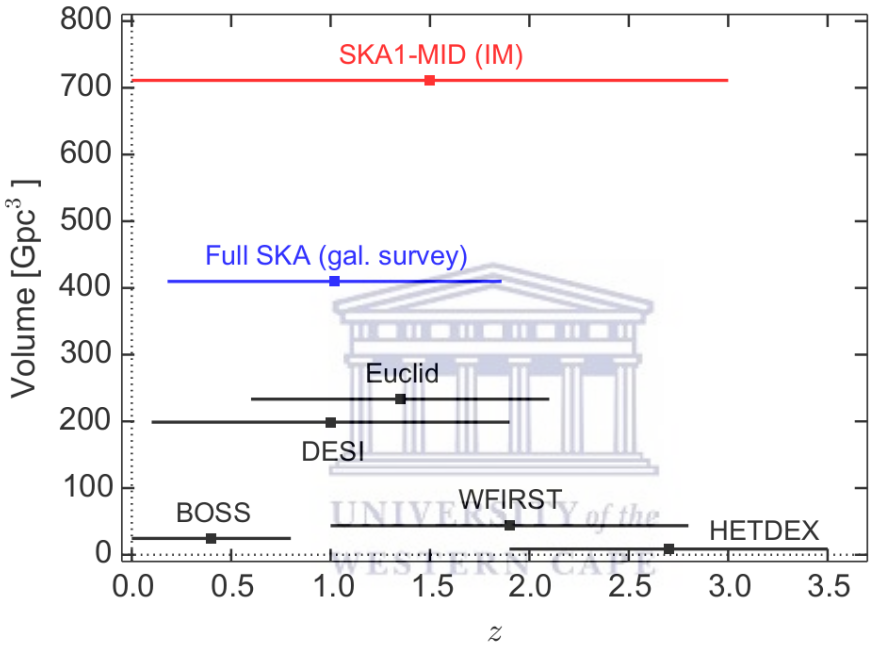


Figure 5.3: The survey volume of SKA1-MID compared to various current and future surveys, at the midpoint of the range in redshift of the survey. Figure taken from Maartens et al. (2015).

Bibliography

Abbott, T. M. C. et al. 2019, MNRAS, 483, 4866

Abbott, T. M. C. et al. 2019, Phys. Rev. Lett., 122, 171301

Abbott, T. M. C. et al. 2019, ApJ, 872, L30

Abolfathi, B. et al. 2018, ApJS, 235, 42

Alam, S. et al. 2015, ApJS, 219, 12

Alam, S. et al. 2017, MNRAS, 470, 2617

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

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

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

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

Barr, J. M., Bremer, M. N., Baker, J. C., & Lehnert, M. D. 2003, MNRAS, 346, 229

Bartelmann, M., & Schneider, P. 1999, Weak Gravitational Lensing (Elsevier Preprint)

Barthel, P. 1989, Scientific American, 260, 20

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

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

Bell Burnell, S. J. 1977, Annals of the New York Academy of Sciences, 302, 685

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

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. 2005a, MNRAS, 362, 25

Best, P. N., Kauffmann, G., Heckman, T. M., & Ivezić, Ž. 2005b, MNRAS, 362, 9

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

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

Bonzini, M. et al. 2012, ApJS, 203, 15

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

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

Buttiglione, S., Capetti, A., Celotti, A., Axon, D. J., Chiaberge, M., Macchetto, F. D., & Sparks, W. B. 2010, A&A, 509, A6

Caldwell, R. R., Dave, R., & Steinhardt, P. J. 1998, Ap&SS, 261, 303

Camera, S., Santos, M. G., Bacon, D. J., Jarvis, M. J., McAlpine, K., Norris, R. P., Raccanelli, A., & Röttgering, H. 2012, ArXiv e-prints

Cid Fernandes, R., Stasińska, G., Schlickmann, M. S., Mateus, A., Vale Asari, N., Schoenell, W., & Sodré, L. 2010, MNRAS, 403, 1036

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C., & Zaritsky, D. 2006, ApJ, 648, L109

Clowe, D., Gonzalez, A., & Markevitch, M. 2004, ApJ, 604, 596

Cochrane, R. K., Best, P. N., Sobral, D., Smail, I., Geach, J. E., Stott, J. P., & Wake, D. A. 2018, MNRAS, 475, 3730

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

Condon, J. J. 1984, ApJ, 287, 461

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

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

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., Kellermann, K. I., Kimball, A. E., Ivezić, Ž., & Perley, R. A. 2013, ApJ, 768, 37

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

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

Crill, B. P. et al. 2003, ApJS, 148, 527

Crocce, M. et al. 2019, MNRAS, 482, 2807

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

Davies, L. J. M. et al. 2018, MNRAS, 480, 768

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

de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, Experimental Astronomy, 35, 25

Dolley, T. et al. 2014, The Astrophysical Journal, 797, 125

Donoso, E., Yan, L., Stern, D., & Assef, R. J. 2014, The Astrophysical Journal, 789, 44

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

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

Dyson, F. W., Eddington, A. S., & Davidson, C. 1920, Royal Society of London Philosophical Transactions Series A, 220, 291

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

Einstein, A. 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, 142

Eisenstein, D. J. et al. 2011, AJ, 142, 72

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

Elvis, M. et al. 1994, ApJS, 95, 1

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

Evans, D. A., Worrall, D. M., Hardcastle, M. J., Kraft, R. P., & Birkinshaw, M. 2006, ApJ, 642, 96

Ewen, H. I., & Purcell, E. M. 1951, Nature, 168, 356

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

Fernandes, C. A. C. et al. 2015, MNRAS, 447, 1184

Ford, J. et al. 2012, ApJ, 754, 143

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

Freedman, W. L. et al. 2001, ApJ, 553, 47

Freedman, W. L. et al. 2019, ApJ, 882, 34

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

Geller, M. J., & Huchra, J. P. 1989, Science, 246, 897

Gendre, M. A., Best, P. N., Wall, J. V., & Ker, L. M. 2013, MNRAS, 430, 3086

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, A. H., Zaritsky, D., & Wechsler, R. H. 2002, ApJ, 571, 129

Gopal-Krishna, & Wiita, P. J. 2000, A&A, 363, 507

Gürkan, G. et al. 2015, MNRAS, 452, 3776

Hale, C. L., Jarvis, M. J., Delvecchio, I., Hatfield, P. W., Novak, M., Smolčić, V., & Zamorani, G. 2018, MNRAS, 474, 4133

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

Hamilton, A. J. S. 2001, MNRAS, 322, 419

Hamuy, M., Phillips, M. M., Suntzeff, N. B., Schommer, R. A., Maza, J., & Aviles, R. 1996, AJ, 112, 2398

Hardcastle, M. J. 2004, A&A, 414, 927

Hardcastle, M. J. et al. 2013, MNRAS, 429, 2407

Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2006, MNRAS, 370, 1893

Hardcastle, M. J. et al. 2016, MNRAS, 462, 1910

Hatfield, P. W., Lindsay, S. N., Jarvis, M. J., Häußler, B., Vaccari, M., & Verma, A. 2016, MNRAS, 459, 2618

Heckman, T. M. 1980, A&A, 500, 187

Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589

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

Heywood, I. et al. 2016, ArXiv e-prints

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

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

Hinshaw, G. et al. 2013, ApJS, 208, 19

Ho, L. C. 2008, ARA&A, 46, 475

Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1993, ApJ, 417, 63

Ho, S., Hirata, C., Padmanabhan, N., Seljak, U., & Bahcall, N. 2008, Phys. Rev. D, 78, 043519

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

Hoekstra, H., & Jain, B. 2008, Annual Review of Nuclear and Particle Science, 58, 99

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

Ineson, J., Croston, J. H., Hardcastle, M. J., Kraft, R. P., Evans, D. A., & Jarvis, M. 2015, MNRAS, 453, 2682

Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., & Lal, D. V. 2011, A&A, 535, A38

Ivezić, Ž. et al. 2019, ApJ, 873, 111

Jackson, N., & Rawlings, S. 1997, MNRAS, 286, 241

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

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

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

Janssen, R. M. J., Röttgering, H. J. A., Best, P. N., & Brinchmann, J. 2012, A&A, 541, A62

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

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

Jarvis, M. J., & Rawlings, S. 2004, New Astronomy Reviews, 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. 2018, PoS, MeerKAT2016, 006

Jiang, L. et al. 2014, ApJS, 213, 12

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

Jungman, G., Kamionkowski, M., & Griest, K. 1996, Phys. Rep., 267, 195

Kaiser, N. 1984, ApJ, 284, L9

Kant Goswami, G., Pradhan, A., & Beesham, A. 2019, arXiv e-prints, arXiv:1905.10801

Kauffmann, G., Heckman, T. M., & Best, P. N. 2008, MNRAS, 384, 953

Kauffmann, G. et al. 2003, MNRAS, 346, 1055

Kellermann, K. I., Condon, J. J., Kimball, A. E., Perley, R. A., & Ivezić, Ž. 2016, ApJ, 831, 168

Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, AJ, 98, 1195

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

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961

Kewley, L. J., Heisler, C. A., Dopita, M. A., & Lumsden, S. 2001, ApJS, 132, 37

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

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

Kimball, A. E., Kellermann, K. I., Condon, J. J., Ivezić, Ž., & Perley, R. A. 2011, ApJ, 739, L29

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, The Astrophysical Journal, 471, 643

Köhlinger, F. et al. 2017, MNRAS, 471, 4412

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

Kukula, M. J., Dunlop, J. S., Hughes, D. H., & Rawlings, S. 1998, MNRAS, 297, 366

Laing, R. A., Jenkins, C. R., Wall, J. V., & Unger, S. W. 1994, in Astronomical Society of the Pacific Conference Series, Vol. 54, The Physics of Active Galaxies, ed. G. V. Bicknell, M. A. Dopita, & P. J. Quinn, 201

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

Lara, L., Cotton, W. D., Feretti, L., Giovannini, G., Marcaide, J. M., Márquez, I., & Venturi, T. 2001, A&A, 370, 409

Ledlow, M. J., & Owen, F. N. 1996, AJ, 112, 9

Leitch, E. M. et al. 2002, Nature, 420, 763

Liddle, A. R. 1999, in High Energy Physics and Cosmology, 1998 Summer School, ed. A. Masiero, G. Senjanovic, & A. Smirnov, 260

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

Lin, Y.-T., Shen, Y., Strauss, M. A., Richards, G. T., & Lunnan, R. 2010, ApJ, 723, 1119

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

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

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

Loveday, J., Maddox, S. J., Efstathiou, G., & Peterson, B. A. 1995, ApJ, 442, 457

Lupton, R. 1993, Statistics in theory and practice

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

Maartens, R., Abdalla, F. B., Jarvis, M., & Santos, M. G. 2015, arXiv e-prints, arXiv:1501.04076

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

Magliocchetti, M., Maddox, S. J., Lahav, O., & Wall, J. V. 1998, MNRAS, 300, 257

Magliocchetti, M., Popesso, P., Brusa, M., Salvato, M., Laigle, C., McCracken, H. J., & Ilbert, O. 2017, MNRAS, 464, 3271

Magliocchetti, M. et al. 2013, Monthly Notices of the Royal Astronomical Society, 433, 127

Makhathini, S., Jarvis, M., Smirnov, O., & Heywood, I. 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14), 81

Mandelbaum, R., Li, C., Kauffmann, G., & White, S. D. M. 2009, Monthly Notices of the Royal Astronomical Society, 393, 377

Markevitch, M., Gonzalez, A. H., Clowe, D., Vikhlinin, A., Forman, W., Jones, C., Murray, S., & Tucker, W. 2004, ApJ, 606, 819

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

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

McCracken, H. J. et al. 2015, MNRAS, 449, 901

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

McLure, R. J., Kukula, M. J., Dunlop, J. S., Baum, S. A., O'Dea, C. P., & Hughes, D. H. 1999, Monthly Notices of the Royal Astronomical Society, 308, 377

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

Mellier, Y. 1999, in NATO ASIC Proc. 541: Theoretical and Observational Cosmology, ed. M. Lachièze-Rey, 211

Ménard, B., Scranton, R., Fukugita, M., & Richards, G. 2010, MNRAS, 405, 1025

Merloni, A., & Heinz, S. 2007, MNRAS, 381, 589

Milgrom, M. 1983, ApJ, 270, 365

Mingo, B. et al. 2019, MNRAS, 488, 2701

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

Mohan, N., & Rafferty, D. 2015, PyBDSM: Python Blob Detection and Source Measurement, Astrophysics Source Code Library

Munshi, D., Valageas, P., van Waerbeke, L., & Heavens, A. 2008, Phys. Rep., 462, 67

Nadathur, S., & Crittenden, R. 2016, ApJ, 830, L19

Narayan, R. 2005, Ap&SS, 300, 177

Norberg, P., Baugh, C. M., Gaztañaga, E., & Croton, D. J. 2009, MNRAS, 396, 19

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

Norris, R. P. 2009, in Panoramic Radio Astronomy: Wide-field 1-2 GHz Research on Galaxy Evolution, 33

Norris, R. P. 2012, in IAU Symposium, Vol. 284, IAU Symposium, 489–493

Ogle, P., Whysong, D., & Antonucci, R. 2006, ApJ, 647, 161

Oliver, S. J. et al. 2010, A&A, 518, L21

Omori, Y. et al. 2019, Phys. Rev. D, 100, 043517

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

Padovani, P., Mainieri, V., Tozzi, P., Kellermann, K. I., Fomalont, E. B., Miller, N., Rosati, P., & Shaver, P. 2009, ApJ, 694, 235

Padovani, P., Miller, N., Kellermann, K. I., Mainieri, V., Rosati, P., & Tozzi, P. 2011, ApJ, 740, 20

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

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. 2010, MNRAS, 401, 2148

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

Planck Collaboration et al. 2018, arXiv e-prints, arXiv:1807.06209

Postman, M., Huchra, J. P., & Geller, M. J. 1992, ApJ, 384, 404

Pracy, M. B., Ching, J., Croom, S., & Sadler, E. M. 2013, in IAU Symposium, Vol. 295, The Intriguing Life of Massive Galaxies, ed. D. Thomas, A. Pasquali, & I. Ferreras, 117–120

Prescott, M. et al. 2018, MNRAS, 480, 707

Rabii, B. et al. 2006, Review of Scientific Instruments, 77, 071101

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

Rengelink, R., & et al. 1998, in Astrophysics and Space Science Library, Vol. 226, Observational Cosmology with the New Radio Surveys, ed. M. N. Bremer, N. Jackson, & I. Perez-Fournon, 143

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., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ, 876, 85

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

Riess, A. G. et al. 2004, ApJ, 607, 665

Rocca-Volmerange, B., Le Borgne, D., De Breuck, C., Fioc, M., & Moy, E. 2004, A&A, 415, 931

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

Röttgering, H. 2003, New Astron. Rev., 47, 405

Röttgering, H. et al. 2011a, Journal of Astrophysics and Astronomy, 32, 557

Röttgering, H. et al. 2011b, Journal of Astrophysics and Astronomy, 32, 557

Rubin, V. C., Ford, W. K. J., & . Thonnard, N. 1980, ApJ, 238, 471

Sachs, R. K., & Wolfe, A. M. 1967, ApJ, 147, 73

Salpeter, E. E. 1964, ApJ, 140, 796

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

Schmidt, M. 1963, Nature, 197, 1040

Scranton, R. et al. 2003, ArXiv Astrophysics e-prints

Scranton, R. et al. 2005, ApJ, 633, 589

Seljak, U. 2009, Phys. Rev. Lett., 102, 021302

Seyfert, C. K. 1943, ApJ, 97, 28

Seymour, N., Dwelly, T., Moss, D., Mhardy, I., Zoghbi, A., Rieke, G., Page, M., & Hopkins, A. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 399, Panoramic Views of Galaxy Formation and Evolution, ed. T. Kodama, T. Yamada, & K. Aoki, 197

Seymour, N. et al. 2007, ApJS, 171, 353

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1

Shimwell, T. W. et al. 2017, A&A, 598, A104

Simpson, C. et al. 2006, MNRAS, 373, L21

Simpson, C. et al. 2012, MNRAS, 421, 3060

Smolcic, V. et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 69

Smolčić, V. et al. 2009a, ApJ, 690, 610

Smolčić, V. et al. 2009b, ApJ, 696, 24

Smolčić, V. et al. 2017, A&A, 602, A1

Smoot, G. F. et al. 1992, ApJ, 396, L1

Sramek, R. A., & Weedman, D. W. 1980, ApJ, 238, 435

Tasse, C., Best, P. N., Röttgering, H., & Le Borgne, D. 2008, A&A, 490, 893

The Dark Energy Survey Collaboration. 2005, arXiv Astrophysics e-prints

Tonry, J. L. et al. 2003, ApJ, 594, 1

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803

Urry, M. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 290, Active Galactic Nuclei: From Central Engine to Host Galaxy, ed. S. Collin, F. Combes, & I. Shlosman, 3

van Haarlem, M. P. et al. 2013, A&A, 556, A2

Verschuur, G. L., & Kellermann, K. I. 1988, Science, 242, 1065

Wake, D. A., Croom, S. M., Sadler, E. M., & Johnston, H. M. 2008, MNRAS, 391, 1674

Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381

Wang, J. M., Ho, L. C., & Staubert, R. 2003a, A&A, 409, 887

Wang, L. et al. 2011, MNRAS, 414, 596

Wang, L., Goldhaber, G., Aldering, G., & Perlmutter, S. 2003b, ApJ, 590, 944

White, S. V., Jarvis, M. J., Kalfountzou, E., Hardcastle, M. J., Verma, A., Cao Orjales, J. M., & Stevens, J. 2017, MNRAS, 468, 217

Whittam, I. H., Prescott, M., McAlpine, K., Jarvis, M. J., & Heywood, I. 2018, MNRAS, 480, 358

Willmer, C. N. A., da Costa, L. N., & Pellegrini, P. S. 1998, AJ, 115, 869

Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M., & Eales, S. A. 2001, MNRAS, 322, 536

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

Windhorst, R. A., Miley, G. K., Owen, F. N., Kron, R. G., & Koo, D. C. 1985, ApJ, 289, 494

York, D. G. et al. 2000, AJ, 120, 1579

Zehavi, I. et al. 2005, ApJ, 621, 22

Zehavi, I. et al. 2004, ApJ, 608, 16

Zehavi, I. et al. 2011, ApJ, 736, 59

Zel'dovich, Y. B., & Novikov, I. D. 1965, Soviet Physics Doklady, 9, 834

Zhang, Y., Ma, H., Peng, N., Zhao, Y., & Wu, X.-b. 2013, AJ, 146, 22

Zheng, Z. et al. 2005, ApJ, 633, 791

Zirbel, E. L. 1997, ApJ, 476, 489

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

Zwicky, F. 1937, Phys. Rev., 51, 679