UNIVERSITY OF THE WESTERN CAPE

MASTERS THESIS

Testing the statistical isotropy of the Universe using radio survey data

---

Author:
Albert BALOYI

Supervisor:
Prof. Roy MAARTENS
Dr. Carlos BENGALY

UNIVERSITY *of the*
WESTERN CAPE

A thesis submitted in fulfillment of the requirements

for the degree of Master of Science

in the

UWC Astrophysics Group

Department of Physics and Astronomy

February 9, 2019

ii

Declaration of Authorship

I, Albert BALOYI, declare that this thesis titled, "Testing the statistical isotropy of the Universe using radio survey data" and the work presented in it are my own. I confirm 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.

Signed:

Date: 09 February 2019

"Once you eliminate the impossible, whatever is left, no matter how improbable, it must be the truth. "

Sir Arthur Conan Doyle, Sherlock Holmes

UNIVERSITY *of the*
WESTERN CAPE

Abstract

The Cosmological Principle forms part of one of the most fundamental hypotheses of modern Cosmology. So it is very important to assess whether it holds true using observational data, or whether it consists of a mathematical simplification. We probe the statistical isotropy of the Universe using the existing radio continuum data, by means of a local variance estimator. In order to investigate this, we analyse the number count variance of the radio catalog by looking at patches of approximately $10°$, $15°$, $20°$ & $25°$ in radii, and thus comparing it to mock catalogs which reproduce the matter density power spectrum, as well as the same sky coverage of the real data. We establish criteria for accepting patches that have more than 90%, 70% & 50% of their pixels not masked. We make use of the NRAO VLA Sky Survey (NVSS), whose operational frequency is 1.4 GHz. We perform statistical tests for detecting possible departures from statistical isotropy using galaxy number counts with flux limits of $20 < S_{NVSS} < 1000$ mJy. We also compare the real data to the mock catalogs of the radio data in order to assess the statistical significance of our results. We use the local variance estimator for testing the statistical isotropy of our data sample. We find that the statistical properties of our sample are in reasonable agreement with the standard cosmological model. The mean of the distribution for the data falls well within the 95% confidence interval of the average of the simulated mocks. For all the radii and acceptance criteria for the patches, we found no significant deviations beyond those allowed by the standard model. As expected there were no large discrepancies between our mocks and the data. The results are consistent with statistical isotropy.

Acknowledgements

I would first like to thank my thesis supervisor Professor Roy Maartens of the Department of Physics & Astronomy at the University of the Western Cape. The door to Prof Maartens's office was always open whenever I ran into trouble or had a question about my research or writing. He consistently allowed this thesis to be my own work but steered me in the right direction whenever he thought I needed it.

I would also like to acknowledge Dr Carlos Bengaly, a postdoctoral researcher at the Department of Physics & Astronomy at the University of the Western Cape as the second reader of this thesis, and I am gratefully indebted to his very valuable comments, assistance with understanding abstract topics and guidance.

Finally, I must express my very profound gratitude to my friends. Special thanks to Reikantse Diretse for allowing me to use his computer in the NASSP lab. To my family for providing me with support and continuous encouragement through the process of researching and writing this thesis. This accomplishment would not have been possible without them.
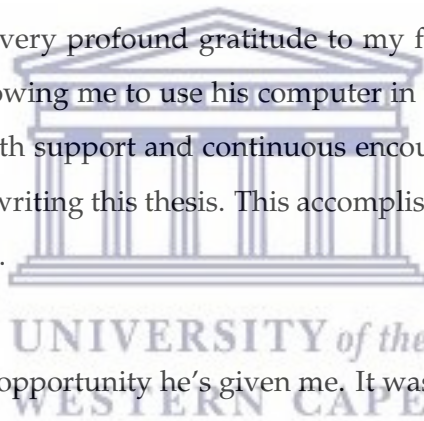
Lastly, I thank God, for the opportunity he's given me. It was through his guidance and strength that I was able to finish this work.

Contents

List of Figures

List of Tables

List of Abbreviations

NVSS	**N**RAO **V**LA **S**ky **S**urvey
CP	**C**osmological **P**rinciple
CMB	**C**osmic **M**icrowave **B**ackground
COBE	**CO**smic **B**ackground Explorer
WMAP	(the) **W**ilkinson **M**icrowave **A**nisotropy **P**robe
LSS	**L**arge **S**cale **S**tructure
VLA	(the) **V**ery **L**arge **A**rray
HEALPix	(the) **H**ierarchical **E**qual **A**rea (iso)-**L**atitude **Pix**elization
ΛCDM	**L**ambda **C**old **D**ark **M**atter
EMU	**E**volutionary **M**ap of the Universe
SKA	**S**quare **K**ilometer **A**rray
ASKAP	the **A**ustralian **S**quare **K**ilometre **A**rray **P**athfinder

For my parents, with love…

Chapter 1

Introduction

Cosmology deals with the study of the physical Universe on very large scales. As a science, it explores theoretical and observational properties of the Universe at very large scales, way beyond the scale of the Milky Way and the local group. Discoveries of galaxies beyond the Milky Way were only made during the late 1920s. Observations of galaxies, large-scale structures and gravitational lensing of the observable Universe all suggest that most of the observable Universe consists of dark matter with the majority of the Universe's energy in the form of dark energy.

The assumption that the Universe is nearly the same everywhere in all directions, at sufficiently large scales, is one that is widely accepted in modern Cosmology (Clarkson, 2012). Due to the advancement of technology and quality of astronomical observations, we are able to have a clearer picture of the Universe on cosmological scales. The Cosmological Principle (CP) states that on large-scales, the distribution of matter in the Universe is homogeneous and isotropic (Dodelson, 2003). This implies that the statistical properties in the Universe must be the same in all directions on the sky (isotropy) and from all positions in the Universe (homogeneity).

For a long time, this has been a necessity for simplification of otherwise complex equations. It is unfortunately difficult to verify the validity of the Cosmological Principle. The high degree of isotropy of the Cosmic Microwave Background (CMB) clearly supports this assumption during the early Universe and for most of the Universe's history (Akrami et al., 2018). Like the thermal radiation of the Cosmic Microwave Background at a temperature of ~ 3 K. Some deviations from this isotropy

in the form of a temperature dipole were depicted at ~ 3 mK. The dipole is interpreted with respect to the Cosmic Microwave Background (CMB) rest frame. This includes the motion of our solar system and even the motion due to our galaxy.

Some unexpected features of the Cosmic Microwave Background temperature fluctuations have been observed by several space missions like the Wilkinson Microwave Anisotropy Probe (WMAP), Planck and the Cosmic Background Explorer (COBE). These anomalies include the shortage of large angle correlation, some hemispherical symmetry and alignment of low multipole moments. These signals require investigation when better data from galaxy surveys become available. The upcoming Square Kilometer Array (SKA) will be able to map neutral hydrogen out to a few redshifts. This will allow us to investigate the cosmological dipole and also be able to test the proper motion hypothesis and if possible, also identify a structure dipole (Schwarz et al., 2016). To our current knowledge, these features seem to be real, but the low statistical significance is likely due to statistical selection biases.

The Cosmological Principle has been a simplifying assumption before it was justified, both theoretically or with observations. It is also a natural prediction of inflation models. Observations show that the Cosmological Principle is quite a reasonable approximation. However, it still exists as a sort of assumption (Gibelyou and Huterer, 2012). Few thorough observational tests have been applied to test homogeneity and isotropy. We will attempt to contribute to this. We will present some work on statistical isotropy tests of the Universe using radio data with the goal of bringing statistical isotropy from the assumption domain into the observational domain.

Other observational projects have made more detailed observations by mapping out the structure of the Universe at its largest scales during the early times, which is roughly 380,000 years after the Big Bang (Deruelle and Uzan, 2018). These results help scientists to further improve the constraints on cosmological parameters. It was at this time when the first hydrogen atoms formed setting photons in motion to travel through the Universe. These photons are today detected as the Cosmic Microwave Background (CMB). The Cosmic Microwave Background provides strong evidence of the Big Bang origin of our Universe.

Analysis of the data from observations of large-scale structures and the microwave background radiation help justify the notion of homogeneity and isotropy of the Universe and they help us constrain the Lambda-CDM parameters better (see figure 2.1). The fluctuations are consistent on angular scales that are larger than the cosmological horizon at recombination. This leads to either of the two, (i) the consistency is causally fine-tuned or (ii) cosmic inflation did happen (Dodelson, 2003).

The aim of this thesis is to test the statistical isotropy of the Universe using radio data, more specifically the NRAO VLA Sky Survey (NVSS) catalog. We test this isotropy by comparing the data to mock realisations of what an isotropic Universe would look like. We use statistical estimators to test the isotropy on different scales. Furthermore, we also test the statistical significance of our results by using our estimators on simulated radio data. Then we conclude on the isotropy of our data sample and the significance of the results we obtain, also how the test could be improved.

Thesis outline

Chapter 1 gives an introduction to the topic, gives a bit of some background to the subject. In Chapter 2, we start off with an overview of cosmology in which we discuss the standard model of cosmology, the parameters associated with the standard model and discuss the expansion of the Universe. We discuss the Cosmological Principle in detail focusing on statistical isotropy. Chapter 3 outlines the radio instrument that was used to collect the data, a discussion on radio continuum data, the preparation of our data set, the statistical estimators used in this work and the analysis to be carried out in testing statistical isotropy. This is followed by Chapter 4 where we present the results, discussion and also comment on how they can be improved. Finally, in Chapter 5 we conclude on our findings and give suggestions on possible future work.

Chapter 2

Cosmology Overview

2.1 Standard Model of Cosmology

The standard model of Cosmology is also known as Lambda-CDM which is short for Lambda-Cold Dark Matter. This is a way of expressing the Big Bang model in which the Universe contains a cosmological constant called lambda, which is associated with dark energy. The Lambda-CDM model accounts for most of the observational properties of the Universe, e.g. the late-time accelerated phase of the Universe preceded by a matter-dominated era.

All these properties allow us to explain the large-scale distribution of cosmic objects, the Cosmic Microwave Background, and the distance to standard candles and rulers. Due to this, Lambda-CDM is hereafter referred to as the standard model of Cosmology. It can also be referred to as the concordance model.

The Λ is a free parameter in Einstein's fundamental equation of general relativity. The conservation of energy-momentum is contained in the equation $T^{\alpha\beta}{}_{;\beta} = 0$, the $T^{\alpha\beta}$ term is the stress-energy tensor. This term together with the Einstein tensor also obeys the same equation $G^{\alpha\beta}{}_{;\beta} = 0$. This leads to a relation between gravitational acceleration and the curvature of space-time. This is a set of equations that describe the gravitational field in terms of curved space-time,

$$R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R + \Lambda g_{\alpha\beta} = 8\pi G T_{\alpha\beta} \tag{2.1}$$

where $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta}$. On the right hand side, G is the gravitational constant and $T_{\alpha\beta}$ can be expressed as

$$T_{\alpha\beta} = (\rho + p)u_\alpha u_\beta - pg_{\alpha\beta} \tag{2.2}$$

The ρ term is the mass-energy density, p is the hydrostatic pressure, u_α is the fluid's four-velocity, and $g_{\alpha\beta}$ is the metric tensor's reciprocal. This equation suggests that a non zero Λ term accounts for the energy of a vacuum in order to have an isotropic and homogeneous Universe. It assumes general relativity to be the correct theory for describing gravity on cosmological scales. The Lambda is associated with dark energy, it drives the accelerating expansion of the Universe and has a negative pressure associated with it. i.e $p = -\rho c^2$.

One of the Friedmann equations, which is derived using Einstein's equations is given by

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} + \frac{\Lambda}{3} \tag{2.3}$$

where the last term contains the cosmological constant Λ, ρ is the total energy density and K is the curvature coefficient which is zero in a flat Universe, negative in an open Universe and positive in a closed Universe. We consider a flat Universe with $K = 0$. The total energy density is

$$\rho = \rho_m + \rho_r + \rho_\Lambda \tag{2.4}$$

for the matter, radiation and dark energy densities. Then equation 2.3 reduces to

$$H^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} \tag{2.5}$$

The Hubble constant for today is defined as H_0, dividing equation 2.5 by the Hubble constant for today, we get

$$\frac{H^2}{H_0^2} = \frac{8\pi G}{3H_0^2}\rho + \frac{\Lambda}{3H_0^2} \tag{2.6}$$

We can define dimensionless density parameters for matter and dark energy as

$$\Omega_m = \frac{8\pi G}{3H_0^2}\rho_0, \qquad \Omega_\Lambda = \frac{\Lambda}{3H_0^2} \tag{2.7}$$

then from equation 2.6 we get

$$1 = \Omega_m + \Omega_\Lambda \tag{2.8}$$
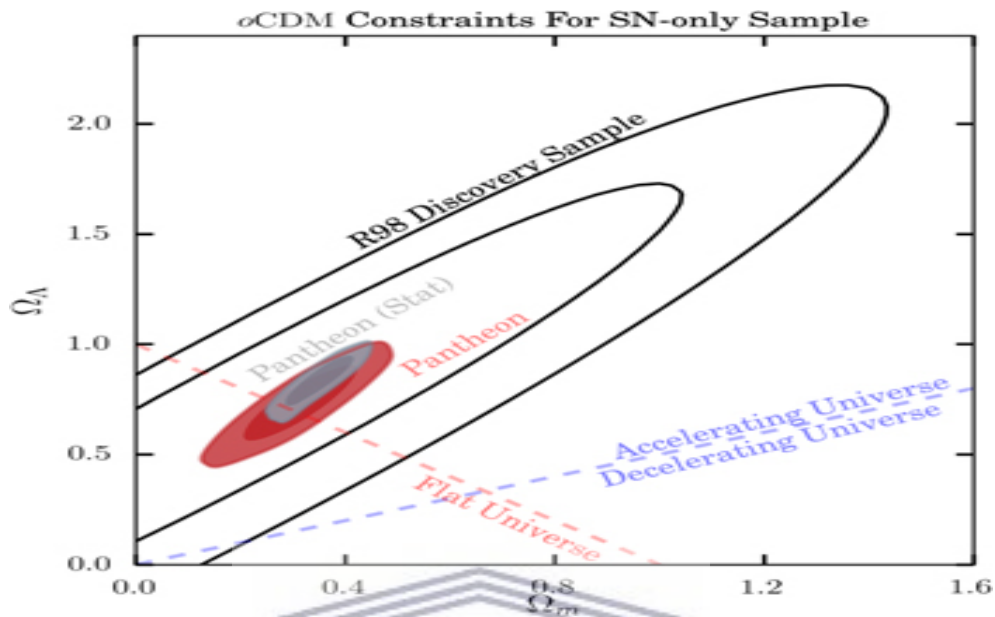
FIGURE 2.1: Evidence for dark energy from SN-only constraints. Here we show confidence contours at 68 % and 95 % for the Ω_m and Ω_Λ cosmological parameters for the **Riess et al. (1998)** discovery sample and the Pantheon sample. The Pantheon constraints with systematic uncertainties are shown in red, and those with only statistical uncertainties are shown in gray (line) (Scolnic et al., 2018).

The Lambda-CDM model has six base parameters. The Planck 2015 analysis showed excellent consistency between the temperature power spectrum and the Lambda-CDM which is specified by six base parameters. Table 2.1 shows the best constraints of the six base parameters from the 2018 Planck collaboration cosmological parameters with 68 % confidence limits for the standard Lambda-CDM model (Aghanim et al., 2018).

The six base parameters for the Lambda-CDM model are, the physical baryon density parameter, a physical dark matter density parameter, the age of the Universe, scalar spectral index, curvature fluctuation amplitude and the re-ionization optical depth.

A complete list of assumptions underlying the Lambda-CDM model are discussed thoroughly by (Akrami et al., 2018), so we will not discuss them here. The reader may look at the paper for some further reading on these assumptions and in-depth

Parameter	Planck	Planck + BAO
$\Omega_b h^2$	0.02237 ± 0.00015	0.02242 ± 0.00014
$\Omega_c h^2$	0.1200 ± 0.0012	0.11933 ± 0.00091
$100\theta_{MC}$	1.04092 ± 0.00031	1.04101 ± 0.00029
τ	0.0544 ± 0.0073	0.0561 ± 0.0071
$ln(10^{10}A_s)$	3.044 ± 0.014	3.047 ± 0.014
n_s	0.9649 ± 0.0042	0.9665 ± 0.0038

TABLE 2.1: The 6-parameter Lambda-CDM model that best fits the combination of data from Planck CMB temperature and polarization power spectra (including lensing reconstruction), with and without Baryon Acoustic Oscillations (BAO) data, for more details see (Aghanim et al., 2018)

information. More cosmological parameters can be derived from those in table 2.1, like t_0, Ω_Λ.

Although the Lambda-CDM model is widely accepted, there are alternative models that have been proposed like dynamical dark energy, quintessence, f(R), in-homogeneous models like Lemaitre-Tolman-Bondi models and back reaction (Clifton et al., 2012). However, looking for such alternatives will not be pursued in this thesis. Early criticism on the dark matter model consists of the Modified Newtonian Dynamics (MOND) (Milgrom, 1983), but not much compelling evidence has been found thus far.

The standard model implies that most of the matter in our Universe is dark matter with most of the energy in the form of dark energy. This is shown in figure 2.2, which is a graphical representation of the standard model of cosmology.

2.2 The Expanding Universe

To better understand how the Universe evolves, we need an expression for the scale factor. We also have to know the curvature of the Universe. Inserting the Friedmann Robertson Walker (FRW) metric into Einstein's field equations they reduce to what is known as the Friedmann equations.
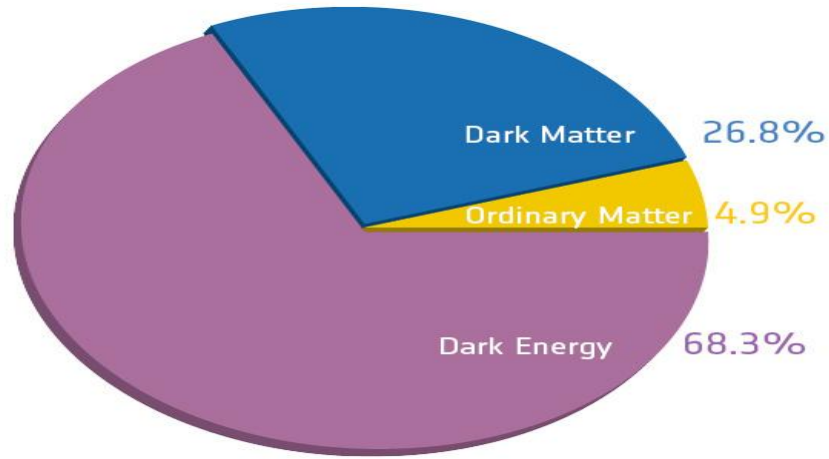
FIGURE 2.2: A pie chart depicting the standard model of cosmology
(Credits:ESA and the Planck Collaboration team)

The Robertson Walker (RW) metric is given by

$$ds_{RW}^2 = c^2 dt^2 - a^2(t)\gamma_{ij}dx^i dx^j \tag{2.9}$$

The γ_{ij} term is the three dimensional spatial metric of constant curvature. We consider the simplest form for a flat Universe with zero curvature. Then the RW metric becomes

$$ds_{RW}^2 = c^2 dt^2 - a^2(t)(dx^2 + dy^2 + dz^2) \tag{2.10}$$

In spherical coordinates $x = r\cos\varphi\sin\theta$, $y = r\sin\varphi\sin\theta$ and $z = r\cos\theta$. The RW metric then becomes

$$ds_{RW}^2 = c^2 dt^2 - a^2(t)(dr^2 + r^2 d\Omega^2), \tag{2.11}$$

the solid angle $d\Omega = d\theta^2 + \sin^2\theta d\varphi^2$. For a generalized case with nonzero spatial curvature, the RW metric becomes

$$ds_{RW}^2 = c^2 dt^2 - a^2(t)(dr^2 + S_K^2(r)d\Omega^2) \tag{2.12}$$

The three cases for the curvature K, are depicted below. From now on we assume a flat Universe, i.e $K = 0$

$$S_K(r) = \begin{pmatrix} r: K = 0 \\ \sin r: K = 1 \\ \sinh r: K = -1 \end{pmatrix}$$

Observations support the idea that the Universe is expanding. Hubble found the relation between the recession velocity and the redshift of galaxies. Assuming that $H_0 R \approx cz$, the distances to these galaxies can be estimated. This led to the formulation of Hubble's Law $v = H_0 R$. The v is the recession velocity, H_0 is the Hubble's constant and $R = a(t)r$ is the distance to the galaxies. If we consider low redshift, then we have

$$v = \frac{dR}{dt} = \frac{d(a(t)r)}{dt} = r\dot{a} = H_0 R,$$

where r is the comoving distance of the receding galaxy and $R = ar$ is the physical distance.

The expansion of the Universe mainly involves two aspects: the energy density of the Universe and the Hubble rate. The latter describes the speeds at which galaxies are moving away from us. This expansion rate is given by

$$H = \frac{\dot{a}(t)}{a(t)}, \tag{2.13}$$

where a(t) is the scale factor and $\dot{a}(t)$ is the time derivative of the scale factor with respect to proper time. Physical separations between two points (galaxies) all expand with the same scale factor a(t). Consider two galaxies that are separated by a distance χ today. If their motion is only due to the Hubble rate, then at a different time t_1, the separation between them was $a(t_1)\chi$. The scale factor today is $a_0 = a(t_0) = 1$.

We can relate the scale factor to the redshift. The comoving wavelength is given by $\lambda_{emi} = a(t)\lambda_{obs}$. This is the wavelength emitted at time t. Redshift is defined as

$$z = \frac{\lambda - \lambda_{obs}}{\lambda_{obs}} \tag{2.14}$$

This equation then implies that

$$1 + z = \frac{\lambda_{obs}}{\lambda_{emi}}, \tag{2.15}$$

by substituting λ_{emi}, we get the following relation between scale factor and redshift

$$1 + z = \frac{1}{a(t)} \tag{2.16}$$

The Friedmann equation 2.3 gives the expansion rate in terms of both the matter, radiation and dark energy.

$$H^2 = (\frac{\dot{a}}{a})^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3}, \tag{2.17}$$

We define the present day density parameters in equation 2.7 where we only consider a late time flat Universe and neglect the radiation contribution ($\Omega_r = 0$). The density scale as a^{-3} for matter, the Friedman equation becomes

$$H(a) \equiv \frac{\dot{a}}{a} = H_0\sqrt{\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_\Lambda}, \tag{2.18}$$

which simplifies to

$$H(z) = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda} \tag{2.19}$$

in terms of redshift for the late Universe. Now we relate the comoving distace $D_c(z)$ to the redshift z. Comoving distance in terms of redshift is given as

$$D_c(z) = c\int_0^z \frac{dz}{H(z)} \tag{2.20}$$

The luminosity distance is given by

$$D_L(z) = (1+z)D_M(z) \tag{2.21}$$

But we know that $D_M(z) = D_c(z)$ for zero curvature $\Omega_k = 0$, so equation 2.20 can be written as

$$D_M(z) = D_c(z) = c \int_0^z \frac{dz}{H(z)}. \tag{2.22}$$

Now the luminosity distance becomes

$$D_L(z) = (1+z) \times c \int_0^z \frac{dz}{H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}} \tag{2.23}$$

Finally, the luminosity-redshift relation for a matter dominated Universe is given by

$$\mu(z) = 5\log D_L(z) + 25 \tag{2.24}$$

The μ is defined as the distance modulus, this is a measure of the distance to an object given by the difference in apparent magnitude and absolute magnitude of the object, $\mu = m - M$.
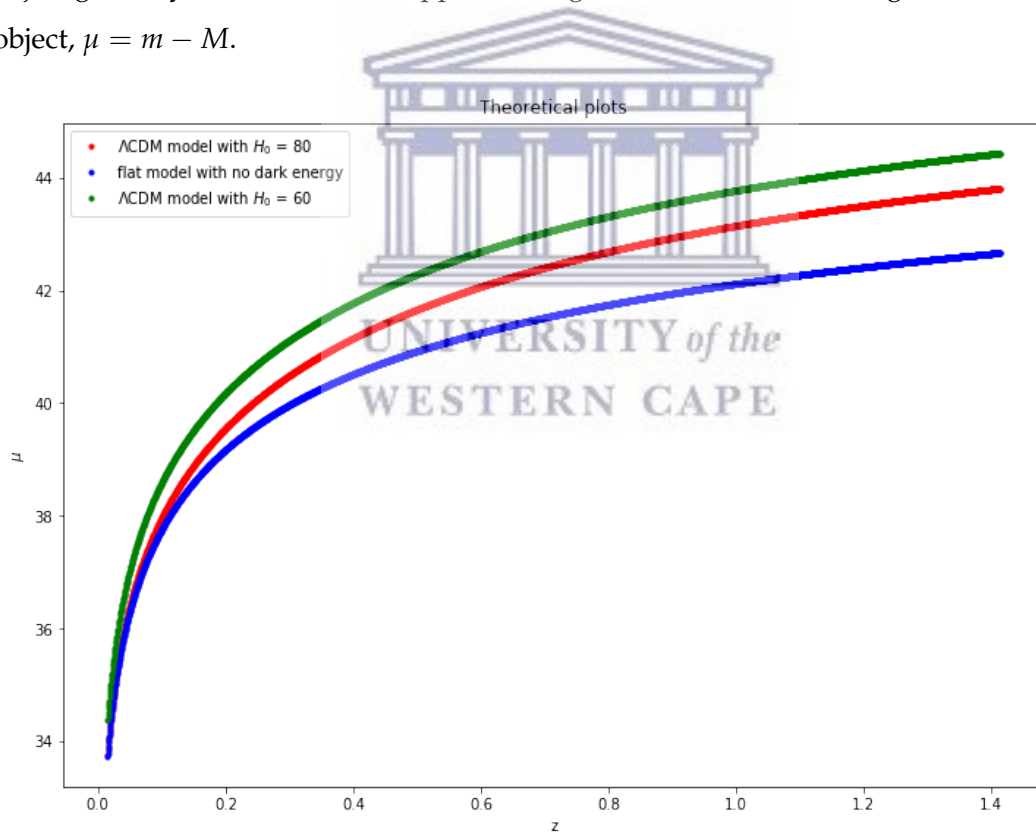
FIGURE 2.3: The luminosity-redshift relation for the late Universe.

The red and green curves in figure 2.3 both have the same parameters with ($\Omega_m = 0.3, \Omega_\Lambda = 0.7$) the only difference is H_0. The curve with a higher value of H_0 (red curve) is below the curve with a lower value of H_0 (green curve). The blue curve represents a flat Universe with no Dark Energy, we can see that it is below the other

two curves, irrespective of having a value of H_0 = 80 km/s/Mpc. Also, note that at higher redshifts the three curves deviate from one another as compared to lower redshift (0 < z < 0.2). The plots in figure 2.3 show the relationship between the luminosity distance and redshift and also how the Hubble parameter affects the luminosity and redshift.

Hubble's constant is the expansion rate of the Universe, that is the rate at which the scale factor changes with time. Recent advancements in technology and high precision experiments have given better constraints on the value of H_0, see figure 2.4 for the different values of H_0. The results are a major advance on that of Edwin Hubble's measurement of approximately 500 km/s/Mpc (Dodelson, 2003).
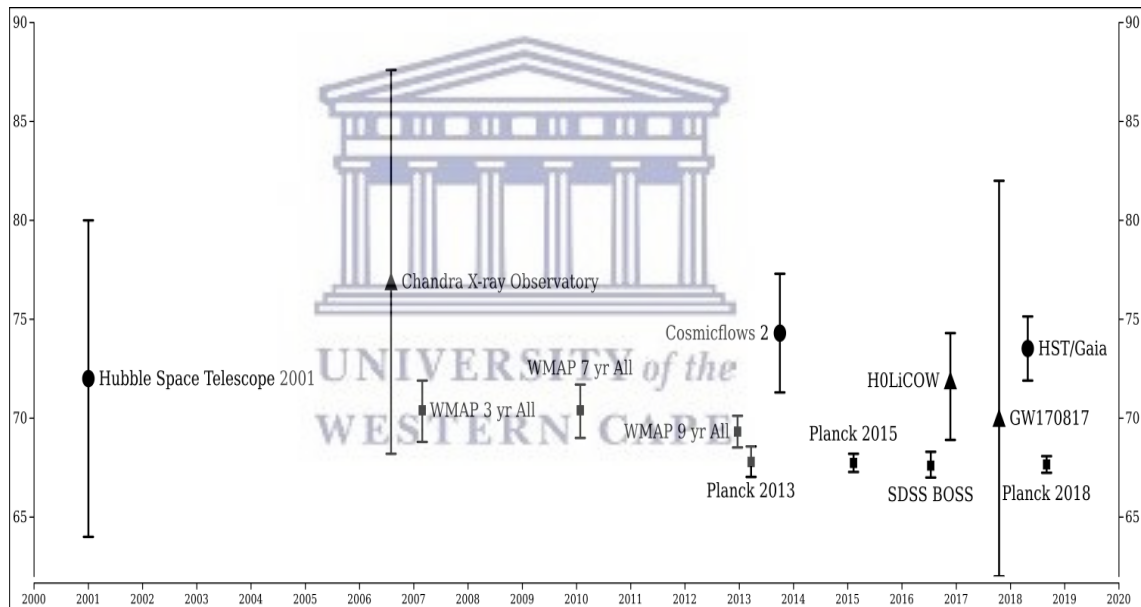
FIGURE 2.4: Estimated values of the Hubble constant H_0, from 2001–2018. Circles represent calibrated distance ladder measurements, squares represent early Universe CMB/BAO measurements with ΛCDM parameters while triangles are independent measurements.(Credit:aasnova.org)

This notion of an expanding Universe gave birth to the theory of the Big Bang. If space is expanding, we can infer that in the past the Universe must have been hotter than it is presently. The temperature tends to infinity in a finite time in the past, this is a phenomenon known as the Big Bang singularity. This was a singularity of infinite density. The Big Bang effectively happened everywhere simultaneously so there should still be traces of it left in the Universe as some background radiation,

now known as the Cosmic Microwave Background radiation (CMB) (Ryden, 2016) (see figure 2.5).

The temperature of the CMB as a function of redshift is given by

$$T(z) = T_0(1 + z). \tag{2.25}$$

This equation shows that the Universe cools as it expands. During the expansion of the early Universe, matter decoupled from radiation at a temperature of 3000 K. The radiation has been Doppler shifted to approximately 2.7 K today. The measurements today show a nearly isotropic radiation background of 2.7 K. It has a spectrum similar to that of a black body. The CMB spectral radiance $dE_\nu/d\nu$ peaks at a frequency of 160 GHz in the microwave range of the electromagnetic spectrum (Dodelson, 2003).
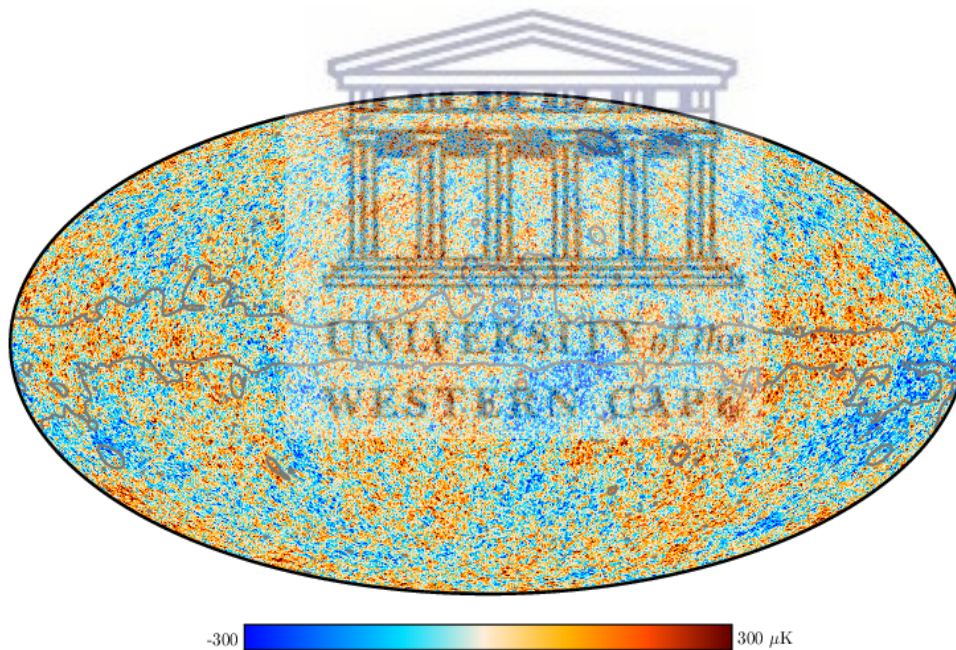
FIGURE 2.5: The Planck CMB sky as shown by the 2018, SMICA temperature map. The CMB map has been masked and is painted in regions where residuals from foreground emission are expected to be substantial. This mask, mostly around the Galactic plane, is shown by a grey line in the full resolution temperature map, the unmasked area covers 80.7 % of the sky (Akrami et al., 2018)

2.3   Cosmological Principle

The Copernican Principle states that we do not occupy a special place in the Universe (Ellis, 2006). It was one of the building blocks for the Cosmological Principle. If our local Universe is isotropic to us as the observer, then the Copernican Principle requires that other observers in the Universe also see isotropy, otherwise we would have a special place in the Universe. A Universe that is isotropic everywhere must also be homogeneous, thus the Cosmological Principle is a conclusion from using isotropy and the Copernican Principle.

It says for an observer on any galaxy at rest relative to the CMB, the Universe on average looks statistically similar in all directions. The term observer means any observer in any part of the Universe. So regardless of wherever one is, the Universe looks nearly the same. This requires the geometry of the Universe to be a simplified one. The three possible geometries of space have been discussed in section 2.1, they are the open, closed and flat geometry. The simplest space-time metric for a flat isotropic and homogeneous Universe is shown by equation 2.10. It shows that the scale factor a(t) describes the evolution of the Universe with time.

Although it has formed the basis of modern cosmology, the Cosmological Principle still needs to be justified with observational data. Obviously, there are noticeable inhomogeneities on the scales of our solar system or galaxy. The assumption of isotropy can be directly tested from observations, but on the other hand, homogeneity cannot be directly tested from observations since we cannot observe within the past light cone (Clarkson and Maartens, 2010). This phenomena is illustrated in figure 2.6 below.

Astronomical data has grown massively, now the Cosmological Principle can be discussed using detailed observations. Galaxy surveys like the Sloan Digital Sky Survey (SDSS) have been used to show that the Universe at very large scales is consistent with isotropy and homogeneity, see Gonçalves et al., 2018 and references therein. These maps of galaxies are also important for the standard model of cosmology.
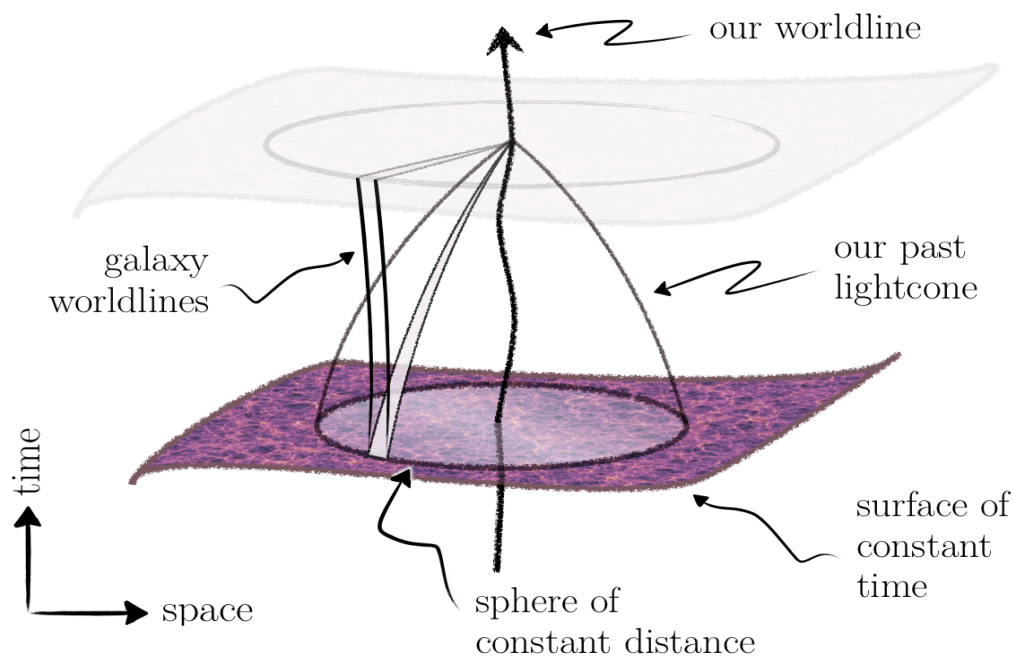
FIGURE 2.6: The Copernican Principle is hard to test because we are fixed to one event in space-time. We make observations on our past null cone which slices through spatial surfaces (Clarkson, 2012)

Detailed studies from Planck showed the isotropy of the Universe by showing how isotropic the cosmic microwave background is, see figure 2.5. Some studies are still trying to find the CMB dipole signal in source number counts across the entire sky, but with very limited precision at the moment. Their results give a dipole amplitude that is larger than the amplitude seen in the CMB, even though the direction is consistent, see Bengaly, Maartens, and Santos, 2018. The dipole is shown in figure 2.7.

The CMB dipole has an amplitude of $\sim 10^{-3}$. We can interpret it to be due to our relative motion so that we get a degree of isotropy to be as high as $\sim 10^{-5}$ after removing both the dipole and monopole (Kogut et al., 1993).

2.4 Large Scale Structure

The Universe is full of matter. But how is all of this matter distributed? Is it all arranged together, like when freezing people gather around a campfire? Or sprinkled evenly, like raisins in a cake? As it turns out, galaxies seem to huddle together in
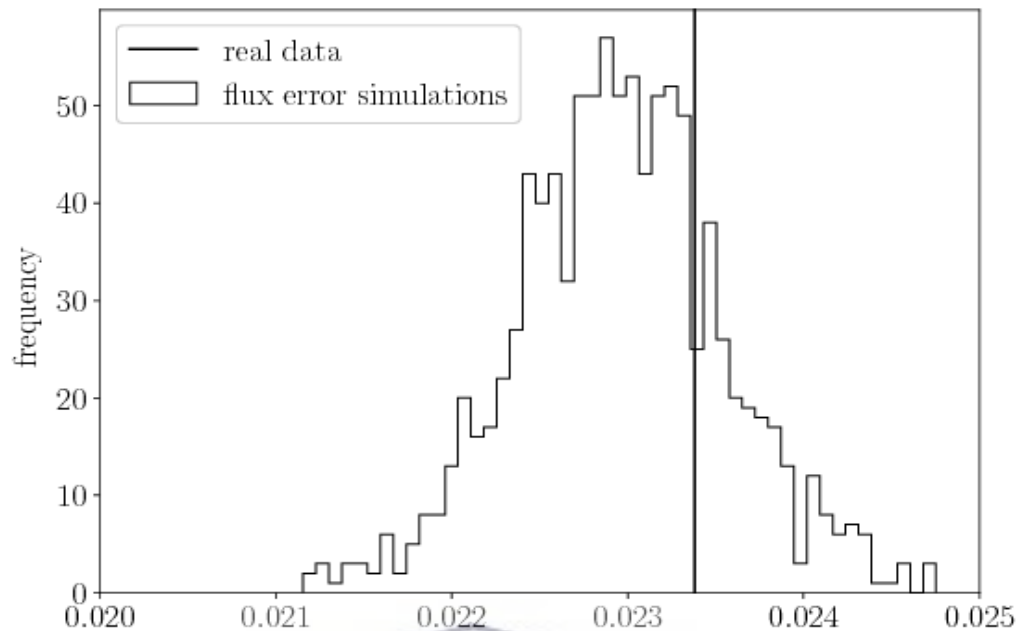
FIGURE 2.7: Dipole amplitude histogram for the NRAO VLA Sky
Survey. The vertical line represents the value of the dipole obtained
from real data (Bengaly, Maartens, and Santos, 2018).

larger groups called galaxy groups and clusters. Many of which contain a mass of
about 10^{14} solar masses and are up to several million light-years across.

After decoupling, the baryonic matter was able to collapse under gravity. This led to
the formation of the first stars, then galaxies and then ultimately clusters of galaxies.
As more matter fell in due to gravity, the over dense regions became more and more
over dense. The over-densities grew big enough for stars to form. In the standard
Lambda-CDM model, quantum fluctuations provide the inhomogeneities needed
for structure formation.

In figure 2.8 one can see that galaxies and galaxy clusters are also not evenly dis-
tributed. Rather, they are arranged in galaxy clusters and super clusters, which are
a result of many galaxy clusters clumped together. The super clusters can contain
thousands of galaxies, and extend over millions of light years. Super clusters are
connected by thin, fibre-like structures known as filaments. There are also large
empty regions called voids, which can also be millions of light years across, and
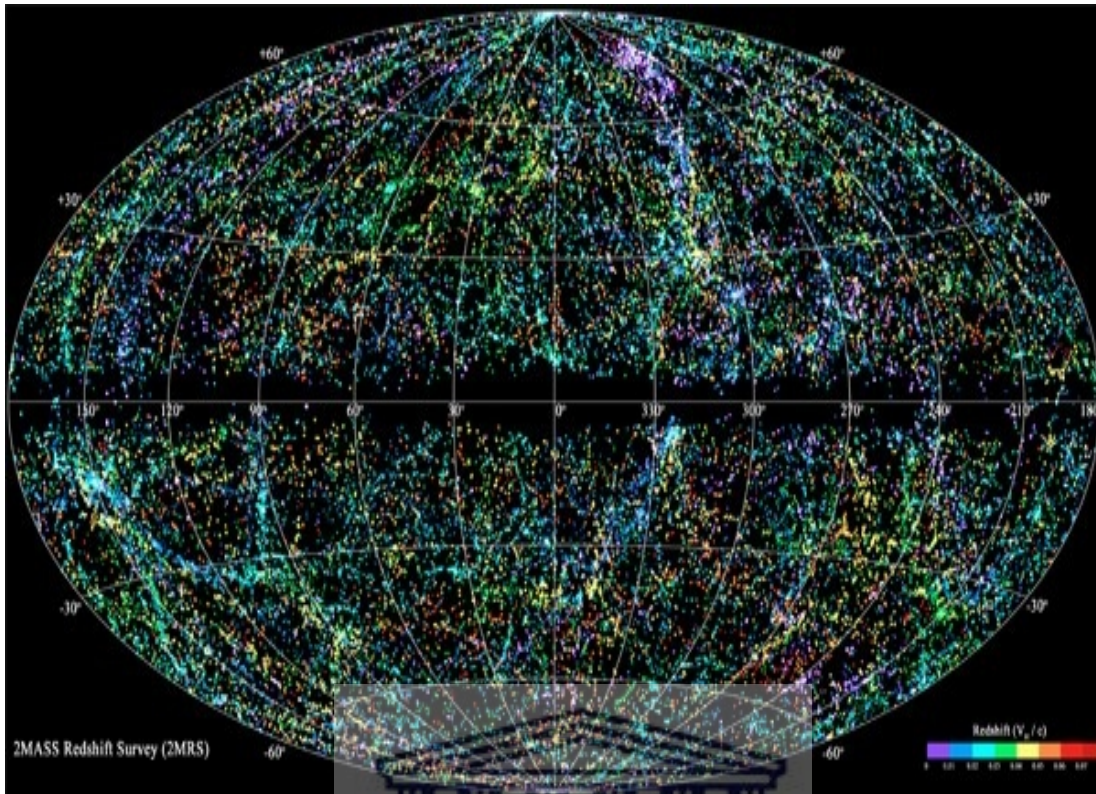
FIGURE 2.8: Distribution of 2MRS catalog galaxies in Galactic coordinates. (Murph, 2011)
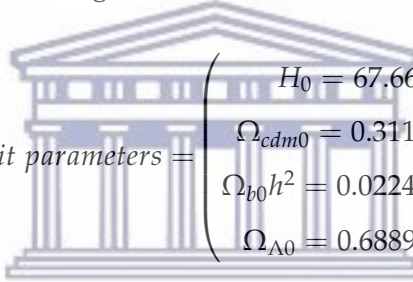
contain none, or at least almost no galaxies.

Large-scale structures are a result of small over-densities which start to grow in a matter dominated era due to gravitational instabilities. For galaxies to grow fast enough, we need a mysterious component in the Universe such as the non-baryonic cold dark matter, otherwise, we can't explain the observations. That is the main reason why it is part of the standard model of Cosmology. Vacuum fluctuations of the inflation field provide the seeds for these small over-densities. No evidence has been found yet against the Gaussian spectrum assumption for the primordial density perturbations (Ade et al., 2016). Future galaxy surveys will be able to put better constraints similar to that of Planck's and thus establish if this assumption is really valid. Refer to chapter 4 of Ballardini et al., 2016 to read up on some of the future galaxy surveys.

Observations of the CMB at decoupling ($z \approx 1100$) have been critical to measure the basic properties of the standard model of the Universe. This standard model of

cosmology is a flat FRW model with dark energy modelled by Λ, cold dark matter dominating over baryonic matter, and initial conditions for structure formation generated by a simple inflation model. Galaxy surveys at low redshift provide a crucial complement to the information from the CMB. In particular, galaxy surveys provide the best constraints on dark energy because dark energy only starts to dominate at $z < 1$.

The best current constraints on Lambda-CDM come from the Planck survey and the Sloan Digital Sky Survey (SDSS-III) Baryon Oscillation Spectroscopic Survey (BOSS), together with smaller data sets such as those measuring the properties of supernovae of type Ia. The various constraints have been combined in the 2018 data released by the Planck collaboration (Aghanim et al., 2018). The basic parameters of the Lambda-CDM model have the following values:

$$Best - fit\ parameters = \begin{pmatrix} H_0 = 67.66 \pm 0.42 \\ \Omega_{cdm0} = 0.3111 \pm 0.0056 \\ \Omega_{b0}h^2 = 0.02242 \pm 0.00014 \\ \Omega_{\Lambda 0} = 0.6889 \pm 0.0056 \end{pmatrix}$$

These constraints rely heavily on the measurement of the power spectrum of CMB temperature anisotropies. The angular power spectrum measures amplitude as a function of wavelength. The power spectrum helps characterize the size of the fluctuations as a function of the angular scales and of the correlation function of galaxies. (see figure 2.9). The two point correlation function gives us the probability of finding a galaxy within a given distance from another galaxy. The two point correlation function is widely used to measure the large scale structure in a galaxy, it measures the amplitude of clustering versus the scale (see figure 2.10).
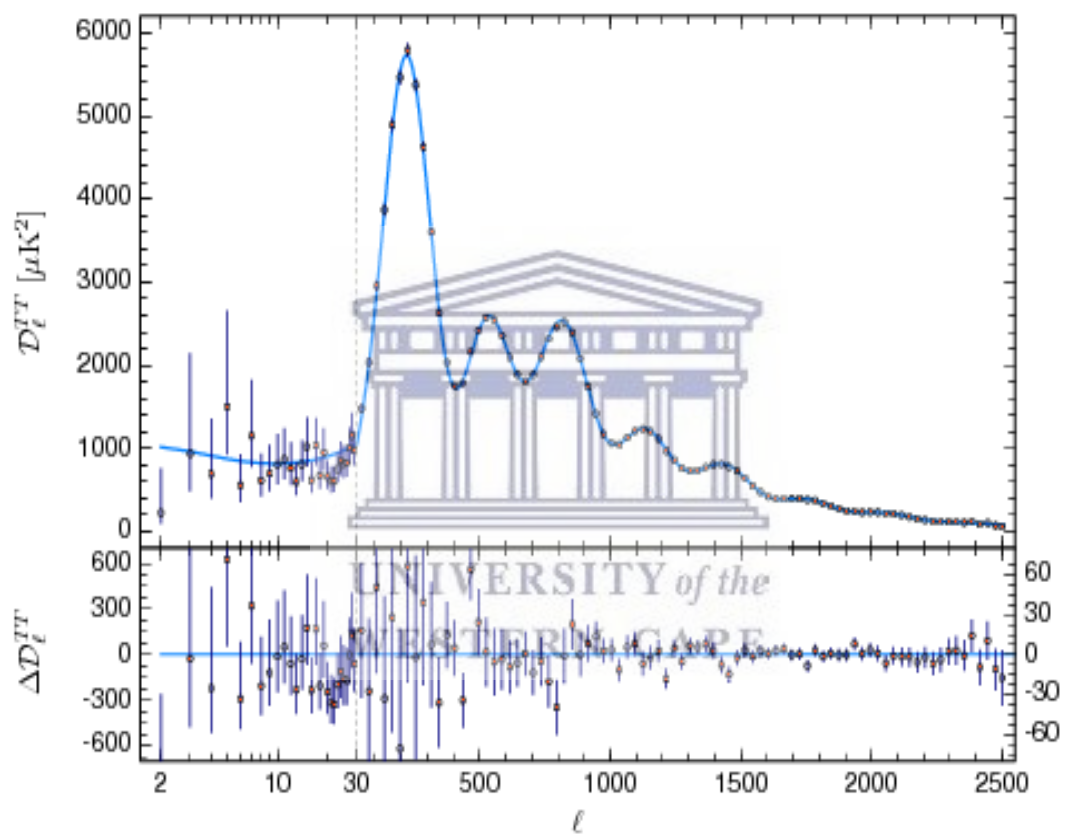
FIGURE 2.9: Planck 2018 temperature power spectrum, with foreground and other nuisance parameters fixed to a best fit assuming the base-ΛCDM cosmology. Taken from Aghanim et al., 2018
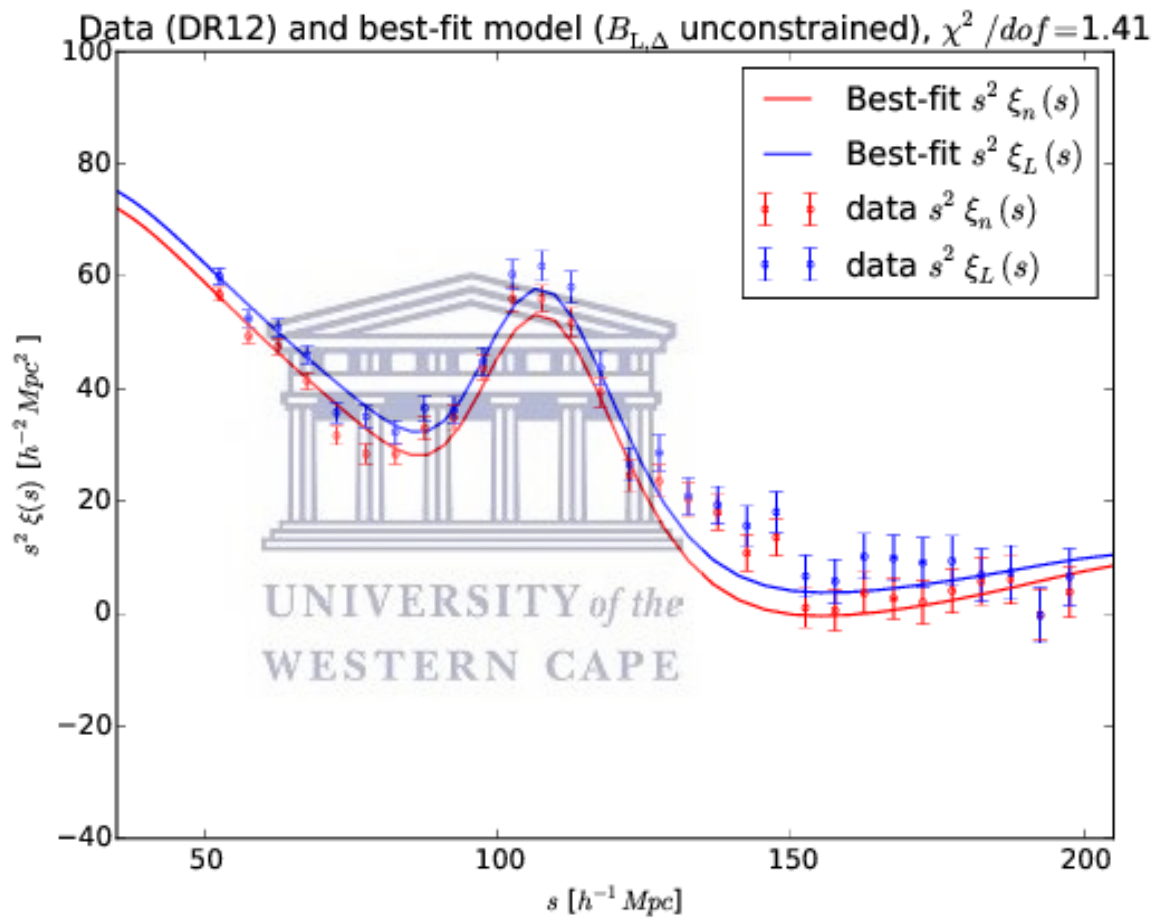
FIGURE 2.10: Measurement of the correlation functions ξ_n and ξ_L using the CMASS-DR12. Taken from Soumagnac et al., 2018

Chapter 3

Radio Surveys, Statistical Tests and Analysis

Introduction

In this chapter, we give an overview of the radio instrument that was used to obtain the data, the data sample and all the methods, including the software we use in calculating galaxy number counts in order to test our hypothesis. We discuss also the limitations associated with each approach.
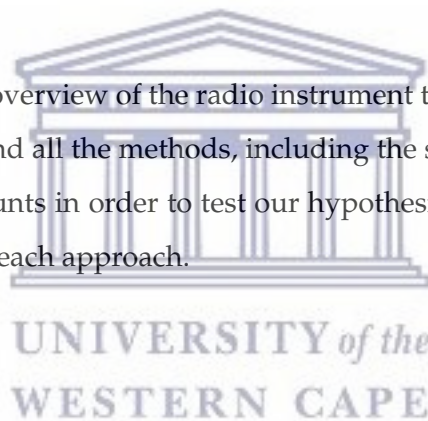
3.1 HEALPix Software

The Hierarchical Equal Area isoLatitude Pixelization (HEALPix) discretizes the celestial sphere into pixels that covers the same area given different resolutions and allows us to perform analysis on data projected on the celestial sphere. It provides a good scheme for distributing points as uniformly as possible over the surface of a unit sphere (Gorski et al., 2005).

Properties

One of the most important features of the HEALPix projections is that they are equi-areal with squared boundaries and straight parallels. The centres define locations

on the sphere, where $\theta \in [0, \pi]$ is the colatitude in radians measured from the North Pole and $\phi \in [0, 2\pi]$ is the longitude in radians measured Eastward. The centres are within smaller equal-sized diamonds that are identified as pixels. The area of each pixel is defined as

$$\Omega_{pix} = \frac{\pi}{3N_{side}{}^2} \tag{3.1}$$

as shown by Gorski et al., 2005, with N_{side} being there resolution of the map.

k	$N_{side} = 2^k$	$N_{pix} = 12N_{side}{}^2$	$\theta_{pix} = \Omega_{pix}{}^{1/2}$
0	1	12	$58.6°$
1	2	48	$29.3°$
2	4	192	$14.7°$
3	8	768	$7.33°$
4	**16**	**3072**	**$3.66°$**
5	32	12288	$1.83°$
6	64	49152	$55.0'$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
29	2^{29}	3.46×10^{18}	$3.93 \times 10^{-4''}$

TABLE 3.1: Table of N_{side}, the associated number of pixels and the corresponding pixel area for HEALPix (Gorski et al., 2005). The values in bold are for this work.

3.2 Radio Continuum

Radio continuum surveys scan the sky in radio frequencies in order to study the distribution and the intensity of radio sources in the sky. The fine details are determined by the amount of resolution of the radio telescope. The angular resolution of a telescope can be approximated by

$$R = \frac{\lambda}{D} \tag{3.2}$$

with λ being the wavelength of radiation and D is the diameter of the telescope (Serway and Jewett, 2018). Continuum surveys have the primary objective of understanding the formation of galaxies over cosmic time, large-scale structures and better constraints on cosmological parameters.

They provide an alternative view to that of optical instruments. They are able to detect hydrogen gas and also reveal some areas of the sky that might be blocked by interstellar dust. Figure 3.1 shows the 21 cm line map of the entire sky taken with a radio telescope. The long wavelength of radio waves enables them to be detectable even in cloudy conditions. For better resolution, the telescope must have a big diameter. This is why the radio dishes have to be very large. However, the bigger size introduces some limitations, for example, the physical size of the telescope, movement of the dish and the ability to point the telescope in certain directions. Moreover, the cost of building radio telescopes that have very big dishes is very high.
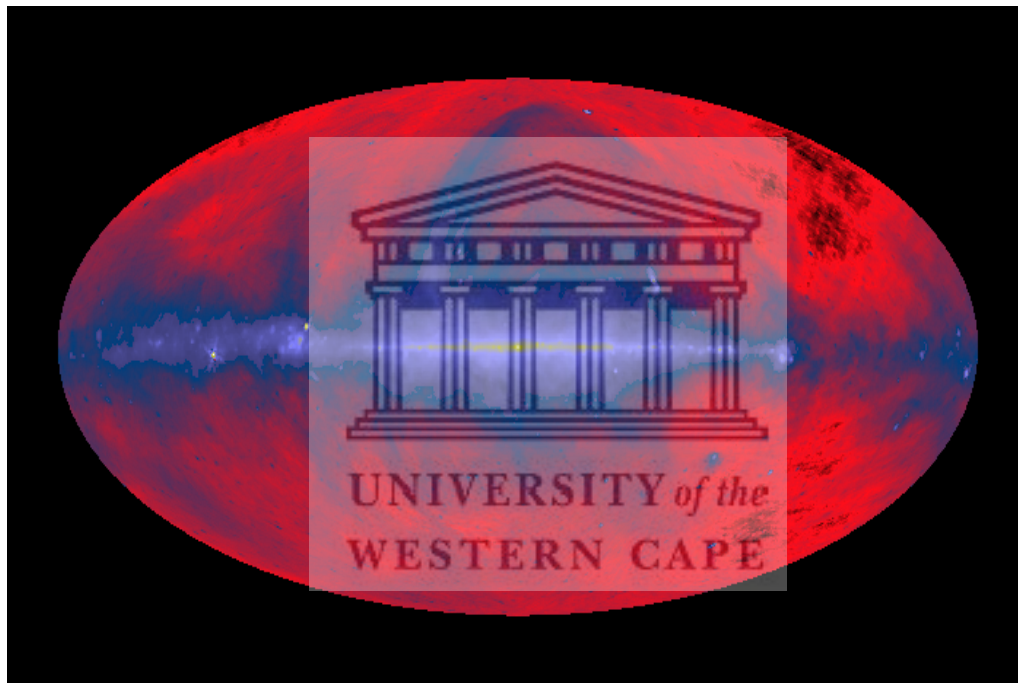
FIGURE 3.1: The sky seen at a radio frequency of neutral hydrogen
emission (21cm). (**haslam1982**)

A solution to some of these limitations is the application of a technique known as interferometry. This involves linking more than one radio telescope to create an interferometric array, see figure 3.2. The resolution of the array depends on the maximum separation between the individual dishes. Increasing the distance between the dishes increases the angular resolution of the array. The signal from the individual dishes is brought together and processed by a correlator, combining the signals to simulate a signal from a larger single telescope.

FIGURE 3.2:   The South African MeerKAT radio telescope
(www.ska.ac.za)

3.2.1   21cm emission

The proton and electron in the hydrogen atom have the spin property. The spin can either be aligned or anti-aligned. The atom will emit radio energy at a wavelength of 21 cm if there is a transition from the aligned to the anti-aligned energy state. Conversely, for the atom to make a transition from the anti-aligned to the aligned state, the atom has to absorb energy from the 21 cm wavelength. Figure 3.3, explains this entire process.
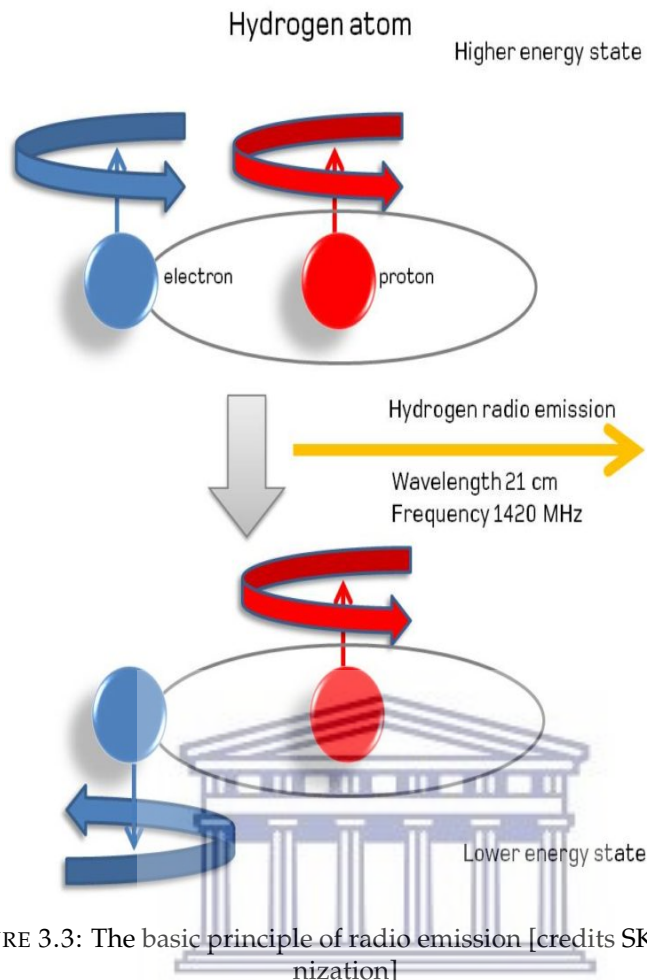
FIGURE 3.3: The basic principle of radio emission [credits SKA Organization]

3.2.2 Synchrotron radiation

Synchrotron radiation is a form of electromagnetic emission. It is emitted by electrons that are spiralling along, and therefore being constantly accelerated at speeds close to the speed of light in a magnetic field. Refer to figure 3.4 for a visual representation of the process. The emitted radio signature reveals the strength of the magnetic field. This form of radiation is continuous and highly polarized. The strength of the magnetic field together with the energy of the charged particles is directly related to the intensity and frequency of the radiation. Strong magnetic fields and high particle energies will produce a higher emitted intensity and frequency of radiation. Synchrotron radiation does not depend on the temperature of the source, even objects that are cool can emit synchrotron radiation. This is why it is sometimes referred to as non-thermal radiation.
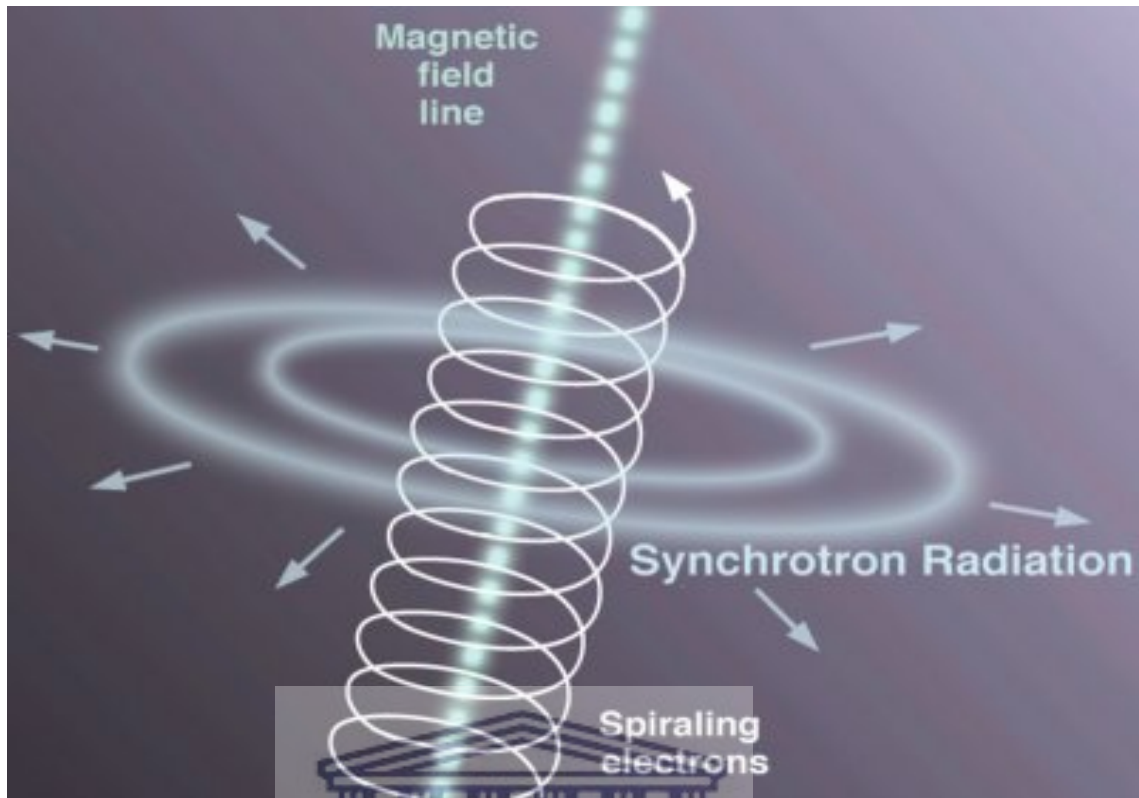
FIGURE 3.4: Synchrotron emission as a result of electrons spiralling
around magnetic field lines. [Source:www.daviddarling.info]

Astronomical sources that emit synchrotron radiation include, acceleration of elec-
trons in stellar systems, supernova remnants, cosmic ray electrons in magnetic fields
of normal galaxies, radio jets, and pulsars just to mention a few. In order to get an
emission spectrum from any of these sources, the emission spectra of individual
electrons must be summed up. An electron moving in a magnetic field emits many
frequencies that all have a peak value of v_0, known as the critical frequency. The
electron will lose more energy if it travels for a longer period in the magnetic field.
This will result in a longer wavelength corresponding to the critical frequency.

Synchrotron emission can be described by a characteristic spectrum, where we can
relate the flux and frequency using a power law:

$$S \propto v^{\alpha}, \tag{3.3}$$

α is the spectral index. Figure 3.5 shows the Crab Nebula which is a supernova
remnant. The blue glow at the centre of the Crab Nebula is believed to be due to
synchrotron emission.

FIGURE 3.5: The Crab Nebula (Source:www.spacetelescope.org)

For radio sources to emit synchrotron radiation, electrons must be spiralling at rela-
tivistic motions. The relativistic motion causes an opening angle that is related to the
Lorentz factor. The Lorentz factor boosts the emission frequency into a range that
can be observed with radio instruments. The maximum frequency can be measured
as

$$\nu_{max} = \nu_e \gamma^2, \tag{3.4}$$

where the Lorentz factor $\gamma = \frac{1}{\sqrt{1+\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-\beta^2}}$ (Forshaw and Smith, 2009).

For compact sources at a lower frequency, the power law spectrum becomes very
steep. These sources are generally optically thick. The source becomes opaque as
the electron temperature equals the brightness temperature. Then a phenomenon
known as synchrotron self-absorption occurs.

3.2.3   Radio flux

To describe the strength of the radiation and how it changes with distance between the source and the observer, we visit the concept of flux density. The specific intensity of a source can be defined by

$$I_v = \frac{1}{\cos\theta d\delta dv}\frac{dP}{d\Omega},$$ (3.5)

over a surface with infinitesimal area $d\delta$, solid angle $d\Omega$ and the angle between the radiation ray and the normal to the surface, θ. The power radiated by a photon over the solid angle denoted by $\frac{dP}{d\Omega}$. Over all frequencies, the intensity is

$$I = \int_0^\infty I_v dv.$$ (3.6)

For a source that subtends a well defined solid angle, the spectral power received by a detector is the flux density, S_v.

Integrating equation 3.5 over the solid angle gives

$$S_v = \int_{source} I_v(\theta, \phi)\cos\theta d\Omega,$$ (3.7)

for small angular sizes, $\cos\theta \approx 1$, thus the density is given by

$$S_v \approx \int_{source} I_v(\theta, \phi)d\Omega$$ (3.8)

measured in units of Jansky, where

$$1 Jansky = 1 Jy \equiv 10^{-26} Wm^{-2}Hz^{-1} \equiv 10^{-23} ergs^{-1}cm^{-2}Hz^{-1}$$

(Burke and Graham-Smith, 2009).

Some of the radio telescopes from around the world are the following APERTIF, ASKAP,eMERLIN, VLA,e-EVN, LOFAR, MeerKAT, MWA just to name a few. Once fully operational, the next generation telescope SKA will be the most powerful, sensitive and largest radio telescope.

3.3 Observational data

The Very Large Array

The Very Large Array (VLA) consists of 27 radio dishes which are 25 meters in diameter, mounted on rails. They are not all arranged in one line, see figure 3.6. The VLA is re-configurable and uses four principal array configurations A to D, depending on how far the individual dishes are moved apart. The A-configuration provides the longest baselines and thus the highest angular resolution for a given frequency, but yields very limited sensitivity to surface brightness. The D-configuration provides the shortest baselines, translating to a high surface brightness sensitivity at the cost of angular resolution. The VLA is located in a desert in New Mexico, USA.

FIGURE 3.6: The Very Large Array in New Mexico (Source:www.images.nrao.edu).

The NRAO VLA Sky Survey

The VLA was used to create the NRAO VLA Sky Survey (NVSS) (Condon et al., 1998) This is a catalog of radio sources with declination values above $-40°$. This survey was observed at a frequency of 1.4 GHz and contains almost two million sources down to a flux of 2.5 mJy. It is assumed that this survey is complete for all sources above 15 mJy (Condon et al., 1998). Due to limitations, some additional

identification problems arise near the galactic plane as well as near very strong radio sources. For declinations below $-10°$ and above $80°$ the DnC configuration of the array was used, while between those declinations the array was in the D configuration. This is because it is harder to observe radio sources near the horizon of the telescope. In order to compensate for this problem to some extent, the DnC configuration was used for most of the North and South areas.
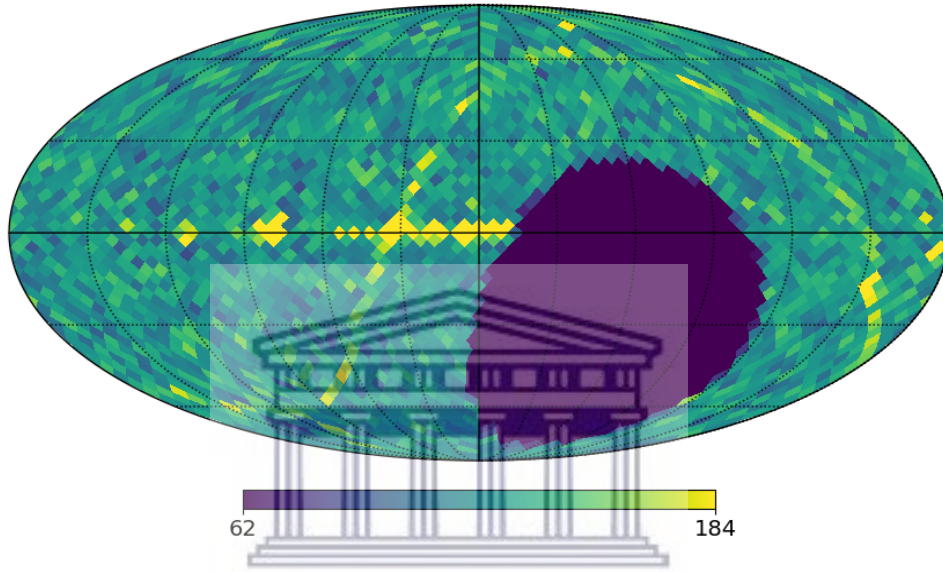
FIGURE 3.7: The NVSS catalog of radio sources in galactic coordinates, generated using HEALPix with $N_{side} = 16$ resolution.

3.3.1   Sample Selection

In order to perform a robust test of isotropy of the radio count angular distribution, one has to do a careful cleaning and selection of the data. We choose the following flux limits for our data,

$$20 \quad mJy < S_{NVSS} < 1000 \quad mJy$$

in order to avoid systematics below and above these values, also to be conservative. In order to purify the NVSS catalog, we adopt a masking procedure similar to the one from Bengaly, Maartens, and Santos, 2018, to deal with well-known contamination. This involves the elimination of pixels in the following regions:

- Close to the galactic plane, i.e., $\mid b \mid \leq 10°$.

- Within $1°$ of the local radio sources given in (Van Velzen et al., 2012).

- Galactic foreground emission above T = 50 K according to the 408 MHz continuum map in (Haslam et al., 1982).

N_{side}	No. of patches (Cetred at each of the 3072 HEALPix pixels)	Patch radii	f_{sky}	N_{Total}
16	3072	$10°, 15°, 20°, \& 25°$	~ 0.657	268 345

TABLE 3.2: Summary of the sample selection parameters used in this work. Column 1 lists the resolution used for our maps, column 2 lists the total number of patches that were selected for the analysis, the corresponding patch radius is depicted in column 3. Column 4 lists the total fraction of the sky, and in the last column we list the total number of sources in our sample.

To get rid of bad pixels referred to above, we use the mask shown in figure 3.8.
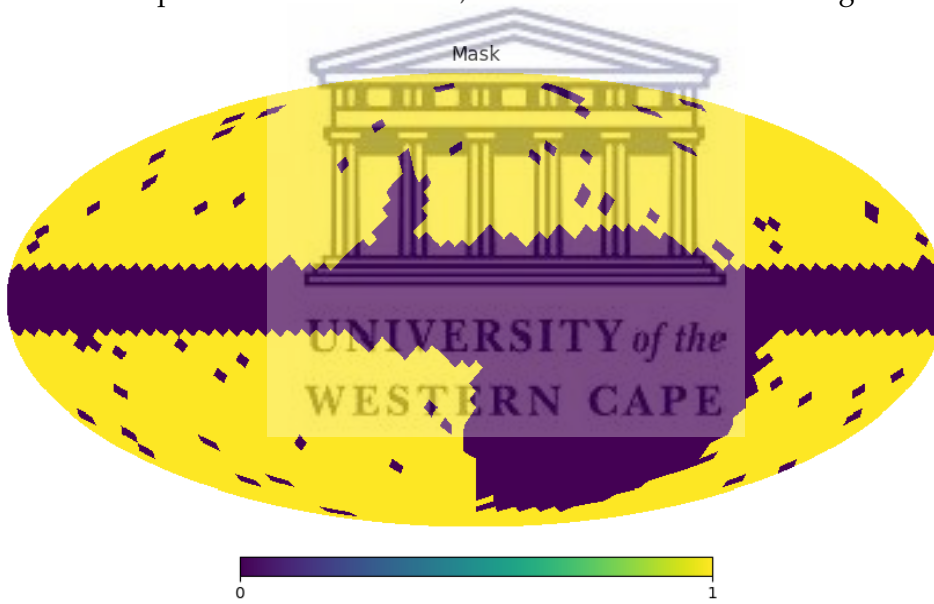
FIGURE 3.8: The mask used to eliminate bad pixels.

The final NVSS map is shown in figure 3.9 for $20 < S_{NVSS} < 1000$ mJy, at a resolution of $N_{side} = 16$, obtained by applying the mask in figure 3.8 to the NVSS catalog map in figure 3.7.

3.3.2 Patches

In order to test the isotropy of our data sample, we propose the following approach. We test the isotropy by analyzing the number counts in each of the pixels contained
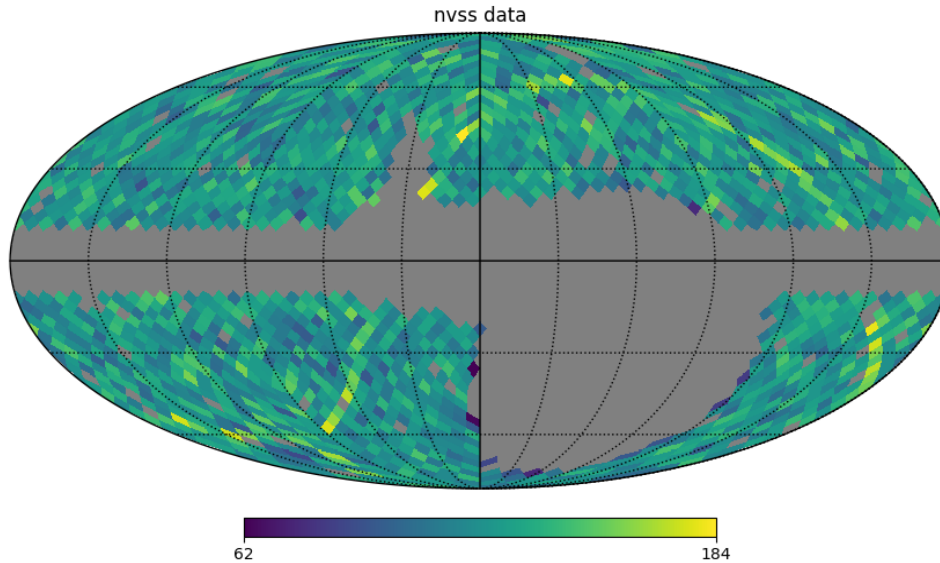
FIGURE 3.9: The pixelized number count map of NVSS radio sources
in the flux ranges $20 < S < 1000$ mJy.

in a patch on the sky, with a radius of $10°$. We repeat the procedure for three other
radii, which are $15°$, $20°$ & $25°$. Since we will apply a mask to bad pixels, we define
criteria for accepting patches that we are going to perform our tests on as follows.
For each patch of a given radius, we shall consider three cases. For each of the three
cases, we perform the test, obtain the results, then do some analysis before moving
on to the next radius. We only accept patches that

   i. have at least 90 % of the pixels unmasked.

   ii. have at least 70 % of pixels unmasked.

   iii. have at least 50 % of pixels unmasked.

For each radius, we compare the galaxy number counts for 3072 patches on the sky,
with the above acceptance criteria. Only pixels whose centre falls within the patch
and are unmasked are accepted. Those whose centre falls outside of the radius are
not considered.
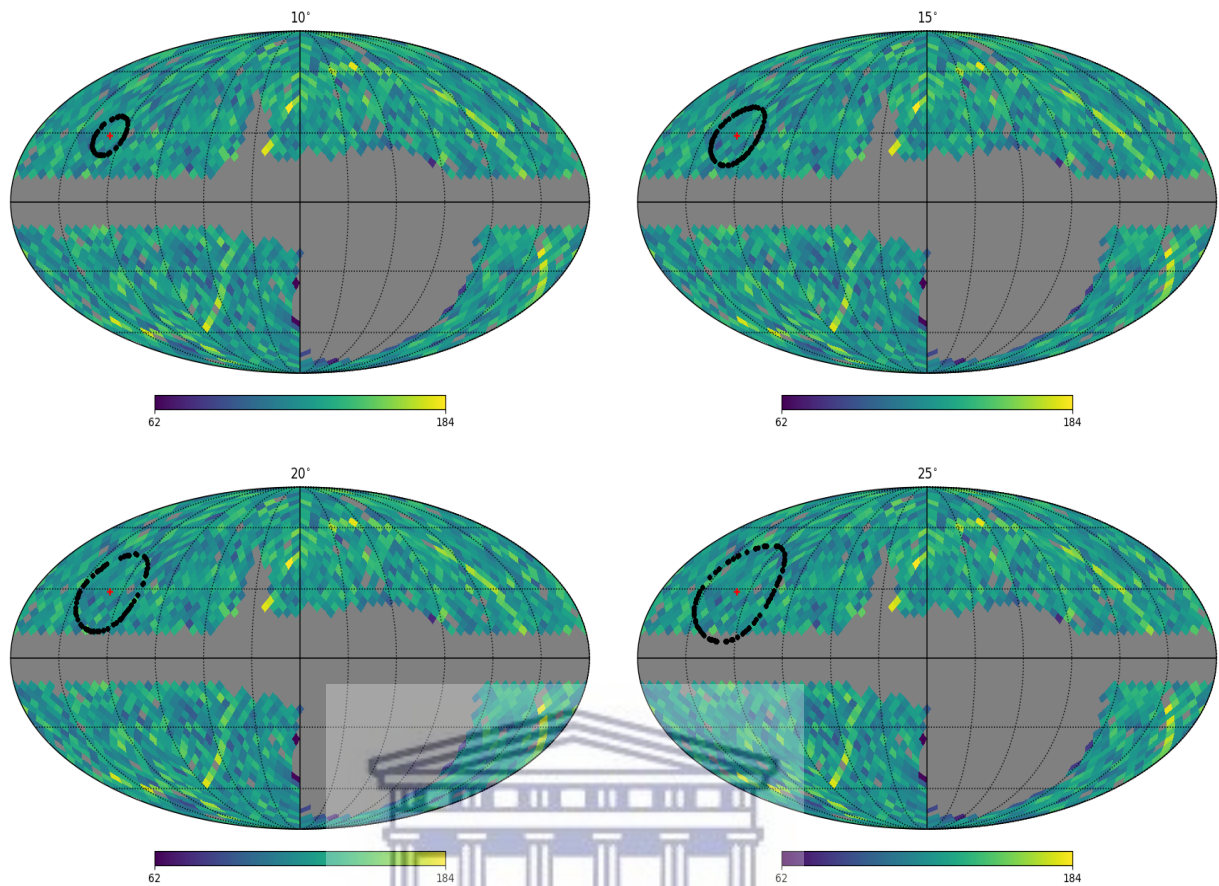
FIGURE 3.10: The maps above show four patches of different radii, the black curve denotes the boundary of the patch and the red cross is its center.

3.3.3 Mock data generation

We shall also run our local variance estimator in equation 3.12, through 1 000 mock realizations of galaxy number counts which are produced using the FLASK code (Xavier, Abdalla, and Joachimi, 2016). Camb sources provide the assumed angular power spectrum (Challinor and Lewis, 2011). For detailed method on how to generate the mock realizations together with the FLASK calibration corrections, see Bengaly, Maartens, and Santos, 2018.

The simulated mocks carry all the information we know about Cosmology, for instance, the number of sources per redshift bin n(z), as well as the clustering bias b(z). The n(z) and b(z) quantities will not be discussed here, the reader is encouraged to read Bengaly, Maartens, and Santos, 2018 for a thorough explanation on them.

3.4   Tests for Statistical Isotropy

We use several test statistics to give us a measure of the differences in galaxy number counts for patches on the sky. Then we compare the obtained galaxy number counts variance. Our testing procedure is as follows:

1. Make 3072 patches of a given radius r, see figure 3.10 for an idea of the sizes of the patches, centred on the pixels of a HEALPix $N_{side} = 16$ grid.

2. For each patch and sky map, we compute the galaxy number count variance

$$\sigma^2 = \frac{\sum(x - \mu_{pix})^2}{N} \tag{3.9}$$

with number counts in a pixel x, mean number count in the pixel μ_{pix} and the mean number count per patch $\bar{n}$, which is summed up over all pixels in a patch

$$\bar{n} = \frac{\sum \mu_{pix}}{N_{pix}}. \tag{3.10}$$

Patches for which less than 90% of the pixels don't have a mask are ignored completely. Similarly, for the other two cases, i.e less than 70% and 50% of the area unmasked.

3. Compute the coefficient of variation, which is a statistical measure of the dispersion of data around the mean, calculated from the mean and standard deviation in galaxy number counts as follows

$$CV = \frac{\sqrt{\sigma^2}}{\bar{n}} \tag{3.11}$$

The coefficient of variation is useful for comparing the degree of variation from one data series to another, in our case from one patch to the next. This particular statistic is a widely used standardized measure for the spread of measurements for a sample, it allows direct comparisons of variations in samples with different means (Marwick and Krishnamoorthy, 2018).

4. For each patch, we compare the coefficient of variation by performing the test statistics in the following section.

5. To establish the expected mean and variance of each patch, we compute the same galaxy number count variance and mean from 1 000 simulated mock data, which have the same assumptions as those on Cosmology and the observational data.

3.4.1 Local Variance Estimator

For this particular statistic, we compute the coefficient of variation in each patch for the real data (NVSS), then compute the coefficient of variation for the equivalent patches in 1 000 simulated mock maps so that we obtain the coefficient of variation for 1 000 simulated mock maps. This is the expected value for the coefficient of variation in galaxy number counts. We take an average of the 1 000 maps because there are some variations in the mocks as well. This average coefficient of variation is then compared to the coefficient of variation of the observed data. We take the difference in galaxy number count between the data sample and the simulated mock data.

We use this Local Variance Estimator (LV) to test the statistical isotropy of the sky distribution of radio sources. For similar methods, see (Akrami et al., 2014). The local variance statistic is defined as

$$\xi = \frac{(\sigma_p/\bar{n}_p)_{\text{data}} - (\bar{\sigma}_p/\bar{n}_p)_{\text{mock}}}{(\bar{\sigma}_p/\bar{n}_p)_{\text{mock}}} \tag{3.12}$$

The ξ is computed from all the pixels, falling inside a circular patch with mean number count $\bar{n}_p$. The term $(\sigma_p/\bar{n}_p)_{\text{data}}$ is the number count coefficient of variation for a patch from the data and $(\bar{\sigma}_p/\bar{n}_p)_{\text{mock}}$ is the average number count coefficient of variation for the 1 000 simulated mock maps for a patch, and $\bar{n}_p, i$ is the mean number counts in a pixel, for a patch p from the data with number of pixels per patch N_{pix};

$$\bar{n}_p = \frac{1}{N_{pix}} \sum_{i=1}^{N_{pix}} n_{p,i} \tag{3.13}$$

3.4.2   Two Sample test

This is a statistical test applied to sets of data to evaluate how likely it is that any observed difference between the sets arose by chance. In our case, the coefficient of variation in galaxy number count between two patches is compared. We used 3072 patches. The χ is defined as

$$\chi = \frac{(\sigma_p/\bar{n}_p)_{\text{data}} - (\sigma_{p+1}/\bar{n}_{p+1})_{\text{data}}}{(\sigma_p/\bar{n}_p)_{\text{data}} + (\sigma_{p+1}/\bar{n}_{p+1})_{\text{data}}} \tag{3.14}$$

The test is performed on pairs of patches, p represent first patch then p+1 will represent the adjacent patch. A perfect uniform distribution corresponds to $\chi = 0$, thus, the smaller the value of χ, the more uniform the distribution is, this is according to Menezes, Pigozzo, and Carneiro, 2017. The same procedure will also be applied to the simulated mocks.

3.5   Analysis

For this section we discuss the analysis of the estimators in section 3.4, also discuss how they show consistency with statistical isotropy for the data. From the method described above, we have estimated the local variance for 3072 patches. Figure 3.11 shows the number of accepted patches corresponding to all the radii for our case. The positions of the patches were chosen using the HEALPix software with $N_{side} = 16$, in order to cover the entire sphere with as many patches as possible.

For every patch, we consider only those pixels whose centres fall within the patch radius. We establish criteria for accepting the patch, by this criteria, we accept all the patches where more than 90%, 70% and 50% of the area remains after applying the mask to the patch in question. Figure 3.11 shows the remaining patches which were used in our work.
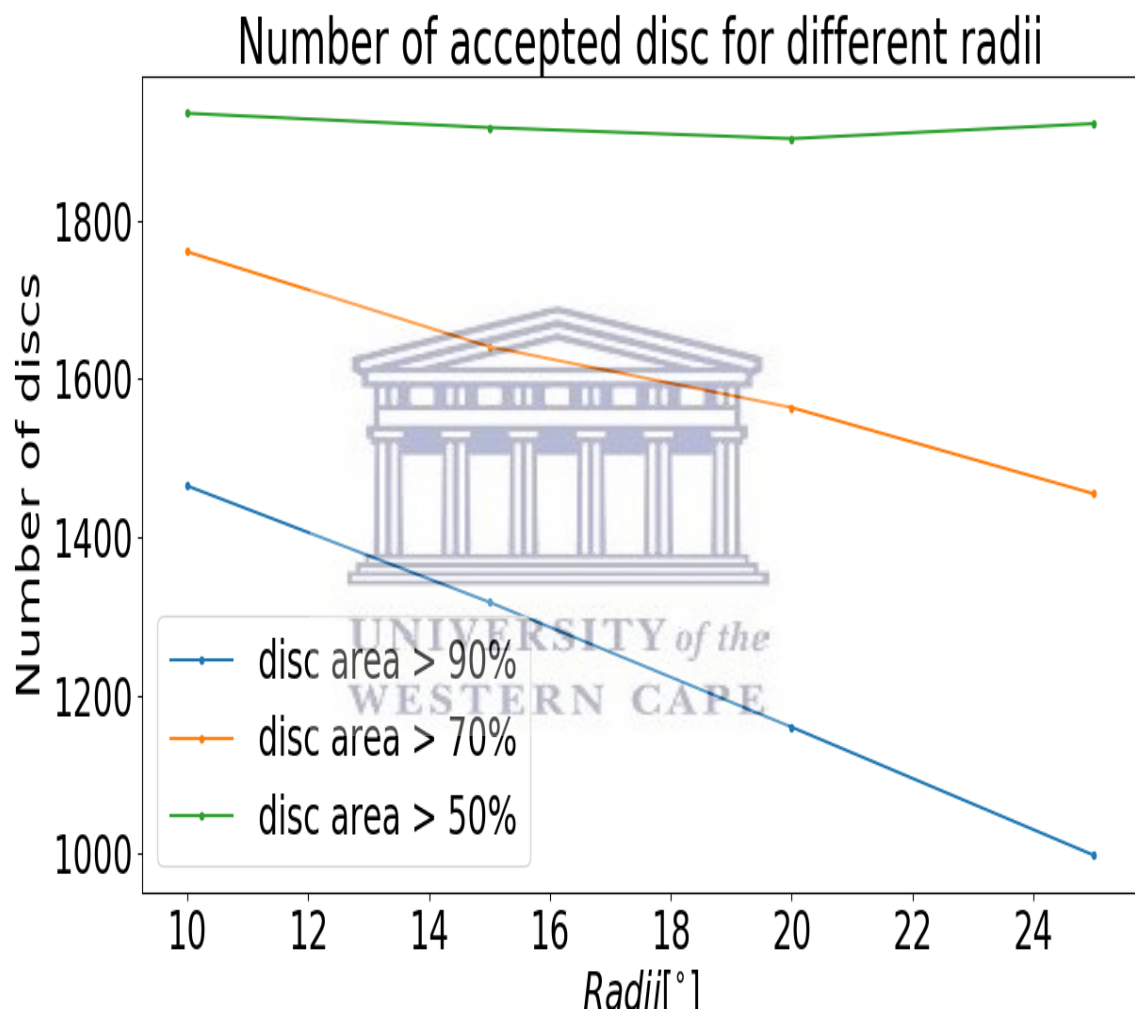
FIGURE 3.11: The total number of discs that were accepted for all 4
radii in this work, using the criteria explained above.

Chapter 4

Results and Discussion

In this chapter we show the results that we obtained from the two estimators described in chapter 3, with results from both our estimators as describe in section 3.4. We also discuss possible reasons for the results and also ways in which they can be improved.

4.1 Local Variance Estimator results

We use equation 3.12 to calculate the local variance values for 3072 pixels, using our rejection criteria to avoid the masked area so that we are left with the number of patches shown in figure 3.11.

The maps in figure 4.1 corresponds to the local variance estimator results for different patch sizes, and by carefully comparing them tells us as to how statistically isotropic our data is. From figure 4.1 the maps become smooth as you move from left to right, that is as the rejection criteria become less rigorous. As the number of patches increases the maps become more uniform. There are also less rejected patches when the rejection criteria are less strict (i.e 50 %).

As an equation, 3.12 shows the difference for the coefficient of variation between the data and mocks. The maps show that $\Delta \xi$ decreases from left to right, which is defined as

$$\Delta \xi = \xi_{max} - \xi_{min} \tag{4.1}$$
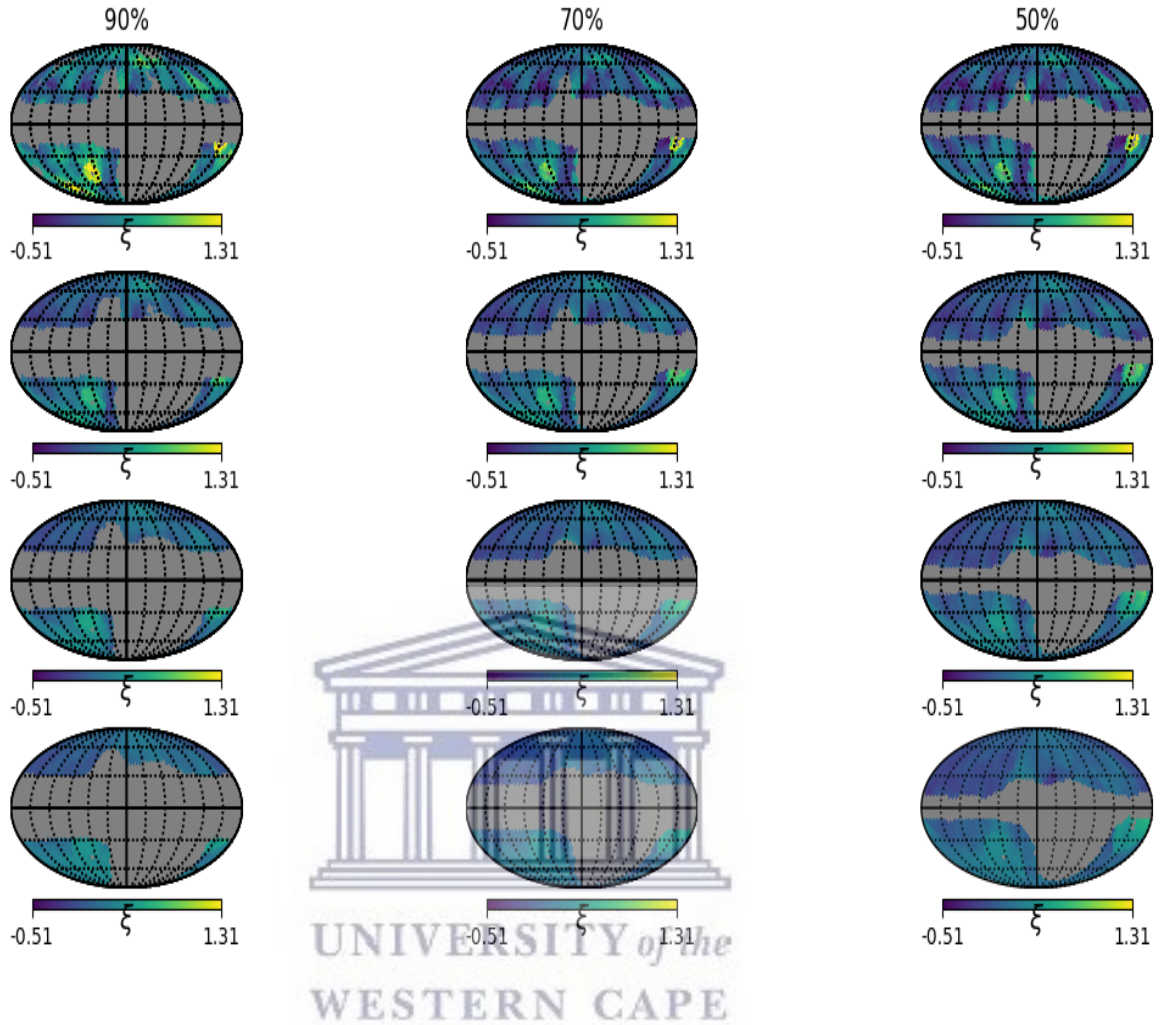
FIGURE 4.1: The local variance maps for the NVSS data with N_{side} = 16 HEALPix grid. The rows represent the patch radii, from top 10° (first row), 15° (second row), 20° (third row) & 25° (forth row). We make a comparison of the three criteria in which to accept the patch, 90 % (left), 70 % (middle) & 50 % (right) of the disc unmasked. All rejected pixels are masked.

The equation measures the spread in the variance of galaxy number counts for the local variance estimator, this will give us an idea of how well our mean is representative of the data. We expect the value of $\Delta\xi$ to decrease as we increase the radii and relax the rejection. This is because a small spread in the data indicates that there are small differences between the individual number count variances in each patch.

This implies that the data and the simulations are becoming statistically similar as we become less rigorous in our rejection. This is due to the decrease in the spread of the data, however, for the 25 ° case, we have the opposite. This may be due to

$PatchRadius(Degrees)$	Criteria of 90%	Criteria of 70%	Criteria of 50%
10	1.6	1.2	1.0
15	0.88	0.71	0.61
20	0.88	0.62	0.49
25	0.50	1.64	1.20

TABLE 4.1: Table showing the spread values for the number count variance for the local variance maps in figure 4.1. It shows the value of $\Delta\xi$ for the 4 radii and rejection criteria respectively.

some error in our calculation or coding for producing the maps. Now if we focus on the size of the patch radius, the smallest patch radius corresponds to the largest value of $\Delta\xi$ (except for the two cases), the associated maps, figure 4.1 top row vary more than the rest of the other maps. Thus the patches with a radius of $10°$ are less statistically similar to the other patches. The second row from top represents patches with a radius of $15°$, these maps are more uniform than the $10°$ maps, they differ less as compared to the smaller radius discs.

From top to bottom we notice also a decrease in the number of accepted patches, so as the patch radii become bigger, more patches are rejected as compared to the previous radius. This is also evident on figure 3.11. Secondly, as the radii increase, the maps become more uniform with the maps corresponding to $25°$ discs being the most uniform of the three radii.

Figure 4.2 shows the local variance estimator results for one mock realization, with the same selected pixel centres as in figure 4.1. We perform this test to check the performance of our estimator. The local variance maps become uniform as the radii of the patch increases. Similar to the local variance results in figure 4.1. Also, notice that $\Delta\xi$ becomes small as the rejection becomes less rigorous.

These local variance maps in figure 4.2 show that our estimator is performing well. We might just need to consider more simulated maps to compare with our data. Considering only one realization might not be getting all the information from the maps.

We can conclude that there are smaller local variance fluctuations across the sky as
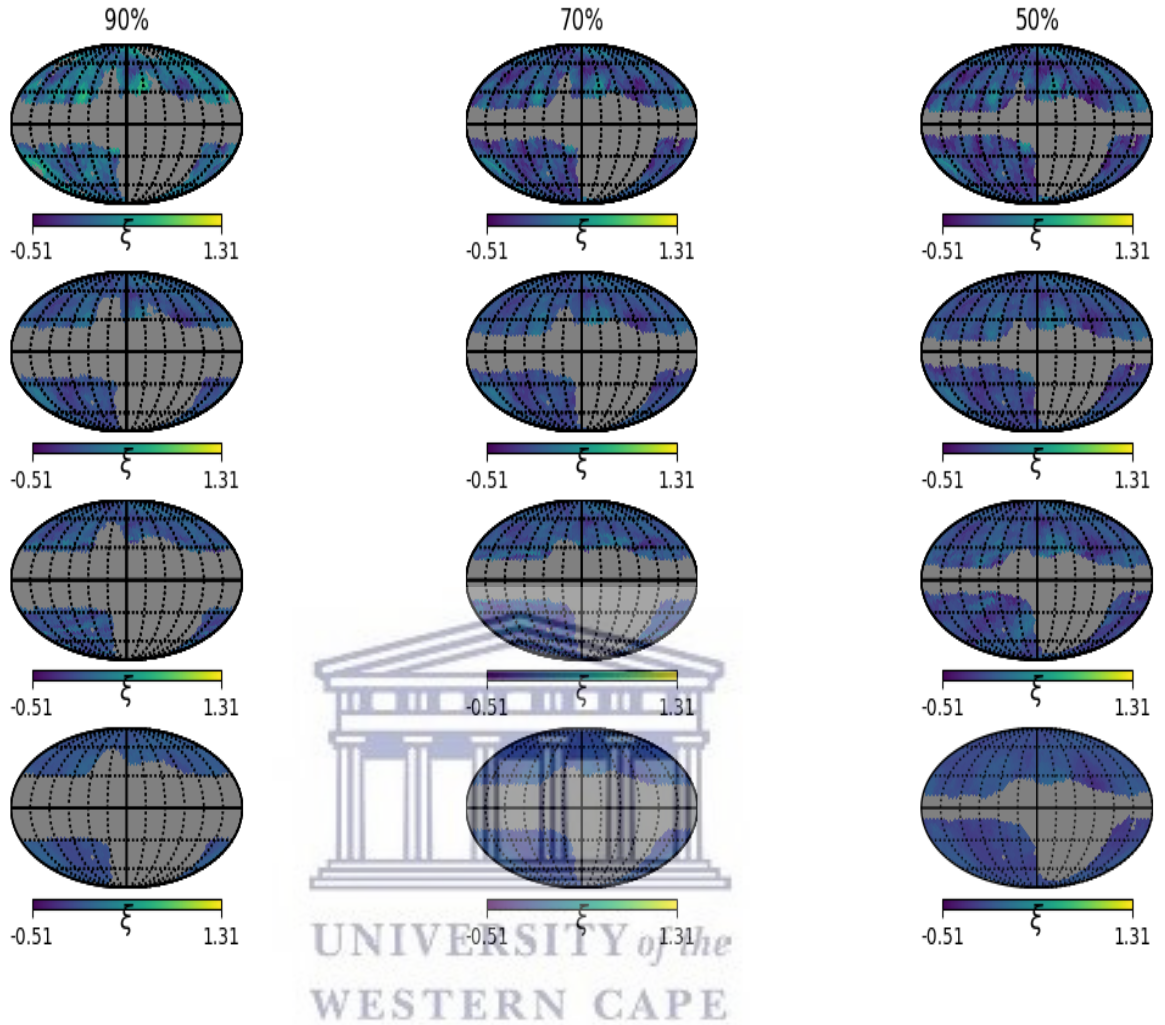
FIGURE 4.2: The local variance maps for one simulated mock realization map with N_{side} = 16 HEALPix grid. The rows represent the patch radii, from top 10° (first row), 15° (second row), 20° (third row) & 25° (forth row). We make a comparison of the three criteria in which to accept the patch, 90 % (left), 70 % (middle) & 50 % (right) of the patch unmasked. All rejected pixels are masked.

we become less rigorous in our rejection, and also as the patch radii increases. The less rigorous the rejection, the more statistically isotropic the data becomes. Increasing the radii means that more pixels are analyzed, this could also introduce more deviations as the masked area also increases.

4.2   Two sample test results

The statistic for the two sample test is done to compare the number count variance of the data with the expected number count variance from the simulated mocks. This is done by comparing patch by patch for the data independently then also independently for the mocks. Then at the end, we can compare variations from all the patches for the data to the averaged variance from 1 000 simulated mocks. If two patches are statistically similar, their χ value will be close to zero. For two perfectly uniform patches $\chi = 0$. We compare the patches in pairs, from patch 1, ... ,3072.

Since we only compare the patches with each other, we also have to run our estimator on the simulated data. Then we will compare the patches from the data to those from the simulated mock maps. To get the average histogram for the mocks, we take an average of the galaxy number counts in each patch for the 1 000 maps, then average that to get only one map with 3072 averaged patches. Then compare those averaged patches against each other to quantify their similarity and differences.
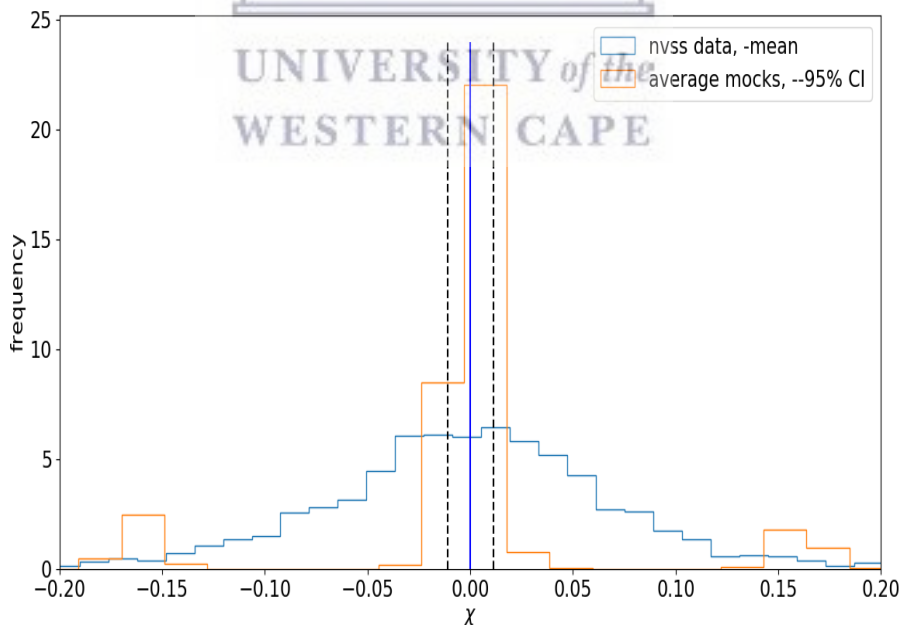

FIGURE 4.3: Two sample test result for r = 10° and at least 90 % of each pixel unmasked, for the NVSS data and the two sample chi-squared average of 1 000 simulated mock maps. The solid blue line represents the mean of the data, the dashed lines represent the 95 % confidence interval for the mocks.

The histogram in figure 4.3 shows both the data and the mocks having a long tail, however, the mocks show a higher frequency around the expected value of $\chi \approx 0$. This is just a representation of one radius and one criterion, the most rigorous rejection criterion with the smallest radius. We also perform the test on all the four radii and the three criteria similar to the one we did for the local variance maps. Also, we show the 95 % confidence interval for the mocks to see if the mean of the data will fall within this confidence interval. Then we can say with confidence if the data agrees well with the simulated mocks.
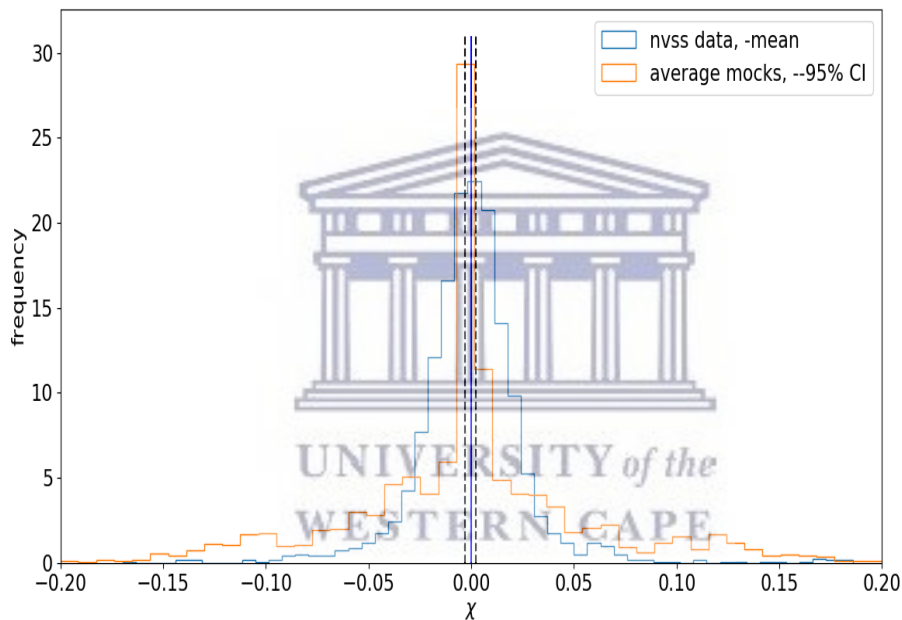
FIGURE 4.4: Two sample test result for r = 25° and at least 50 % of each pixel unmasked. For the NVSS data and the average of 1 000 simulated mock maps. The solid blue line represents the mean of the data, the dashed lines represent the 95 % confidence interval for the mocks.

Now we plot our less strict criteria with the largest radius,i.e Fig. 4.4 so we can compare with the most strict and smallest radius plot in figure 4.3. Notice that the histogram for the data has a shorter tail now. The frequency has increased towards the expected χ value. The mean for the data is still within the 95 % confidence interval of the mocks. The data is more statistically isotropic in this plot as compared to the earlier plot in figure 4.3. The average mocks still have a long tail, with a high frequency close to the expected χ value.
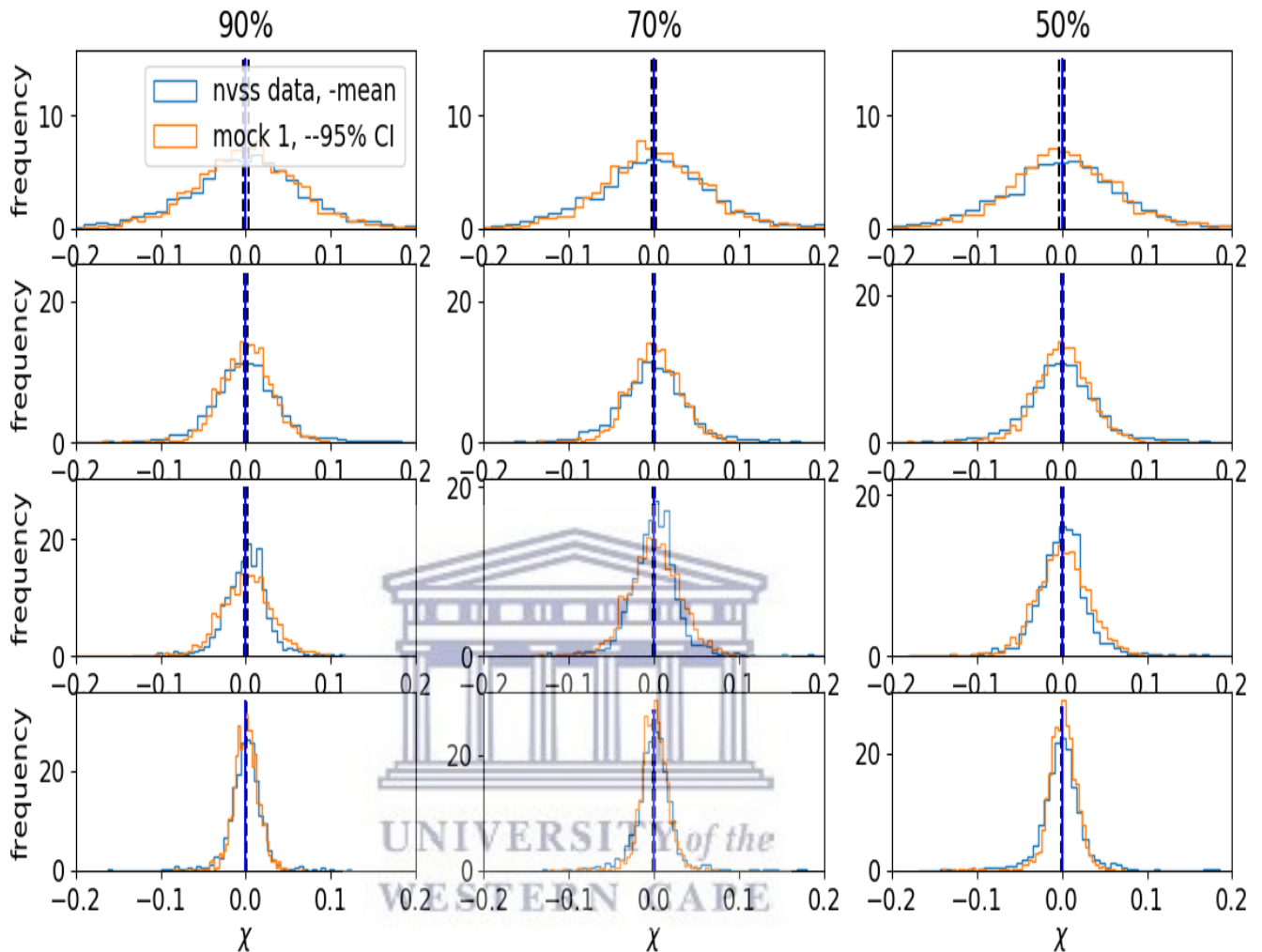
FIGURE 4.5: The two sample test results for the NVSS data (blue) versus one simulated mock realization (orange). The rows represent the patch radii, from top 10° (first row), 15° (second row), 20° (third row) & 25° (forth row). We make a comparison of the three criteria in which to accept the patch, 90 % (left), 70 % (middle) & 50 % (right) of the patch unmasked. Also plotted with the 95 % confidence interval for the mock, and the mean of the data.

Figure 4.5 makes a comparison between the data and one mock realization. We check how the variance in the data between the patches compares with the variance between the patches in one of the mock realizations. The agreement between simulated mocks and data gets better as we become less strict with the rejection, this could be due to the fact that we are analyzing more patches compared to the other two cases (90%, 70%). Also as the radius of the patch increase, there is a higher frequency of values around $\chi = 0$. Analyzing bigger patches results in more statistically isotropic results.

The results in figure 4.5 is only a comparison between the data and one mock realization. It is not enough for us to conclude whether the data is statistically isotropic or not. We now have to compare the data with an average of 1 000 simulated mock maps, only then we can conclude if the data is statistically isotropic or not.
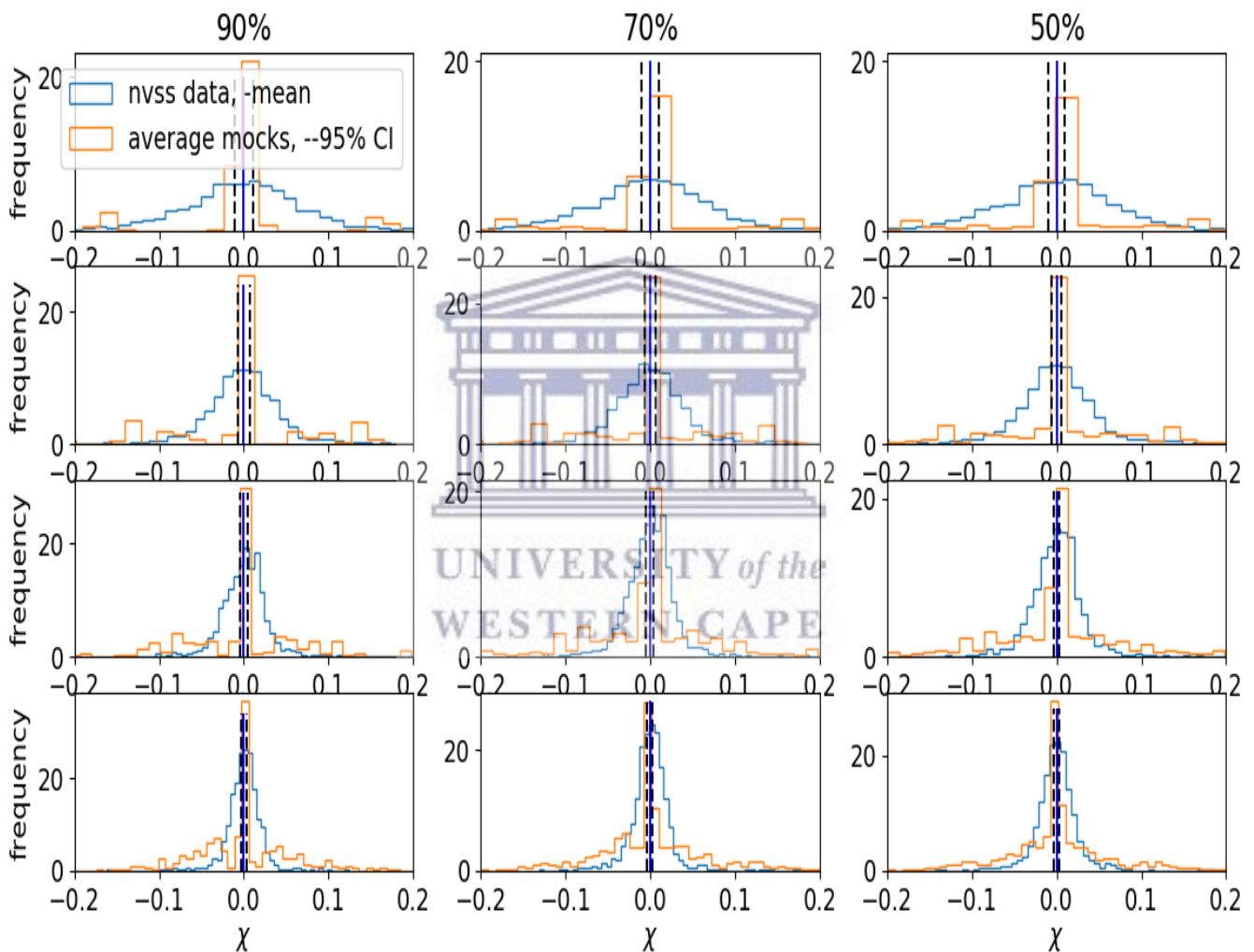
FIGURE 4.6: The two sample test results for the NVSS data (blue) and the average of 1000 simulated mock (orange). The solid blue line represents the mean of the data, the dashed lines represent the 95 % confidence interval for the mocks. The rows represent the patch radii, from top 10° (first row), 15° (second row), 20° (third row) & 25° (forth row). We make a comparison of the three criteria in which to accept the patch, 90 % (left), 70 % (middle) & 50 % (right) of the disc unmasked.

The simulated mocks are generated using the Lambda-CDM model as the fiducial power spectrum, with the same sky coverage as the data. Thus, if the distribution

between the data and simulated mocks are similar, we can infer statistical isotropy for the data. Figure 4.6 shows the comparison between data and mocks for a radius of 10° the data is more spread out than the mocks, this may have been due to few patches accepted because of the radius size. The test doesn't perform well with 10° patches, however, the mean for the data is within the 95 % confidence interval of the averaged mocks.

The spread of the data decreases as the patch radii increases. That is, as the number of patches accepted decreases. Increasing the patch radius means that the patch will cover a larger fraction of the sky. We expect the χ statistic to perform better as the radii increases since we're analyzing larger fractions of the sky. For all our cases the mean of the data falls well within the 95 % confidence interval of the mocks. The histograms are symmetric for the data.

Chapter 5

Conclusion

We started with the Cosmological Principle, which assumes that the Universe is isotropic and homogeneous. We test this assumption on galaxy number counts across the sky to verify if our data is statistically isotropic. The analysis is done by performing two test statistics, the local variance estimator and two-sample test. For both these tests, we compare the galaxy number count variance for different patches on the sky. We perform them on different radii, to check if our tests are biased towards a certain radii size. We also establish criteria for accepting the patches on the sky, this is done to check if the tests are sensitive to the number of patches we accept or not.

To test the statistical significance of the results, we compare the results obtained from the data with those that are obtained from simulated mocks. We generate a total of 1 000 mocks for the same sky coverage as the data, using the fiducial Lambda-CDM power spectrum. These mocks are expected to be statistically isotropic, thus a good way to compare with our data to detect any departures from statistical isotropy. The simulated mocks have some fluctuations in number counts so we take an average number count variance of 1 000 mocks to compare it with the data.

We start with the local variance estimator ξ. For this particular test, we compute the number count variance in each patch for the data, then compare it with the expected number count variance from the 1 000 mocks. With the expectation that for our data to be statistically isotropic, we should have uniform local variance maps.

Figure 4.1 show the local variance maps that were obtained from using equation 3.12. For the patch size of $10°$, the local variance maps have a smaller spread as you move from the more strict criteria for accepting patches (90% unmasked pixels in the patch) to the less strict criteria (50% unmasked pixels in the patch), this can be seen in table 4.1 with an exception for the last row. This means that for this particular radius, the maps become more uniform as our estimator becomes less strict with rejecting patches. For this case, we can say that the less strict criteria is the one that produces statistically isotropic maps.

This feature is also noticeable in the other radii as well. The local variance maps become statistically isotropic as the rejection criteria changes (as we accept more patches). This leads us to conclude that, based on the local variance maps in figure 4.1. The data is statistically isotropic when we accept many patches that are larger in the sky.

Next, we use the two sample test. This test statistic helps us compare the individual patches. We use it to compare the number count variance between patches for the data. Then we use it to also compare the average number count variance for the mocks. Both the test for the data and the mocks are done independently of each other. For statistically isotropic patches we expect to get a χ value that is as close as possible to zero. Thus, a uniform distribution around zero would imply statistical isotropy between the patches.

Figure 4.3 shows a comparison between the data and the average mocks for one case of r = $10°$ and 90 % of the pixels unmasked. This plot shows a large spread for the data compared to the mocks. The mean of the data is well within the 95 % confidence interval of the mocks. So we can say that there is some agreement between the data and the mocks. Comparing the plot to figure 4.4, we notice that in figure 4.4 the spread of the data is less. The histogram is more symmetric and the mean of the data is within the 95 % confidence interval of the mocks. The data is more statistically isotropic in figure 4.4, which corresponds to a large radius and less strict acceptance of patches.

Comparing only two cases might not be enough, so in figure 4.5 we do a comparison between the data and only one mock realization. The plot shows symmetric histograms with an equal spread. The data and this particular mock are in good agreement, with better performance noticeable for larger radii.

We perform a full comparison between data and the average of 1 000 simulated mocks, see figure 4.6. We notice the histograms are not as symmetric as in figure 4.5, however, there is still a reasonable agreement between the mocks and the data. The mean is well within the 95 % confidence interval for all radii. The histograms are also less spread for larger patches.

We can conclude that from the results we have obtained from our two estimators, the test statics perform better as the radii become larger. Thus the data becomes statistically isotropic for larger radii and when we do the analysis over a larger sample of patches. We do not notice any huge discrepancies between the data and the mocks. We did not find evidence supporting larger than expected radio count anisotropy at angular scales smaller than the dipole. As the dipole is higher than expected, but at smaller scales, it weakens its statistical significance. This result indicates that the large dipole may happen due to uncleaned systematics affecting large angles, rather than a true cosmological signal. We can conclude that the statistical properties of our sample are in reasonable agreement with the standard cosmological model.

More tests still need to be done for us to conclusively measure the violations of statistical isotropy. Other radio continuum data sets like the TIFR GMRT Sky Survey (TGSS) conducted by Giant Metrewave Radio Telescope (GMRT) at a low frequency of 150 MHz should also be used to test statistical isotropy (Bengaly, Maartens, and Santos, 2018). Also, with observations from surveys like Evolutionary Map of the Universe (EMU) on the Australian Square Kilometre Array Pathfinder (ASKAP) (Johnston et al., 2007). We can further test statistical isotropy as well as homogeneity with much better precision. The upcoming Square Kilometer Array (SKA) survey will also be able to test statistical isotropy with better precision (. Bengaly et al., 2018).

Chapter 6

Bibliography

Ade, P.A.R et al. (2016). "Planck 2015 results-xiii. cosmological parameters". Astronomy Astrophysics 594, A13.

Aghanim et al. (2018). "Planck 2018 results. VI. Cosmological parameters". arXiv: 1807.06209 [astro-ph.CO].

Akrami, Y. et al. (2018). "Planck 2018 results. I. Overview and the cosmological legacy of Planck".

Akrami, Y. et al. (2014). "Power asymmetry in WMAP and Planck temperature sky maps as measured by a local variance estimator". Astrophys. J. 784, L42.

Ballardini, M, et al (2016), "Probing primordial features with future galaxy surveys". JCAP. Volume 1610. arXiv:1606.03747 [astro-ph.CO]

Bengaly, C.A.P, Maartens, R and Santos, G (2018). "Probing the Cosmological Principle in the counts of radio galaxies at different frequencies". JCAP. arXiv:1710.08804 [astro-ph.CO].

Bengaly, . et al. (2018). "Testing the standard model of cosmology with the SKA: the cosmic radio dipole". arXiv:1810.04960 [astro-ph.CO].

Burke, B.F and Francis G.S (2009). An introduction to radio astronomy. Cambridge University Press.

Challinor, A and Antony L (2011). "The linear power spectrum of observed source number counts". Phys. Rev. D84. 043516.

Clarkson, C (2012). "Establishing homogeneity of the universe in the shadow of dark energy". Comptes Rendus Physique 13, 682. arXiv:1204.5505 [astro-ph.CO].

Clarkson, C and Maartens R (2010). "Inhomogeneity and the foundations of concordance cosmology". Class. Quant. Grav. 27. arXiv: 1005.2165 [astro-ph.CO].

Clifton, T et al. (2012). "Modified gravity and cosmology". Physics reports 513.1-3, 1–189.

Condon, J.J. et al. (1998). "The NRAO VLA Sky survey". Astron. J. 115.

Deruelle, N and Jean-Philippe U (2018). Relativity in Modern Physics. Oxford Graduate Texts. Oxford University Press.

Dodelson, Scott (2003). Modern Cosmology. Amsterdam: Academic Press.

Ellis, G.F.R. (2006). "Issues in the philosophy of cosmology". Philosophy of physics.Ed. by Jeremy Butterfield and John Earman. arXiv:astroph/0602280 [astro-ph].

Forshaw, J.R. and Smith A.G (2009). Dynamics and relativity.

Gibelyou, C and Dragan H (2012). "Dipoles in the Sky". Mon. Not. Roy. Astron. Soc. 427.

Gonçalves, R.S. et al. (2018). "Cosmic homogeneity: a spectroscopic and model- independent measurement". Mon. Not. Roy. Astron. Soc. 475.1. arXiv:1710.02496 [astro-ph.CO].

Gorski, K.M et al. (2005). "HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere". The Astrophysical Journal 622.2, p. 759.

Haslam, C.G.T et al. (1982). "A 408 MHz all-sky continuum survey. 2-The atlas of contour maps". Astronomy and Astrophysics Supplement Series 47.

Johnston, S. et al. (2007). "Science With The Australian Square Kilometre Array Pathfinder". PoSMRU. [Publ. Astron. Soc. Austral.24,174(2007)], p. 006. arXiv:0711.2103[astro-ph].

Kogut, A. et al. (1993). "Dipole anisotropy in the COBE DMR first year sky maps". Astrophys. J. 419, arXiv:astro-ph/9312056 [astro-ph].

Marwick, B and Kalimuthu K (2018). Tests for the Equality of Coefficients of Variation from Multiple Groups. R package version 0.1.3.

Menezes, R. S., C. Pigozzo, and S. Carneiro (2017). "Testing cosmic isotropy with galaxies position angles distribution". arXiv:1705.02323 [astro-ph.CO].

Milgrom, M (1983). "A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis". The Astrophysical Journal 270.

Murph, D. (2011). "Visualized: a decade of mapping the universe, and all we got was this bloated JPEG".

Ryden, Barbara (2016). Introduction to cosmology. Cambridge University Press.

Schwarz, D.J. et al. (2016). "CMB Anomalies after Planck". Class. Quant. Grav. 33.18, arXiv:1510.07929 [astro-ph.CO].

Scolnic, D.M et al. (2018). "The complete light-curve sample of spectroscopically con- firmed SNe Ia from Pan-STARRS1 and cosmological constraints from the combined pantheon sample". The Astrophysical Journal 859.2.

Serway, R.A and J.W Jewett (2018). Physics for scientists and engineers with modern physics. Cengage learning.

Soumagnac, M.T. et al. (2018). "Large scale distribution of mass versus light from Baryon Acoustic Oscillations: Measurement in the final SDSS-III BOSS Data Release 12". arXiv:1802.10368 [astro-ph.CO].

Xavier, H.S., F.B. Abdalla, and B Joachimi (2016). "Improving lognormal models for cosmological fields". Mon. Not. Roy. Astron. Soc. 459.4.