Detecting Baryon Acoustic Oscillations with HI Intensity Mapping using MeerKAT

Brandon Engelbrecht

Supervisor: Prof. M Santos
Dr. J Fonseca (Co.)

Department of Physics and Astronomy
University of the Western Cape

This dissertation is submitted for the degree of
Master of Science

August 2019

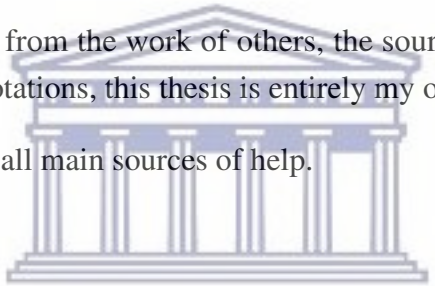
I would like to dedicate this thesis to my Mother...

Declaration

I, *Brandon Engelbrecht*, declare that this thesis titled, "Detecting Baryonic Acoustic Oscillations wit HI Intenisty Mapping for MeerKAT" and the work presented in it are my own. I confirm that:

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Brandon Engelbrecht
August 2019



Acknowledgements

Firstly I would like to thank my supervisor Prof. Mario Santos, for giving me the opportunity to work under him as well as being part of the Centre for Radio Cosmology (CRC). He has given his time and effort to aide me through the MSc and I am grateful for that.

To my co-supervisor Dr. José Fonseca, whose help has been enormous in the completion of this MSc. Always ready to assist when available especially in making the work understandable.

Next I would like to thank the simulation team: Amadeus Wild, Dr. Marta Spinelli, Siyambonga Matshawule and Dr. Sean February; without them I would be drowning in a computational mess. Thanks to them, I was able to progress with the simulations smoothly.

To Eliab Malefahlo and the MSc office, for making sure the work-space was always lively and ready to assist whether it be personal or academic issues they were always ready to give advice.

I would like to thank the CRC, the South African Radio Astronomy Organization (SARAO) and National Research Foundation (NRF) for the funding me throughout this project. A thanks to the Centre for High Performance Computing (CHPC) for the use of the Lengau cluster as well as the technical support.

To Kimeel Sooknunan and Jake Gordin, who have been with me through my academic career since Honours and became the best friends I could have asked for. They have always given help with computational and theoretical knowledge respectively. However their biggest support has been emotional and mental and for that I am truly grateful for their friendship.

Lastly I would like to give thanks to my family whom I have placed on hold in order to complete this project. They have always stood behind me and giving me encouraging support and giving me a productive working environment.

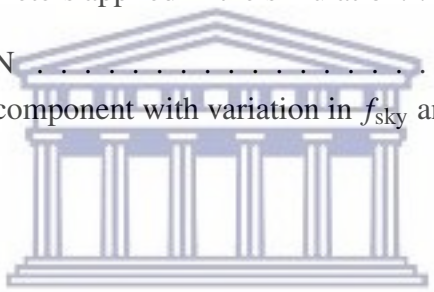
Abstract

Future radio surveys as the Square Kilometer Array (SKA) and its precursor, the "Meer" Karoo Array Telescope (MeerKAT), will map the Neutral Hydrogen (HI) in large areas of the sky using the intensity mapping (IM). HI IM is currently one of the most promising ways of accessing the Large-Scale Structure of the Universe. The distribution of matter in the Universe not only encodes its composition but also how it evolves and its initial conditions. An effect on the matter distribution that will be detected by the SKA on the post re-ionization Universe are the Baryonic Acoustic Oscillations (BAO). While it has been shown that in single dish mode the SKA can measure the BAO peak in the radial 21cm power spectrum at low redshifts, this possibility has not yet been studied in detail for the MeerKAT. In this thesis we construct a set of full sky simulations to test how well MeerKAT will be able to extract the BAO wiggles along the line of sight. These simulations are done for the frequencies corresponding to MeerKAT L-band. The maps combine the cosmological HI signal, systematic noise, cosmological foregrounds and the instrumental telescope beam. A model-independent estimator is used to extract the BAO wiggles by subtracting a smooth polynomial component from the 21cm radial power spectrum. We test with simulations if this estimator is biased and the signal to noise of the extraction. We conclude that we are able to remove contaminants and recover the cosmological HI signal while not risking the recovery of the BAO signal. We investigate the effects of varying the sky area and the observational hours on the signal to noise ratio for the BAO wiggles. We found that for a HI IM experiment using MeerKAT, the optimal sky area to detect the BAO along the line of sight is 50% of the sky. With a signal-to-noise ratio of 3.37. This can be achieved with 2000 hours of exposure time.

Table of contents

List of figures

List of tables



Chapter 1

Introduction

1.1 Isotropic Universe

This section introduces the ΛCDM cosmological model which is applied throughout the thesis and is based on information found in "Modern Cosmology" by S.Dodelson [1] and the lecture notes "Cosmology Part III Mathematical Tripos" of D. Bauman [2].

The Friedmann-Robertson-Walker Metric

The *Cosmological Principle* states that on very large scales (scales $>$ 100 Mpc), the Universe is homogeneous and isotropic. Homogeneity is the notion that the Universe is the same regardless of position and isotropy is the notion that the Universe is the same regardless of direction.

With the idea of the Cosmological Principle, we need to establish mathematical and physical models of the Universe in which this principle holds. Assume the Universe can be modelled as a continuous fluid and assign spatial coordinates to each fluid element x^i $(i = 1,2,3)$. Therefore any point in a four-dimensional space-time can be labelled by the x^μ $(\mu = 0,1,2,3)$ coordinates, corresponding to the fluid element which is passing through the point and a time parameter (x^0). The space-time metric that describes our Universe in which the Cosmological Principle holds can be shown by

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-Kr^2} + r^2(sin^2\theta d\theta^2 + d\phi^2)\right], \tag{1.1}$$

with r, θ and ϕ being the spherical co-moving coordinates. t is the *cosmological proper time* and a is the *cosmic scale factor* which describes the general expansion of the Universe as

a function of time. c is the *speed of light* and K the *curvature parameter*. We assume a spatially flat Universe ($K = 0$) in this thesis. The metric tensor measures the length between two points in any space. The *cosmological redshift* can be described as

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}}, \tag{1.2}$$

where λ_{emit} is the emitted wavelength from the source and λ_{obs} is the observed wavelength. The standard convention is to set $a = 1$ at time today and relate redshift with the cosmic scale factor as:

$$1 + z = \frac{1}{a} \tag{1.3}$$

The Friedmann Equations

We assume that the evolution of our Universe can be described by Einstein's equations of General Relativity

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.4}$$

where $G_{\mu\nu}$ is the Einstein tensor, G is Newton's constant and $T_{\mu\nu}$ is the stress-energy tensor. These tensors describe the geometry and the energy-momentum content of the Universe respectively. The Einstein tensor is given by

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\mathscr{R} g_{\mu\nu}, \tag{1.5}$$

where $R_{\mu\nu}$ is the *Ricci tensor* (Eq. A.1) and $\mathscr{R}$ is the *Ricci scalar* (Eq. A.2). These equations are defined in Appendix A as well as the metric connection (Eq. A.3). The metric is given by $g_{\mu\nu}$.

If we consider a perfect fluid in which we ignore the effects of shear stress, viscosity and heat conduction. The stress-energy tensor is defined as

$$T_\mu{}^\nu = T_{\mu\alpha} g^{\alpha\nu} = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}, \tag{1.6}$$

with ρ being the density in Universe and P being the pressure.

Inserting the Friedmann-Robertson-Walker (FRW) metric (Eq. 1.1) into Einstein's field equations (Eq. 1.4) we can obtain the Friedmann equations:

- Constraint equation:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho, \tag{1.7}$$

- Evolution equation:

$$\frac{\ddot{a}}{a} + \frac{1}{2}\left(\frac{\dot{a}}{a}\right) = -4\pi GP, \tag{1.8}$$

combining Eq. 1.7 and Eq. 1.8 we obtain the Acceleration equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P). \tag{1.9}$$

The critical energy density for a Universe which is spatially flat has the form:

$$\rho_c \equiv \frac{3H_0^2}{8\pi G}, \tag{1.10}$$

with H_0 being the Hubble parameter at t=0 (today). At t=0, $H_0 = 100h$ km/s/Mpc. From this we can determine that if $\rho_{tot}/\rho_{crit} > 1$ the Universe is open, if $\rho_{tot}/\rho_{crit} < 1$ the Universe is closed and if $\rho_{tot}/\rho_{crit} = 1$ the Universe is flat. A closed Universe is one where space-time is curved back on itself, with a finite space and no edge. An open Universe is that where space time is curved but does not curve back on itself and is infinite. A flat Universe is an infinite un-curved, i.e; zero spatial curvature ($\Omega_K = 0$).

The relation between pressure and energy density can be assumed by the following relation [1], also known as the Equation of state (EoS):

$$P_i = w_i\rho_i. \tag{1.11}$$

where $i \equiv (m, r, \Lambda)$. m, r, Λ are matter, radiation and dark energy contributions in the Universe. We have set a flat Universe, $\Omega_K = 0$ If $w_m = 0$, this describes a pressure-less fluid; when $w_r = 1/3$, the energy content represents radiation and $w_\Lambda = -1$ refers to a cosmological constant filled Universe. When combining the acceleration equations (Eq. 1.9) and the EoS (Eq. 1.11) we see that the energy density evolves with the scale factor

$$\rho_i \propto \frac{1}{a^{3(1+w_i)}}. \tag{1.12}$$

ρ_i describes the densities for matter, radiation and dark energy. Knowing this we can see how the different densities evolved throughout the Universe: $\rho_m \propto a^{-3}$, $\rho_r \propto a^{-4}$ and $\rho_\Lambda \propto a^0 \equiv const$. We can define the dimensionless density parameters today (0) as

$$\Omega_{I,0} \equiv \frac{\rho_{I,0}}{\rho_{crit,0}}, \tag{1.13}$$

where I represents the different components (m, r, Λ). The density parameters become

$$\Omega_{m,0} = \frac{8\pi G\rho_{m,0}}{3H(t)^2}, \tag{1.14a}$$

$$\Omega_{r,0} = \frac{8\pi G\rho_{r,0}}{3H(t)^2}, \tag{1.14b}$$

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2}, \tag{1.14c}$$

we can now rewrite the first Friedmann equation (Eq. 1.7) as

$$H = H_0\sqrt{\Omega_{m,0}a^{-3} + \Omega_{r,0}a^{-4} + \Omega_\Lambda}, \tag{1.15}$$

with $H_0 = 100h$ kms^{-1}Mpc^{-1} being the Hubble constant today and $h \approx 0.67$ is a dimensionless quantity.

Distances

There are a few ways in which we can measure the distance between two points in an expanding Universe. However, unless otherwise stated, we will be using the co-moving distance throughout the thesis. The furthest distance which is accessible today to an ob-

server, is the point from which light has traveled. This distance separates the observable and unobservable regions of the Universe. This is called the co-moving horizon (η) defined by

$$\eta \equiv \int_0^t \frac{c\,dt'}{a(t')}. \tag{1.16}$$

The co-moving distance can be defined as a function of the scale factor or as a function of redshift:

- In terms of the scale factor

$$\chi(a) = \int_a^1 \frac{c\,da'}{(a')^2 H(a')}. \tag{1.17}$$

- In terms of redshift, recalling that $z = 1/a - 1$

$$\chi(z) = \int_0^z \frac{c\,dz}{H(z')}. \tag{1.18}$$

One way to measure distances is through the angular diameter distance (d_A). This method measures the angle θ subtended by an object of known physical size l, provided that $\theta \ll 1$:

$$d_A = \frac{l}{\theta}. \tag{1.19}$$

In an expanding Universe, the co-moving size is given by l/a, where l is the length (distance) and a the scale factor. The co-moving distance out to an object is given by Eq. 1.17 or Eq. 1.18 therefore for a spatially flat Universe the d_A [3] is:

$$d_A = a\chi(a) = \frac{\chi(z)}{1+z}. \tag{1.20}$$

Another method is the luminosity distance d_L, which can be measured from standard candles (Type IA supernovae). If we have a source at a fixed co-moving distance $\chi(z)$, then we can relate the luminosity L and the observed flux F by

$$F = \frac{L}{4\pi\chi(z)^2}. \tag{1.21}$$

This is true for Euclidean space, for FRW space-time the formula is changed to,

$$F = \frac{L}{4\pi\chi^2(z)(1+z)^2} \equiv \frac{L}{4\pi d_L^2} \tag{1.22}$$

for the observed source at redshift. Therefore

$$d_L = \chi(z)(1+z), \tag{1.23}$$

since d_L in Eq 1.22 is defined such that the relation between L, F and d_L is the same as Eq 1.21. The reason for the change is due to the cosmic expansion of the Universe resulting in $(1+z)^{-2}$ term. This term stems from 1. The rate at which photons are emitted and arrive differ by a factor of $1/(1+z)$ and 2. The energy of photons emitted compared to when they are received differ by the same factor $1/(1+z)$.

1.2 Large-Scale Structure

The structure of the Universe in terms of the ΛCDM model, began with small density fluctuations [4]. These fluctuations then coalesced to form denser galaxies and ultimately led to our current Universe. In order to describe how density fluctuations produce large-scale structures, we need to make use of *cosmological perturbation theory*. In this section we shall consider linear Newtonian perturbation theory only in which we perturb the homogeneous cosmological background.

We begin with the energy and momentum conservation for a Newtonian fluid [1, 2],

$$\partial_t \rho + \nabla \cdot (\rho \bar{v}) = 0 \tag{1.24a}$$

$$\partial_t \bar{v} + (\bar{v} \cdot \nabla)\bar{v} = -\frac{1}{\rho}\nabla P - \nabla\phi, \tag{1.24b}$$

where ρ is the density, P is the pressure, $\bar{v}$ is the vector for fluid velocity ϕ is gravitational potential. Eq. 1.24a is the continuity equation and Eq. 1.24b is the Euler equation. The gravitational potential is related to the density by the Poisson equation

$$\nabla^2\phi = 4\pi G\rho. \tag{1.25}$$

Next we define perturbations with y being a generic scalar quantity

$$y = \bar{y} + \delta y, \tag{1.26}$$

where $\bar{y}$ is the homogeneous background and δy is a first order perturbation. To ensure linearity we neglect any perturbations of higher order.

With the introduction now complete we start with the static space while ignoring the effects of gravity ($\phi = 0$), the background terms under this assumption becomes $\bar{\rho} = const., \bar{P} = const.$ and $\bar{v} = 0$. The continuity and Euler equations for first order perturbations now become

$$\begin{aligned} \partial_t(\bar{\rho} + \delta\rho) + \nabla \cdot ((\bar{\rho} + \delta\rho)(\bar{v} + \delta v)) &= 0 \\ \iff \partial_t \delta\rho + \nabla \cdot (\bar{\rho}\delta v + \bar{v}\delta\rho) &= 0 \\ \iff \partial_t \delta\rho + \nabla \cdot (\bar{\rho}\delta v) &= 0 \end{aligned} \tag{1.27}$$

and

$$\begin{aligned} \partial_t \delta v + ((\bar{v} + \delta v) \cdot \nabla)(\bar{v} + \delta v) &= -\frac{1}{\bar{\rho}} \nabla(\bar{P} + \delta P) + 0 \\ \iff \partial_t \delta v &= -\frac{1}{\bar{\rho}} \nabla \delta P \\ \iff \bar{\rho}\partial_t \delta v + \nabla \delta P &= 0. \end{aligned} \tag{1.28}$$

Combining the partial time derivative of continuity equation (Eq. 1.27) and the gradient for the Euler equation (Eq. 1.28):

$$\begin{aligned} \partial_t^2 \delta\rho &= -\partial_t \nabla \cdot (\bar{\rho}\delta v) \\ \partial_t \nabla \cdot (\bar{\rho}\delta v) + \nabla^2 \delta P &= 0 \\ \implies \partial_t^2 \delta\rho &= \nabla^2 \delta P. \end{aligned} \tag{1.29}$$

In adiabatic fluctuations, the pressure fluctuations are proportional to the density fluctuations therefore[2]; $\delta P = c_s^2 \delta\rho$ where c_s is the speed of sound in a fluid,

$$\partial_t^2 \delta\rho - c_s^2 \nabla^2 \delta\rho = 0. \tag{1.30}$$

Turning on gravity we obtain the source term for the Euler equation

$$\partial_t^2 \delta\rho - c_s^2 \nabla^2 \delta\rho = 4\pi G \bar{\rho} \delta\rho, \tag{1.31}$$

with the perturbed Poisson equation $\nabla^2\delta\phi = 4\pi G\delta\rho$.

Since we live in an expanding Universe, we need to be able to describe the fluid in an expanding space. We define again (for this section we use the following notation) the relation between **r** the physical coordinate and **x** the co-moving coordinate

$$\mathbf{r} = a\mathbf{x}. \tag{1.32}$$

The velocity field can be described by

$$\mathbf{v}(t) = H\mathbf{r} + \mathbf{u}, \tag{1.33}$$

with $H\mathbf{r}$ being the Hubble flow, which describes the motion of galaxies as result of the expansion of the Universe. **u** is the proper velocity of an object (galaxy) with the cosmological reference frame set by the expansion of the Universe [5] (pg 137). The gradient in co-moving coordinates becomes

$$\nabla_{\mathbf{r}} = a^{-1}\nabla_{\mathbf{x}} \tag{1.34}$$

and the relation between time derivatives at a fixed **x** and fixed **r** becomes

$$\left(\frac{\partial}{\partial t}\right)_{\mathbf{r}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} + \left(\frac{\partial \mathbf{x}}{\partial t}\right)_{\mathbf{r}} \cdot \nabla_{\mathbf{x}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} + \left(\frac{\partial a^{-1}\mathbf{r}}{\partial t}\right)_{\mathbf{r}} \cdot \nabla_{\mathbf{x}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} - H\mathbf{x}\cdot\nabla_x. \tag{1.35}$$

Substituting Eq. 1.34 and Eq. 1.35 into Eq. 1.24a, we obtain the continuity equation in an expanding Universe,

$$\left[\frac{\partial}{\partial t} - H\cdot\nabla\right][\bar{\rho}(1+\delta)] + \frac{1}{a}\nabla\cdot[\bar{\rho}(1+\delta)(Ha\mathbf{x}+\mathbf{v})] = 0 \tag{1.36}$$

where

$$\delta \equiv \frac{\delta\rho}{\bar{\rho}} \tag{1.37}$$

is defined as the fractional density perturbation or density contrast. At zeroth order fluctuations we have

$$\frac{\partial\bar{\rho}}{\partial t} + 3H\bar{\rho} = 0, \tag{1.38}$$

note that $\nabla_{\mathbf{x}}\cdot\mathbf{x} = 3$. At first order fluctuations (linear in δ and $\mathbf{v} = 0$) we get

$$\left[\frac{\partial}{\partial t} - H\cdot\nabla\right][\bar{\rho}\delta] + \frac{1}{a}\nabla\cdot[\bar{\rho}Ha\mathbf{x}\delta + \bar{\rho}\mathbf{v}] = 0, \tag{1.39}$$

which can be re-written as

$$\left[\frac{\partial\bar{\rho}}{\partial t}+3H\bar{\rho}\right]\delta+\bar{\rho}\frac{\partial\delta}{\partial t}+\frac{\bar{\rho}}{a}\nabla\cdot\mathbf{v}=0. \tag{1.40}$$

The first term is zero by Eq. 1.38, therefore

$$\dot{\delta}=-\frac{1}{a}\nabla\cdot\mathbf{v}, \tag{1.41}$$

$\dot{\delta}$ is the derivative with respect to time. A similar process for the Euler equation, Eq. 1.24b gets us

$$\dot{\mathrm{v}}+H\mathbf{v}=-\frac{1}{a\bar{\rho}}\nabla\delta P-\frac{1}{a}\nabla\delta\phi \tag{1.42}$$

and the Poisson equation, Eq. 1.25 becomes

$$\nabla^2\delta\phi=4\pi Ga^2\bar{\rho}\delta. \tag{1.43}$$

Combining the partial time derivative of Eq. 1.41 with the gradient dot product of Eq. 1.42 and Eq. 1.43 we obtain the time evolution for matter perturbations

$$\ddot{\delta}+2H\dot{\delta}-\frac{\mathrm{c}_\mathrm{s}^2}{a^2}\nabla^2\delta=4\pi G\bar{\rho}\delta. \tag{1.44}$$

Matter Power Spectrum

Based on the nature of Eq. 1.44 we can decompose the density contrast (δ) as a plane wave,

$$\delta(\vec{x},t)=\int\frac{d^3k}{(2\pi)^3}\hat{\delta}(\vec{k},t)e^{i\vec{k}\cdot\vec{x}} \tag{1.45}$$

$\hat{\delta}$ is the Fourier transform of the density contrast

$$\hat{\delta}(\vec{k})=\int d^3x\delta(\vec{x})e^{-i\vec{k}\cdot\vec{x}}. \tag{1.46}$$

The statistical properties of the density contrast do not change under rotations and translations. The reason is that δ is an isotropic and homogeneous random field. Therefore

$$\langle\delta(\vec{x})\rangle=\left\langle\frac{\rho(\vec{x})-\rho_0}{\rho_0}\right\rangle=\frac{\langle\rho(\vec{x})\rangle}{\rho_0}-1=0, \tag{1.47}$$

the mean of the density contrasts vanishes. The variance in Fourier space defines the power spectrum at a given value of k

$$\langle \hat{\delta}(\vec{k})\hat{\delta}^*(\vec{k'})\rangle = (2\pi)^3 P(k)\delta^D(\vec{k}-\vec{k'}), \tag{1.48}$$

where δ^D is the Dirac delta function which accounts for the power spectrum symmetry. Isotropy implies that the wavenumber k is not directional dependent, therefore the power spectrum is represented as $P(\vec{k}) = P(k)$. In real space the power spectrum is called the 2-point correlation function:

$$\langle \delta(\vec{x}_1)\delta^*(\vec{x}_2)\rangle = \xi(|\vec{x}_1 - \vec{x}_2)|) = \xi(\vec{r}) = \xi(|\vec{r}|), \tag{1.49}$$

which is related to the power spectrum via a Fourier transform.

We can describe the shape of the initial power spectrum (P_0) which is imprinted in the early Universe by power a law function, $P_0(k) = Ak^n$ [6] . The need now arises to process the power spectrum from $P_0(k) \rightarrow P(k)$. This is achieved with the aid of the *transfer function* ($T(k)$). The transfer function describes the evolution of perturbations through the epochs of horizon crossing and radiation or matter transition [1]. However the evolution of the power spectrum at late-times is described by the *growth function* ($D_1(z)$). T and D solve the general form of the equation of motion Eq. 1.44. We are now able to define the power spectrum for matter distribution at late-times as [1]:

$$P_m(k,z) = P_0 T(k)^2 \left(\frac{D_1(0)}{D_1(z)}\right)^2. \tag{1.50}$$

In reality we can not measure the CDM directly, however since galaxies form in large CDM halos (peaks in the density contrast) we can use them as a proxy. This results in a bias factor (b) in order to relate the CDM density with galaxy density contrast:

$$b \equiv \frac{\delta_g}{\delta}, \tag{1.51}$$

$$P_x(k,z) = b_x^2 P_m(k,z), \tag{1.52}$$

where x relates to different biased tracers such as "g" for galaxies or "HI" for neutral hydrogen.

Redshift-Space Distortion (RSD)

When observing galaxies along the line of sight in redshift space, there is a squeezing effect which perturbs the measured redshift of these galaxies. This effect is caused by the peculiar velocities of these galaxies which perturb their position in redshift space. Then correction for this effect in linear theory is known as the Kaiser formula [7] which is given by:

$$\delta_{g,obs}(\mathbf{k}) = (b + f\mu^2)\delta_m(\mathbf{k}), \tag{1.53}$$

we refer the reader to section 9.4 [1] for the detailed derivation. The logarithmic growth rate is defined as $f \equiv dlnD/dlna \sim \Omega_M(z)^\gamma$ where $\gamma \sim 0.55$ in GR [1, 8]. The power spectrum now becomes anisotropic as well and by substituting Eq. 1.53 into Eq. 1.48

$$\langle \hat{\delta}_{g,obs}(\vec{k})\hat{\delta}^*_{g,obs}(\vec{k'})\rangle = (b + f\mu^2)^2\langle \hat{\delta_m}(\vec{k})\hat{\delta_m}^*(\vec{k'})\rangle, \tag{1.54}$$

we can express the observed galaxy power spectrum as:

$$P_{g,obs}(z,\mu,k) = (b + f\mu^2)^2 P_m(z,k). \tag{1.55}$$

1.3 Baryonic Acoustic Oscillations as a standard ruler

Understanding the expansion history of the Universe can place better constraints on the current dark energy measurements in our Universe. To measure the expansion, we can rely on the luminosity distance, Eq. 1.21 or the angular diameter distance. Eq. 1.19, which are functions of redshift. In the case of luminosity distance, observers utilize type Ia Supernovae (SNE). These objects have an intrinsic brightness which is known and we can measure their apparent brightness (refer to Eq. 1.22). These are known as standard candles and their measurements have shown the expansion of the late Universe is accelerating [9, 10]. Another method of measuring the cosmological expansion is by using an objects size, if the intrinsic size of an object is known then one can measure the angular size of that object and because size changes as a function of redshift, we are able to calculate the expansion. These are known as standard rulers.

Standard rulers have been around since the mid 90's and consists of various astrophysical species such as ultra compact radio sources or double lobed radio sources (Fanroff-Riley TypeIIb radio sources) [11]. A variation of standard rulers is called Statistical Standard Rulers (SSR) which utilizes the idea that galaxies cluster together at a preferred length scale. Therefore if we were to observe these galaxies at various redshifts we would be able to

constrain the angular diameter distance [11]. In this thesis we are interested in one SSR called the Baryonic Acoustic Oscillation (BAO).

In the early Universe, $z \geq 1000$ the coupling (induced by Thomsom scattering) of the photons and baryons in the ionized Universe resulted in perturbations behaving as sound waves [12]. The speed at which these acoustic waves moved outward for a single density perturbation is given by

$$c_s = c/\sqrt{3(1+R)}. \tag{1.56}$$

where c_s is the speed of sound in a fluid, c is the speed of light and R is the baryon-to-photon relation given by $R \equiv 3\rho_b/4\rho_\gamma \propto \Omega_b/(1+z)$ [13]. The comoving distance that a sound wave can travel in a plasma since the Big Bang [14] is given by the sound horizon (r_s) [15]:

$$\begin{aligned} r_s &= \int_0^{t_{rec}} c_s(1+z)dt \\ &= \int_{z_{rec}}^{\infty} \frac{c_s dz}{H(z)} = \frac{1}{\sqrt{\Omega_m H_0^2}} \frac{2c}{\sqrt{3z_{zeq}R_{eq}}} ln \frac{\sqrt{1+R_{rec}} + \sqrt{R_{rec}+R_{eq}}}{1+\sqrt{R_{eq}}}, \end{aligned} \tag{1.57}$$

where "*rec*" and "*eq*" refer to recombination and equality respectively. During the recombination period, decoupling of the photon-baryons occur; this causes the Universe to move into a state of neutrality. It is at this point that the baryon wave stalls and photons are able to propagate freely away, giving rise to the Cosmic-Microwave-Background (CMB).

Figure 1.1 (obtained from [13]) graphically explains the evolution of the oscillations, after recombination when baryons and photons decouple. The *top left* panel, shows the initial point in time when baryons and photons are tightly coupled together, which is before recombination. The *top right* panel shows around the time of recombination, that the propagation of baryons and relativistic species induces a pull effect on CDM, which increases its perturbation. The *middle left* panel represents some time after recombination, the point at which baryons and photons decouple. The *middle right* panel occurs well after recombination, the baryonic perturbation (blue) produces a bump that occurs around 150 Mpc scale while the CDM perturbation is near the centre. The *bottom left* panel shows the gravitational attraction between the baryons and dark matter over-densities, where baryons now mainly follow the distribution of the CDM. In the *bottom right* panel, most of the growth is drawn from the homogeneous bulk, the baryon fraction converges toward the cosmic mean at late times [8]. This implies that there are two points for galaxy growth; at the origin and at 150 Mpc radius. The BAO 'bump' (Figure 1.1: *bottom right*) can be measured from large-scale structure surveys, that map out the matter distribution (a tracer for the underlying dark matter) by

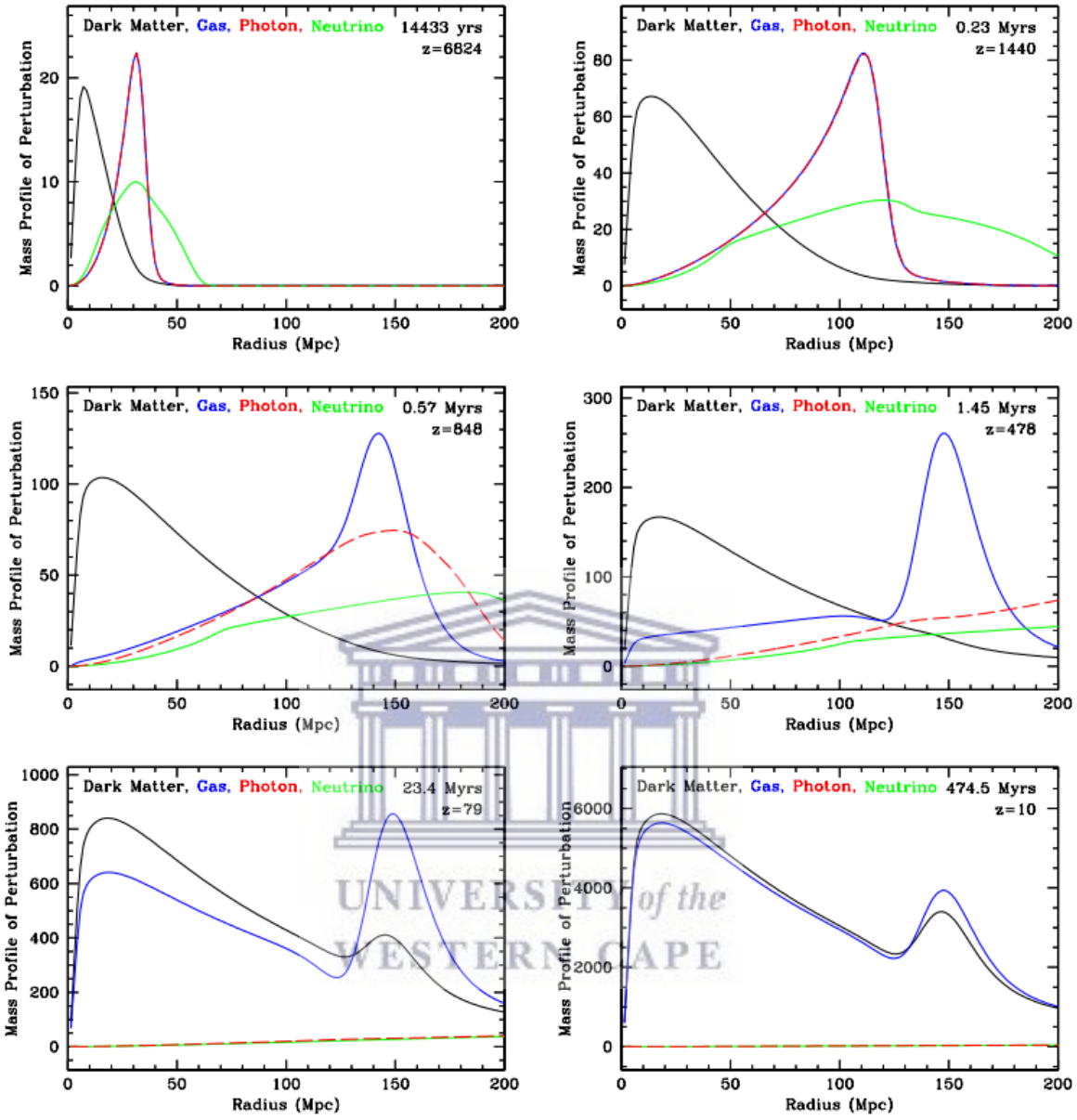


Fig. 1.1 Figure taken from [13]. Each panel shows the radial perturbed mass profile in each of the four species: dark matter (black), baryons (blue), photons (red), and neutrinos (green). The redshift and time after the Big Bang are given in each panel.

applying the matter correlation function

$$\xi(\mathbf{r}) = \langle \delta_m(\mathbf{x}) \delta_m(\mathbf{x}+\mathbf{r}) \rangle. \tag{1.58}$$

The main advantage of the BAO feature when measuring distances is its robustness towards systematic effects. Since the distance information is encoded into the location of the power

spectrum or correlation function rather than the amplitude; it becomes difficult for non-cosmological effects to mimic or shift the BAO. Detection of the BAO feature has been accomplished in the optical wavelength thus far. By measuring the position and redshift of enough galaxies, surveys such as 2dFGRS[1] [16] (2dF Galaxy Redshift Survey, Figure 1.2) and the SDSS[2] [17] (Sloan Digital Sky Survey, Figure 1.3) have obtained the first results.

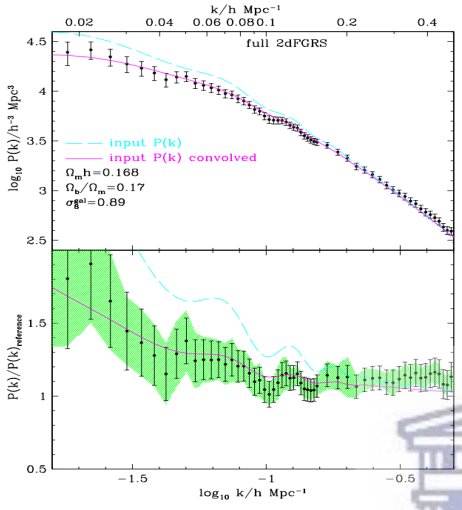


Fig. 1.2 The recovered 2dFGRS redshift space galaxy power spectrum (top) and after division of the expected ΛCDM smooth power spectrum (bottom) [16].

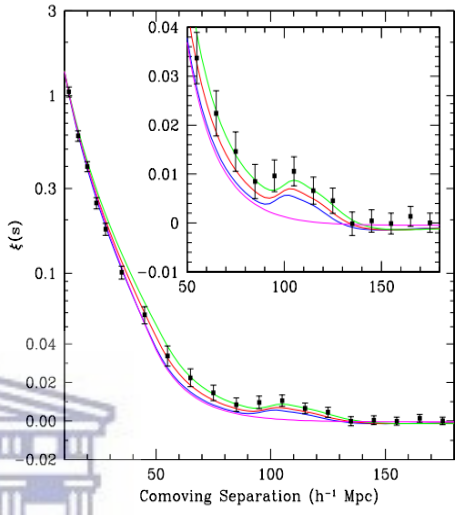


Fig. 1.3 The large-scale redshift space correlation function, Eq 1.58 of the SDSS LRG sample [17].

Figure 1.2 shows that in Fourier space, the BAO feature is represented as wiggles. The smooth power spectrum refers to a power spectrum which does not include baryons and therefore do not possess the BAO wiggles. While in Figure 1.3 the BAO is shown in real space (by mean of the correlation function) as a bump.

1.4 Radio Telescopes

Radio telescopes are instruments designed to measure the radio emission of sources that exist outside of Earth's atmosphere. These sources emit in the radio frequency range, which extends from 30 MHz to 300 GHz and vary in morphology (stars, galaxies, quasars).

In 1933, Karl Jansky, named as the "Father of Radio Astronomy"[18] was the first person to investigate these radio emissions. Jansky constructed a mounted antenna in order

[1] http://www.2dfgrs.net/
[2] https://www.sdss.org/

to detect shortwave interference for Bell Telephone Laboratories and in order to find the direction of interference he made the instrument fully steerable. He began pointing at the sky and found that a source of radio noise was coming from the centre of the Milky Way Galaxy. In 1937 Grote Reber continued the work that Janksy started, Reber built the world's first radio telescope, a 9.5 metre parabolic reflector from which he was able to produce the first radio sky map [19].

Since Jansky and Reber, radio astronomy and radio telescopes have grown tremendously with the later varying widely depending on the science it wants to investigate. Even with the variation, the basis of any radio telescope is to have a large radio antenna and a sensitive radio receiver. A telescopes ability to measure faint radio sources relies on the antenna area and efficiency as well as the sensitivity of the amplifier (receiver) used in signal detection. Over time radio telescopes have grown in both size and sensitivity levels, stretching to sizes over hundreds of meters across with state of the art receivers in order to probe fainter cosmic radio signals.

The structural design of radio telescopes has also grown throughout the decades, the most common design is that of a reflector. A reflector consists of a parabolic antenna (dish) which reflects the incoming radio waves towards the guide horns. The guide horns channel and propagate the radio waves to sensitive receivers. These receivers were originally positioned directly at the focal point however the positioning created limitations in the form of accessibility to the receivers and the weight (limits size and number of receivers). In order to alleviate these limitations a secondary reflector can be introduced which focuses the radiation near to the centre or vertex of the dish. This allows for more receivers to be positioned as the weight restriction is more relaxed. The advantage of this design when compared to radio telescopes of the past which used a symmetric tripod or quadrapod structure is that, there is no obstruction for the incoming radiation [20]. Therefore off-axis telescopes are more sensitive and have less structural interference coming from the feeds support structures.

The ability for a radio telescope to look at an astronomical object with great detail (angular resolution) is determined by the wavelength observed divided by the size of the collecting area. This is problematic for radio telescopes as the largest of these instruments operating at the shortest wavelength is notably weaker in resolution than their optical counterparts. Radio telescopes therefore require much larger sizes since they operate at larger wavelengths, however building gargantuan size radio dishes is not a feasible idea and thus we introduce interferometry.

The concept of interferometry is to combine a number of smaller antennae together in order to create a large aperture. In the case of a two antenna radio interferometer, the signal of a source is measured by the antenna pair as an interference pattern. Due to the rotation of the Earth the path from source to receiver changes creating interference fringes, the difference in path length of each antenna would vary across the source if the source had a finite angular size. The measurements in the interference fringes would therefore be dependent on the radio source brightness across the sky.

We have come along way since Jansky and Reber, creating more sensitive instruments capable of seeing further with more much sensitivity meaning we can see fainter sources not yet detected. A new era of radio observations awaits for instance likes of the Square Kilometre Array and its vanguard the Meer Karoo Array Telescope (MeerKAT).

1.5 SKA and MeerKAT

The Square Kilometre Array (SKA), once completed will be the world's largest radio telescope with a collecting area of approximately 1km^2, consisting of thousands of dishes and up to a million low frequency antennas [21]. It will be built across the African continent (majority in South Africa) and Australia. South Africa and it's African partner countries will host the mid to high frequency antennae and Australia will host the low frequency antennae. Currently the SKA programme in its initial stages where the two host countries have precursors already running; the Australian Square Kilometre Array Pathfinder (ASKAP) and MeerKAT which is a successor of the Karoo Array Telescope (KAT-7) in South Africa. Phase 1 (SKA1) and Phase 2 (SKA2) will spread across the African and Australasian Continent.

- Phase 1
  SKA1 will be split into two different instruments: an interferometer array-SKA1-LOW and a dish array SKA1-MID. SKA1-LOW will be made of 512 stations located in Australia and operating in the $50 < \nu < 350$ MHz range. SKA1-MID will consist of 197 dishes (64 of MeerKATs' 13.5m dishes and 133 SKA1-MIDs' 15m dishes) located in South Africa, it will have a compact core comprising of MeerKATs dishes and a maximum baseline of 150km. The operating frequency will comprise of three frequency bands: 350MHz $< \nu <$ 1050MHz, 950MHz $< \nu <$ 1750MHz and 4.6GHz $< \nu <$ 15.8GHz (See [22] for more details).

- MeerKAT

  MeerKAT located in Canarvon South Africa, is the precursor to SKA1-MID and is currently the world's most sensitive radio telescope in operation. It consists of 64 interlinked receptors with 48 receptors being in a 1 km diameter making up the core and having a maximum baseline of 8 km. The antenna design is that of an Offset Gregorian with the main reflector being 13.5 m in diameter and the secondary being 3.8 m [23]. MeerKAT will conduct a range of large survey projects, which have already been allocated observation time, including the following:

  – Radio Pulsar Timing [24];
  – LADUMA - Looking at the Distant Universe with the MeerKAT Array [25, 26];
  – MALS - MeerKAT Absorption Line Survey [27];
  – MHONGOOSE - MeerKAT HI Observations of Nearby Galactic Objects: Observing Southern Emitters [28, 29];
  – TRAPUM - Transients and Pulsars with MeerKAT [30];
  – A MeerKAT HI Survey of the Fornax Cluster [31];
  – MIGHTEE - MeerKAT International GigaHertz Tiered Extragalactic Exploration Survey [32];
  – ThunderKAT - The Hunt for Dynamic and Explosive Radio Transients with MeerKAT [33].

1.6 Cosmology with SKA1-MID and MeerKAT

The various cosmological surveys that MeerKAT and eventually SKA1-MID will conduct would be to probe the Large-Scale Structure (LSS). The SKA1-MID cosmology surveys include, Continuum galaxy [34], HI galaxy redshift [35] and HI Intensity Mapping [36–38].

The above mentioned survey classes have been proposed for SKA1-MID as cosmology surveys [22]. They have been named *Medium-Deep Band 2 Survey* and *Wide Band 1 Survey*.

- Medium-Deep Band 2 Survey
  A survey covering 5000 deg^2 field with an observation time of 10000 hours. This survey should accomplish the weak continuum lensing survey and a HI galaxy survey up to $z \sim 0.4$.

- Wide Band 1 Survey
  This survey will look at a 20000 deg^2 sky area with observational time of ≈10000 hours. The aim would be to do a wide continuum galaxy survey and a HI intensity mapping survey ranging from $z = 0.35 - 3$

These surveys will cover a volume that is unparalleled to any previous or planned radio survey. Figure 1.4 shows the volume and the redshift dominance of these surveys.

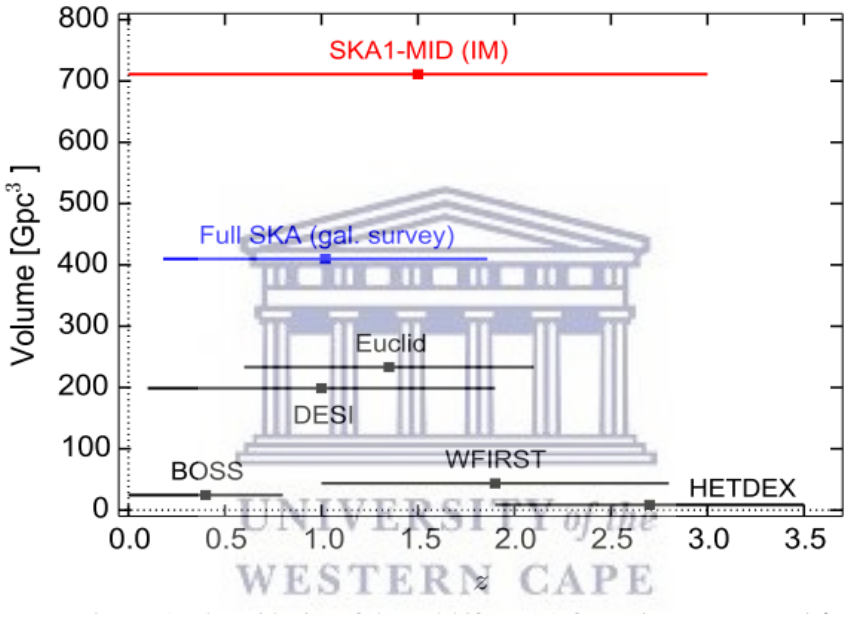


Fig. 1.4 Survey volumes (calculated at central redshift) for various current and future surveys [37]. IM refers to intensity mapping, which will be discussed in Chapter 2. The remaining surveys, Euclid [39] (space based telescope) and DESI [40], BOSS [41], HETDEX [42] (ground based telescopes) are optical telescopes. WFIRST [43] (space based telescope) operates in the infrared.

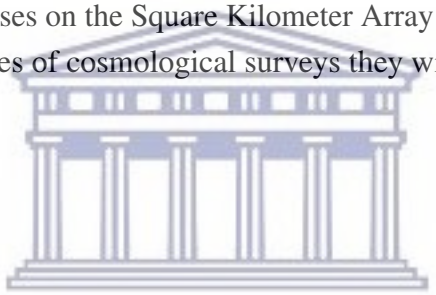
Before SKA1-MID becomes operational there exists its precursor MeerKAT, which has a proposal for its own cosmological survey called *MeerKLASS* [38]. The MeerKAT Large Area Synoptic Survey (MeerKLASS) is a large area survey covering ∼4000 deg^2 for ∼4000 hours and will use the intensity mapping technique to detect the BAO and RSD effects in auto-correlation. Auto-correlation refers to the MeerKAT telescope being in single dish mode, where each pointing of the telescope gives one single pixel on the sky [38]. The advantage of using single dish mode to scan the sky, is the ability to probe very large cosmological scales. While in interferometry mode, the maximum scale is given by the field of view (FoV) of the dish. This would limit the cosmological constrains as one would be bound to consider

mainly non-linear scales which are hard to model and carry less cosmological information. Therefore the best is to look at larger scales which are linear.

MeerKLASS is not reserved as a cosmological survey, the survey will conduct other astronomical sciences. In the realm of cosmology, it will conduct dark matter detection, continuum galaxy surveys and 21cm intensity mapping. The 21cm IM survey for MeerKLASS will be the main focus of this thesis which is done in the Chapter 2 and 3.

1.7 Summary

In this chapter we review the FRW Universe and define the matter power spectrum from perturbation theory. We give a brief overview of Baryon Acoustic Oscillations and how we can measure it in Universe. A brief history of radio telescopes is given in the next section. The remaining two sections focuses on the Square Kilometer Array and the MeerKAT telescopes, in particular the various types of cosmological surveys they will undertake.



Chapter 2

HI Intensity Mapping with MeerKAT

In this chapter we discuss why HI IM is used as a means to map out the Large-Scale Structures (LSS) of the Universe. First we discuss the method and how it benefits large-scale observations. Second we investigate the effects that a telescope beam would have on the 21cm radial power spectrum. Then we consider the effects of cosmological foregrounds on the overall 21cm signal and how we remove them. Lastly we discuss the simulation process in producing intensity maps.

HI Intensity Mapping [14, 37, 44, 45], is a novel technique that uses cosmic neutral hydrogen to map the large-scale structure of the Universe in three spatial dimensions [46]. The underlying basis being that we do not detect individual galaxies, but rather measure the combined flux of a large area which could contain more than one HI galaxy source [44]. This technique can be seen in Figure 2.1 and offers a map of the sky similar to the CMB map, with the added benefit that the 21cm signal would now be a function of redshift as well [36].

The IM technique can be carried out by two different telescope configurations, namely single dish and interferometry. Single dish (auto-correlation) mode results in each pointing of the telescope array corresponding one pixel on the sky. The advantage is that we are able to scan large-scale modes, however we are limited in angular resolution due to smaller dish sizes [38]. The latter configuration (interferometry) measures the Fourier modes, this provides an advantage in high angular resolution and low sensitivity to systematic noise. The drawback is that the largest angular scale that can be probed is set by the minimum baseline.

Figure 2.1[1] gives a cartoon description of IM process. The first successful attempt, us-

[1]https://www.thphys.uni-heidelberg.de/~trr33/pub/Main/WinterSchool12Slides/Pritchard_6.pdf

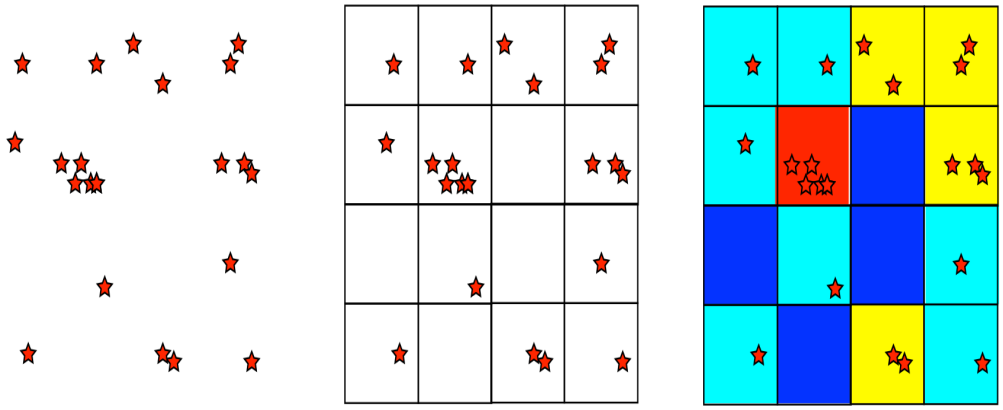


Fig. 2.1 The left most figure reveals sources in the field of view of the telescope which would not be resolved. The center image shows the sky being sliced into a grid of cells. Each cell contains an 'n' amount of sources. The right most image gives a visual of how IM can be demonstrated. Cells with higher number of sources would produce higher intensity and this is depicted in the cells. Cells which have a low source count would give off low temperature (intensity) response.

ing intensity mapping to measure the HI signal was accomplished with the Green Bank Telescope (GBT) in 2010 [47, 48] in cross-correlation only.

2.1 The HI brightness temperature

During the evolution of the Universe there existed a period known as the Dark ages, which took place after Recombination and before the first stars began emitting radiation (Re-ionization). During this period the Universe existed in a state of charge neutrality. Hydrogen existed in its atomic state and resided in two mediums: the Inter-galactic Medium (IGM) between galaxies and the Inter-Stellar Medium (ISM) inside the galaxies. Once the first stars began emitting radiation (UV), the Universe moved to a period of re-ionization. Atomic hydrogen which existed in the IGM became ionized by the radiation while the atomic Hydrogen which existed in the ISM was protected by the high density gas clouds [37]. This stage is known as the Era of Re-ionization (EoR). HI therefore predominantly resides in the ISM of galaxies. To this day we can use the HI as a probe to map out galactic structures as well as the distribution of HI rich galaxies throughout the Universe. As a proxy we can utilize the HI distribution to measure the distribution of the underlying dark matter [44].

The 21cm emission line is a strong contender for radio observations and experiments. The

emission, a result of a rare spin transition otherwise known as hyperfine transition, occurs when the spin of the proton and electron align (same direction) and then transition to skew (opposite direction). During this process a photon is emitted at 1420.4 MHz or 21cm, as seen in Figure 2.2 and therefore the observed wavelength is $\lambda_{\text{obs}} = (1+z) \times 21\text{cm}$, which means large wavelengths and low frequencies (radio band). Photons at this wavelength do not get obstructed from cosmic dust and the Earth's atmosphere, so it can be detected by ground based radio telescopes.

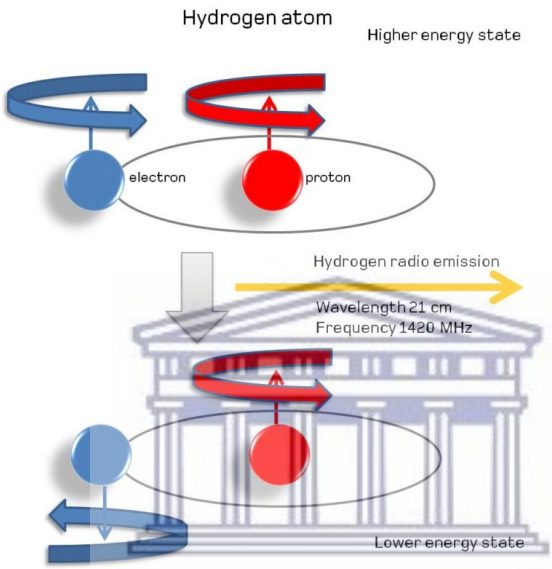


Fig. 2.2 An artistic view of the 21cm Hyperfine transition [49].

Determining the HI brightness requires a model to explain how the HI density evolves over redshift. The derivation of the HI density (Ω_{HI}) is summarized from Appendix A of [44],

$$\Omega_{HI}(z) = \frac{1}{\rho_c^0} \int_0^\infty n(M,z) M_{HI}(M,z) dM \tag{2.1}$$

Where ρ_c^0 is the critical density of the Universe today and $n(M,z)$ is the halo mass function at redshift z [50]. The density is dependant on the average HI mass function $M_{HI}(M,z)$, which takes in M (dark matter halo mass) at a given redshift (z) and produces the HI density.

The $M_{HI}(M,z)$ can be modeled by a power law function, based upon hydro-dynamical simulations of [51] as follows,

$$M_{HI} = M^{3/4} exp\left[\gamma(z) - \left(\frac{M_{min}(z)}{M}\right)^2\right]. \tag{2.2}$$

Where $e^{\gamma(z)}$ is the general normalization and $M_{min}(z)$ represents the minimum halo mass that can sustain HI [44, 52].

We are then able to model the Ω_{HI} as:

$$\Omega_{HI} = 4 \times 10^{-4}(1+z)^{0.6} \tag{2.3}$$

The HI temperature brightness for a given redshift in mK can be described by:

$$\bar{T}_b(z) = 190\frac{H_0(1+z)^2}{H(z)}\Omega_{HI}(z)h \tag{2.4}$$

and is obtained from [37, 44, 46].

2.2 21cm Power Spectrum

In order to obtain the linear 21cm power spectrum we need to be able to understand how the HI bias (refer to Eq. 1.51) evolves with redshift (similar to Ω_{HI}, the derivation is summarized from Appendix A of [44]). The HI bias (b_{HI}) is defined below

$$b_{HI}(z) = \frac{1}{\rho_c^0 \Omega_{HI}(z)} \int_0^\infty b(M,z) n(M,z) M_{HI}(M,z) dM, \tag{2.5}$$

where $b(M,z)$ is the dark matter halo bias for a given redshift [53]. From Eq. 2.5 we can now model b_{HI} [44] as,

$$b_{HI}(z) = 0.904 + 0.135(1+z)^{1.696}. \tag{2.6}$$

With the HI bias now defined we can describe the HI power spectrum as

$$P_{HI}(k,z) = \bar{T}_b^2(z) b_{HI}^2(z)[1+\beta_{HI}\mu^2]^2 P_m(k,z), \tag{2.7}$$

where $\mu = \hat{k}\cdot\hat{z}$, $\beta_{HI} = f(z)/b_{HI}(z)$ is the RSD parameter [46] and $f(z)$ is the linear growth rate [44]. The 21cm power spectrum shown in Figure 2.3 does not include RSD effects only

the 21cm bias. The two vertical lines at $k = 0.02$ and 0.45 h/Mpc, following [36] shows the BAO wiggle feature in wavenumber.

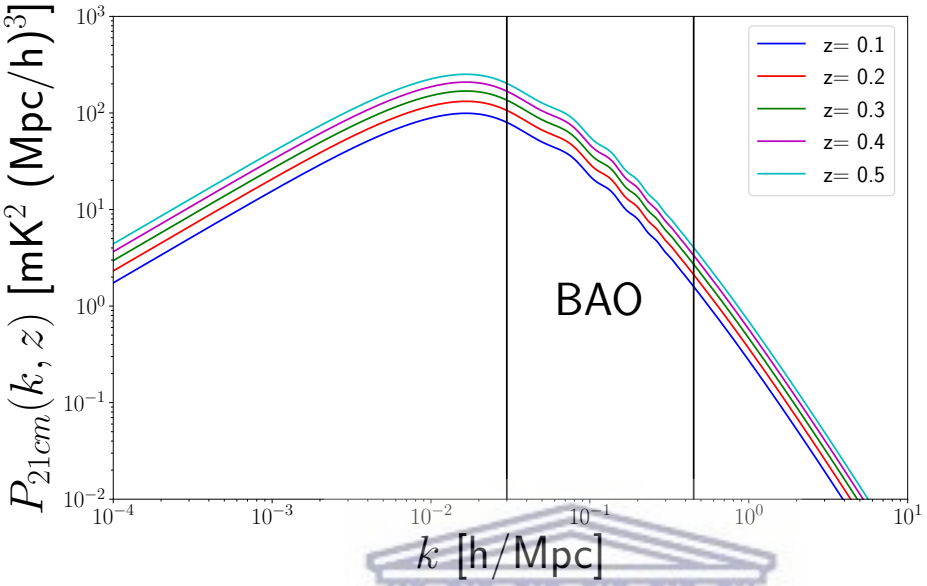


Fig. 2.3 The 21cm Power Spectrum for different redshift increments in the MeerKAT L-band range. We applied Eq. 2.7 and have ignored the RSD effect in this instance. The section between the vertical lines ($0.02 \leq k \leq 0.45$) indicate the position of the BAO. Image produced using CAMB [54]

An increase in z results in an increase in the volume which causes and amplitude increase in the power spectrum ($P_{21cm}(k,z)$).

2.3 Experiments

There are various planned experiments across the globe to probe the 21cm signal using IM. To name a few: the Giant Meterwave Radio Telescope (GMRT)[2], the Ooty Radio Telescope (ORT)[3], the Canadian Hydrogen Intensity Mapping Experiment (CHIME)[4], the Five hundred Aperture Spherical Telescope (FAST)[5], Tianlai [6], Baryonic acoustic oscillations In Neutral Gas Observations (BINGO)[7], Australian SKA Pathfinder (ASKAP) [8], South African

[2]http://gmrt.ncra.tifr.res.in/
[3]http://rac.ncra.tifr.res.in/
[4]http://chime.phas.ubc.ca/
[5]http://fast.bao.ac.cn/en/
[6]http://tianlai.bao.ac.cn
[7]http://www.jb.man.ac.uk/research/BINGO/
[8]http://www.atnf.csiro.au/projects/askap/index.html

SKA Pathfinder (MeerKAT)[9], the Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX)[10]and the SKA [55].

2.4 Beam effects

Radio telescopes (such as MeerKAT) have a primary beam which induces instrumental effects on the power spectrum signal and more importantly the BAO signal. We can show that for a Gaussian (used for simplicity) primary beam, the BAO signal behaves differently in the transverse and radial (line of sight) directions. For now we have ignored the effects of RSD as they are irrelevant for the conclusion of this section, however they will be taken into consideration in the simulation which discussed in chapter 3. The observed 21cm mode ($\delta_{21cm,obs}$) that includes the primary beam, is defined as:

$$\delta_{21cm,obs}(\mathbf{k}_\perp, k_\parallel, z) = e^{-k_\perp^2 R^2/2}\delta_{21cm}(\mathbf{k}_\perp, k_\parallel, z), \tag{2.8}$$

where δ_{21cm} is the 21cm cosmological mode and R is defined as the smoothing scale. The smoothing scale is related to the angular resolution of the beam as:

$$R = r(z)\theta_{FWHM}/(2\sqrt{2ln2}), \tag{2.9}$$

where $r(z)$ is the co-moving distance and the angular resolution, θ_{FWHM} which is the full width half maximum,

$$\theta_{FWHM} = \frac{\lambda}{D_{dish}}, \tag{2.10}$$

with D_{dish} being the diameter of the dish. In the case of MeerKAT, the angular resolution becomes:

$$\theta_{FWHM} = \frac{\lambda}{D} = \frac{0.21(1+z)}{13.5} = 0.9(1+z)\ \text{deg} \tag{2.11}$$

Isotropic Power Spectrum

The 21cm isotropic power spectrum can be obtained from Eq. 2.8 by taking $\langle \delta_{21cm,obs}\delta^*_{21cm,obs}\rangle$:

$$P_{21cm,obs}(k,\mu,z) = e^{-k^2R^2(1-\mu^2)}P_{21cm}(k,z), \tag{2.12}$$

[9]http://www.ska.ac.za/meerkat/
[10]https://www.acru.ukzn.ac.za/~hirax/

where $\mu = \hat{n} \cdot \hat{k} = \frac{k}{k_\parallel}$. Averaging over μ we obtain

$$P_{21cm,obs}(k,z) = P_{21cm}(k,z)\frac{1}{2}\int_{-1}^{1} e^{-k^2R^2(1-\mu^2)}d\mu = \frac{\mathscr{D}(kR)}{kR}P_{21cm}(k,z), \tag{2.13}$$

where $\mathscr{D}(x)$ is the Dawson function [11].

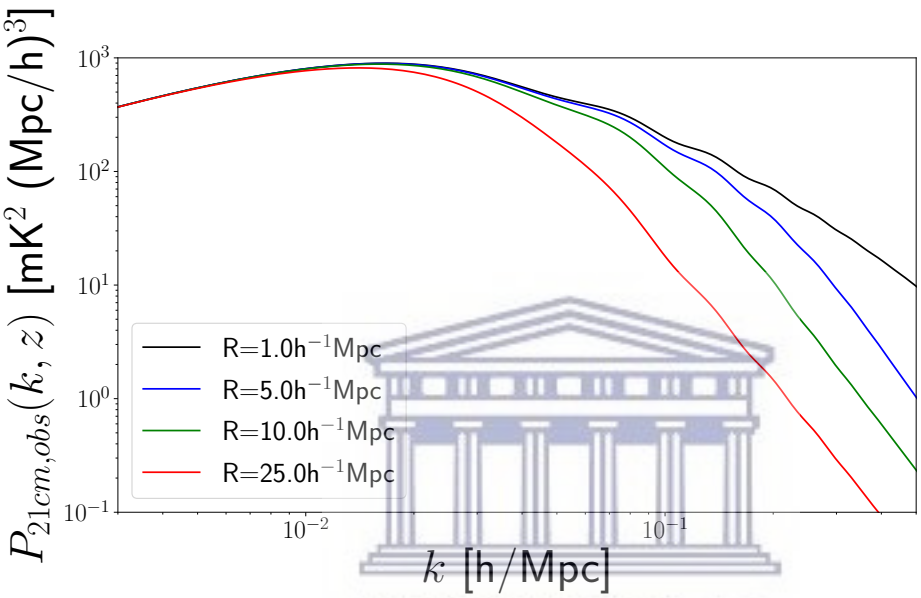


Fig. 2.4 Observed 21cm Power Spectrum for different smoothing scales at z=0.4. The smoothing scales in degrees are given as R = (1, 5, 10, 25) h^{-1}Mpc = (0.12, 0.62, 1.25, 3.12) degrees respectively.

Figure 2.4 shows the observed 21cm power spectrum at z=0.4 for our ΛCDM fiducial cosmology ($\Omega_M = 0.315,\ \Omega_\Lambda = 0.685,$ and $\Omega_k = 0$) that is used throughout this thesis. We can see that the angular smoothing scale (R) starts to smear out (smooth away) the BAO signal as we move to larger smoothing scales. We can show the effects in real space with the 21cm correlation function:

$$\xi_{21cm,obs}(r,z) = \frac{b_{21cm}^2}{2\pi^2}\int_0^\infty k^2 P_m(k,z)\frac{sin(kr)}{kr}\frac{\mathscr{D}(kr)}{kR}dk. \tag{2.14}$$

[11] http://mathworld.wolfram.com/DawsonsIntegral.html

Radial Power Spectrum

The radial power spectrum (also known as the 1D or line of sight power spectrum) follows directly from averaging out the transverse contributions in the 3D power spectrum. The 3D power spectrum that is applied, is taken from Eq. 2.13 which will be substituted into P_{3D} of Eq. 2.15, after which we integrate over the transverse components.

$$P_{1D}(k_{\parallel},z) = \int \frac{d\mathbf{k}_{\perp}}{(2\pi)^2} P_{3D}(k_{\parallel},\mathbf{k}_{\perp},z). \tag{2.15}$$

The line of sight power spectrum receives power from transverse scales larger than the smoothing scale for a given wavenumber k [44].

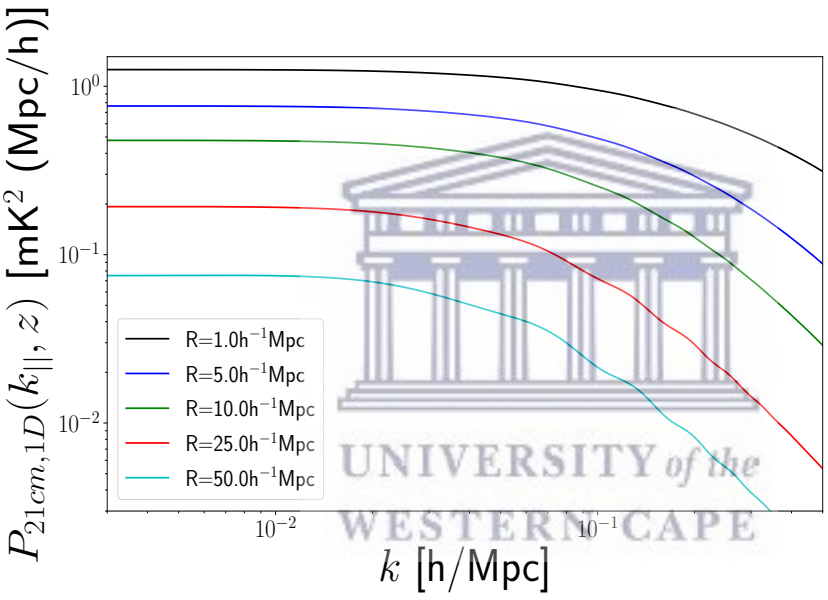


Fig. 2.5 Radial 21cm Power Spectrum for different angular smoothing scales at z=0.4; the impact of the telescope beam can be seen on the wiggles. The smoothing scales in degrees are given as R = (1, 5, 10, 25, 50) h^{-1}Mpc = (0.12, 0.62, 1.25, 3.12, 6.24) degrees respectively.

Figure 2.5 shows how the radial telescope beam (from Eq. 2.15) is affected by the different angular smoothing scales. As the angular smoothing scale increases we see a drop in amplitude however we also see the emergence of BAO feature in the form of wiggles. Figure 2.6 shows the ratio between wiggle and no wiggle (no-BAO) 1D power spectrum for different angular smoothing scales. This shows that the wiggles are visible in all cases but they are more pronounced in larger angular smoothing scales.

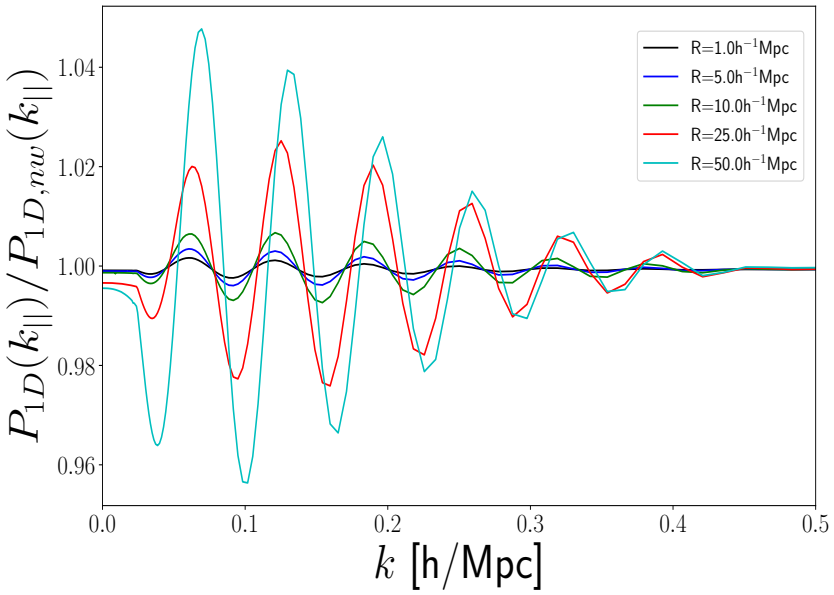


Fig. 2.6 Normalized ratio of radial 21cm power spectrum with and without wiggles. $P_{1D,w}/P_{1d,nw}$ for the radial power spectrum at z=0.4. The expression of $P_{1D,nw}$ is the same $P_{1D,w}$, however we applied the smoothing technique which is described in [36], in order to produce the no wiggle power spectrum

2.5 Foregrounds and Cleaning methods

The existence of foreground sources both terrestrial and extra-terrestrial (E.T) is of great concern to HI IM experiments. Terrestrial sources of contamination can result from the atmospheric noise and radio frequency interference (RFI); while E.T sources are a result from galactic and extra-galactic sources of contamination. E.T sources (foreground sources) are our main focus of contamination since the desired cosmological HI signal is buried beneath these foreground signals which are several orders of magnitude larger. The term foregrounds is a misnomer as it also applies to sources which lie behind the cosmological HI signal as well, such as point sources. It can be shown that the types of foreground that are relevant to this analysis are: Galactic synchrotron, Free-free emission (galactic and extra-galactic) and Point sources. Galactic synchrotron emission, originates from the centre of the Milky Way and in terms of HI IM experiments it is considered to be the dominant foreground. These emissions are a result of high-energy cosmic-ray electrons being accelerated by the magnetic field around the plane of our galaxy [56]. Free-free emission or Bremsstrahlung radiation is a result of free electrons being accelerated by ions. This occurs either inside our galaxy known as galactic free-free or outside known as extra-galactic free-free emission. Finally

Point sources are foregrounds which exist outside our own galaxy, these can comprise of Active Galactic Nuclei (AGN) and star-forming galaxies.

In order to provide forecasts for these foregrounds, we need to be able to model them, which are described in chapter 3. Other sources of foreground contamination were ignored or taken into account when running the simulations. The effects of atmospheric noise are ignored, RFI was taken into account by shifting our survey depth to a higher redshift. Contamination by other frequency lines (for instance OH at $\nu_{OH} \sim 1600$ MHz) would be limited due to spectral isolation of the 21cm line and therefore the HI signal should be robust against line confusion [57].

Removing the effects of foregrounds can be challenging as over/under-estimating the cleaning strength can result in a loss of HI signal or contamination of HI signal. Various cleaning methods have been produced to retrieve the HI signal, these can be categorized as blind and non-blind techniques. Non-blind refers to our understanding of the foregrounds, while blind removal suggests that we do not make any assumptions of the foregrounds. The method we will be imposing is the blind foreground technique called principle component analysis (PCA), which will be discussed in the next chapter.

Foreground Removal

Foreground emissions are a nuisance to intensity mapping experiments, they overwhelm the 21cm signal by 4 to 5 orders of magnitude [58], fortunately there exists methods to remove them. Foregrounds tend to have a smooth dependence with frequency [57], as compared to the cosmological HI signal. The HI signal follows the large-scale structure resulting in having power across a large range of Fourier scales [59]. The foreground removal method that is applied falls under the term blind foreground cleaning, since the method does not assume anything from the foregrounds except that it is smooth in frequency. The blind foreground cleaning method applied in this thesis is called, Principle Component Analysis (PCA) which is described further below.

Below we summarize the technique of foreground removal employed in [57] and applied in the foreground removal software called *fg_rm*[12] [60]

[12]https://github.com/damonge/fg_rm

The total signal that is recovered from the experiment can be understood with the following:

$$T(\nu,\hat{\mathbf{n}}) = \sum_{k=1}^{N_{fg}} f_k(\nu) S_k(\hat{\mathbf{n}}) + T_{cosmo}(\nu,\hat{\mathbf{n}}) + T_{noise}(\nu,\hat{\mathbf{n}}) \tag{2.16}$$

Equation 2.16 gives the sky brightness temperature for a given direction $\hat{\mathbf{n}}$ and frequency ν. The term N_{fg} refers to the number of foreground degrees of freedom to subtract, $f_k(\nu)$ is a set of smooth frequency functions, $S_k(\hat{\mathbf{n}})$ represents the foreground maps and T_{cosmo} and T_{noise} are the cosmological signal and instrumental noise components respectively.

We can rewrite Equation 2.16 for any distinct number of frequency channels N_ν along a given line of sight (LoS) $\hat{\mathbf{n}}$ as:

$$\mathbf{x} = \hat{\mathbf{A}} \cdot \mathbf{s} + \mathbf{r} \tag{2.17}$$

where $\mathbf{x}$ represents $T(\nu,\hat{\mathbf{n}})$ the total signal received, $r_i = T_{cosmo}(\nu,\hat{\mathbf{n}}) + T_{noise}(\nu,\hat{\mathbf{n}})$, $\hat{\mathbf{A}} = f_k(\nu)$ and $\mathbf{s} = S_k(\hat{\mathbf{n}})$. In order to recover the cosmological signal and noise term: $\mathbf{r} = \mathbf{x} - \hat{\mathbf{A}} \cdot \mathbf{s}$, one needs to be able to accurately determine $\hat{\mathbf{A}}$ and $\mathbf{s}$, this is the aim of the blind foreground subtraction algorithm.

Principle Component Analysis (PCA)

The PCA is, a blind foreground cleaning method which takes advantage of the main characteristics of foregrounds, i.e: their smooth frequency behavior and large amplitude. This helps in finding the foreground components s_k and the basis functions A_{ik} simultaneously. The algorithm was invented by Karl Pearson [61] with the idea of being a dimension-reduction tool capable of preserving the information when moving from a larger number of variables to a smaller number.

The algorithm can be summarized into three steps:

- The frequency covariance matrix is first calculated by averaging over all pixels that we have access to:

$$C_{ij} = \frac{1}{N_{pix}} \sum_{n=1}^{N_{pix}} T(\nu_i,\hat{\mathbf{n}}_n) T(\nu_j,\hat{\mathbf{n}}_n) \tag{2.18}$$

  where N_{pix} is the number of pixels, $T(\nu_i,\hat{\mathbf{n}})$ is the brightness temperature for a given frequency channel ν_i

- We diagonalize the covariance matrix:

$$\hat{\mathbf{U}}^T \hat{\mathbf{C}} \hat{\mathbf{U}} = \Lambda \equiv diag(\lambda_1 ... \lambda_{N_v}) \tag{2.19}$$

where the eigenvalues $\lambda_i > \lambda_{i+1}$ are the eigenvalues for $\hat{\mathbf{C}}$ and $\hat{\mathbf{U}}$ is the orthogonal matrix with columns that represent the respective eigenvectors.

- Next we select the number of eigenvalues (n_{fg}) which corresponds to the number of foregrounds. A matrix is then constructed from these eigenvalues which models the brightness temperature for each LoS as

$$\mathbf{x} = \hat{\mathbf{U}}_{\rm fg}\mathbf{s} + \mathbf{r} \tag{2.20}$$

Projecting $\mathbf{x}$ on to the eigenvector basis of $\hat{\mathbf{C}}$ we can get the foreground maps. From here we get

$$\mathbf{s} = \hat{\mathbf{U}}_{\rm fg}^{\mathbf{T}}\mathbf{x} \tag{2.21}$$

Recall that $\hat{\mathbf{U}}$ is an orthogonal matrix, therefor $\hat{\mathbf{U}}_{\rm fg}^T \hat{\mathbf{U}}_{\rm fg} = 1$ is true and that Eq. 2.17 and Eq. 2.20 must coincide. We can rewrite Eq. 2.21 to have the form of $s = (\hat{\mathbf{U}}_{\rm fg}^T \hat{\mathbf{N}}^{-1} \hat{\mathbf{U}}_{\rm fg})^{-1} \hat{\mathbf{U}}_{\rm fg}^T \hat{\mathbf{N}}^{-1} \mathbf{x}$, where $\hat{\mathbf{N}}$ is the covariance matrix of $\mathbf{r}$ (cosmological signal and noise). Including $\hat{\mathbf{N}}$ requires weighting to be placed on the temperature maps of Eq. 2.18 ($T(\nu_i, \hat{\mathbf{n}}_n)/\sigma_i$). The standard deviation of combining both components (noise and cosmological signal) is given as $\sigma_i = \sqrt{\sigma^2_{cosmo,i} + \sigma_{noise,i}}$. In order to obtain the de-weighted maps we multiply σ_i before the projection of $\mathbf{x}$ which introduces $\mathbf{N}^{-1}$ thus:

$$\mathbf{s} = \hat{\mathbf{U}}_{\rm fg}^{\mathbf{T}}\mathbf{x} \equiv (\hat{\mathbf{U}}_{\rm fg}^T \hat{\mathbf{N}}^{-1} \hat{\mathbf{U}}_{\rm fg})^{-1} \hat{\mathbf{U}}_{\rm fg}^T \hat{\mathbf{N}}^{-1} \mathbf{x}. \tag{2.22}$$

2.6 Simulating Pipeline

In order to produce HI intensity maps we made use of the publicly available code called *CRIME* (Cosmological Realizations for Intensity Mapping Experiments) [13] [60, 62]. *CRIME* comprises of three codes:

- *GetHI*- produces the HI maps,
- *ForGet*- produces various foreground maps,

[13] https://github.com/damonge/CRIME

- *JoinT*- combines the HI intensity maps with foreground, introduces white Gaussian noise and convolves with a telescope beam.

These three codes are summarized in the different sections below: HI maps (GetHI), Foreground maps (ForGet) and Observation maps (JoinT); which is taken from [62].

HI maps

The cosmological signal is produced within a box of equal side lengths L (co-moving size), the box is then divided into smaller cells of equal size l_c. The amount of cells contained within the box is given by the N_{grid} parameter and the size $l_c \equiv L/N_{grid}$. In Fourier-space Gaussian random numbers are produced with a variance

$$\sigma^2(k) \equiv \left(\frac{L}{2\pi}\right)^3 P_0(k), \tag{2.23}$$

where $P_0(k)$ is the initial matter power spectrum given at z=0. The velocity potential is then calculated for the same redshift value

$$\psi_{\mathbf{k}}(z=0) = f_0 H_0 \frac{\delta_{\mathbf{k}}(z=0)}{k^2}. \tag{2.24}$$

This is done for a grid with $\mathbf{k} = \mathbf{n}2\pi/L$. Next the Gaussian overdensity δ_G and the radial velocity v_r are calculated for the same redshift by taking the Fast Fourier Transform and moving into configuration space. Using the *radial Co-moving distance* $\chi(z)$, the δ_G and v_r fields can evolve to any redshift inside the lightcone. A lognormal transformation[62] is done on δ_G, therefore if $\mathbf{x}$ defines a cells' position and z($\mathbf{x}$) the redshift, the overdensity and radial velocity are given as:

$$1 + \delta_{HI}(\mathbf{x}) = exp\left[G(z)\delta_G(\mathbf{x}, z=0) - G^2(z)\sigma_G^2/2\right] \tag{2.25}$$

$$v_r(\mathbf{x}) = \frac{f(z)H(z)D(z)}{(1+z)f_0 H_0} v_r(\mathbf{x}, z=0) \tag{2.26}$$

with $\sigma_G^2 \equiv \langle \delta_G^2 \rangle$ and $G(z) \equiv D(z)b(z)$ a combination of the growth factor as well as the galaxy bias. The total mass for each cell is then calculated by

$$M_{HI} = (2.775 \times 10^{11} M_\odot)\left(\frac{l_c}{Mpc/h}\right)^3 \frac{\Omega_b x_{HI}(z)}{h}(1+\delta_{HI}). \tag{2.27}$$

Where $M_\odot$ is solar mass, Ω_b is the density of baryons, x_{HI} is neutral fraction which is modeled as $x_{HI} = 0.008(1+z)$[62] and $\delta_{HI} = b_{HI}\delta_{cdm}$ is the HI density contrast.

Lastly the box is sliced into spherical shells and pixelized to produce the 21cm brightness temperature images for different frequency bins.

For the simulation we ran *GetHI* with the following parameters shown in Table 2.1:

Parameters	
N_{side}	256
N_{grid}	2048
Smoothing scale (r_{smooth}[Mpc/h])	2

Table 2.1 Parameters for the cosmological box.

N_{side} is the resolution parameter corresponding to the pixel resolution (θ_{pix}) of the Healpix map, we have used an $n_{side} = 256$ which correspond to a $\theta_{pix} = 0.22^{\circ}$. The initial linear matter power spectrum (P_0), needed to generate the matter field is produced with pyCAMB [14] (python version of CAMB[15]) [54] . The *Planck 2016* cosmological parameters [51, 63] are applied in the simulation :

Parameters	Ω_M	Ω_b	h	n_s	σ_8
Fiducial values	0.315	0.049	0.67	0.96	0.83

Table 2.2 Cosmological parameters [51, 63]

The number of HI IM maps produced in the simulation is 318, each map corresponded to a frequency (ν) bin with a width in co-moving distance of 5Mpc/h. We created a ν-list that spanned the MeerKAT L-band frequency range (900 MHz $> \nu >$ 1300 MHz). To ensure that the L-band was being effectively produced by the simulation we gave a larger ν-list which ranged from 836.75 MHz $< \nu <$ 1352.76 MHz. We used the following equation in order to create the ν-list:

$$\Delta z = \frac{H(z_i)}{c} \times 5\text{Mpc/h}, \tag{2.28}$$

Δz is the redshift interval, $H(z_i)$ is the Hubble parameter at the initial redshift, this parameter changes as the redshift increases and c is the speed of light.

[14] https://camb.readthedocs.io/en/latest/
[15] https://camb.info/

In summary our simulation generates a box with length size of $l_{\text{box}} = 3477.68$ Mpc/h. A grid size (n_{grid}) of 2048 is chosen, which grids the simulation box into 2048 cells. Each cell has a length size of $l_{\text{box}}/n_{\text{grid}} \approx 1.7$ Mpc/h. The simulation box is then pre-smoothed with a smoothing scale of 2 Mpc/h, this ensures that the sky maps are realistic and contains no extreme pixels (sources). Each map is a Healpix map with a pixel resolution of $\theta = 0.22^{\circ}$, this corresponds to $n_{\text{side}} = 256$. These parameters are listed below in Table 2.3

Generating these maps was accomplished through the Centre for High Performance Computing (CHPC [16]) Lengau cluster machine located in Cape Town, South Africa. We used a single node of 24 cores in parallel which took approximately 7 minutes of run time and a disk space of approximately 0.9 gigabytes.

Parameters	Values
z	$0.05 > z > 0.698$
ν (MHz)	$836.75 > \nu > 1352.76$
Maps	318
n_{grid}	2048
n_{side}	256
$r_{\text{smoothing}}$	2 Mpc/h

Table 2.3 Parameters used to obtain the Simulation box.

Foreground maps

Four types of Foregrounds are included in the simulation: galactic synchrotron, galactic free-free, extra-galactic free-free and point sources. These foregrounds are integrated into the simulation, with the following power spectrum model [64]:

$$C_l(\nu_1, \nu_2) = A\left(\frac{l_{ref}}{l}^2\right)^{\beta}\left(\frac{\nu_{ref}^2}{\nu_1\nu_2}\right)^{\alpha} exp\left(-\frac{log^2(\nu_1/\nu_2)}{2\xi^2}\right). \tag{2.29}$$

The parameters in Table 2.4 represent the behavior of the power spectrum model (Eq. 2.29). A refers to the amplitude, β controls the distribution of foregrounds on angular scales,α denotes the mean frequency dependence for the foregrounds and ξ is the correlation length in frequency space. If $\xi \to 0$ then this implies that the foregrounds are ideally correlated

[16]https://www.chpc.ac.za/

Foreground	A (mk^2)	β	α	ξ
Galactic synchrotron	700	2.4	2.80	4.0
Galactic free-free	0.088	3.0	2.15	35
Extragalactic free-free	0.014	1.0	2.10	35
Point sources	57	1.1	2.07	1.0

Table 2.4 Foreground model parameters for Equation 2.29 taken from [64].

which corresponds to a trivial foreground cleaning case [44]. The foreground maps are generated with the *ForGet* command in *CRIME*. Shown below in Figure 2.8, 2.9, 2.10 and 2.11 are the Healpix maps for the various foregrounds. It should be pointed out that the foreground temperature fluctuates, resulting in negative temperatures on the colour bar. The largest of these foreground contaminants comes from our own galactic centre, this is known as synchrotron emission. The effect is simulated by taking the Haslam 408 MHz All Sky Map [65] and projecting it to different frequency bins using the spectral index. However due to the poor resolution of the Haslam map, the code applies Eq. 2.29 to generate Gaussian realizations which model the synchrotron emission on small scales.

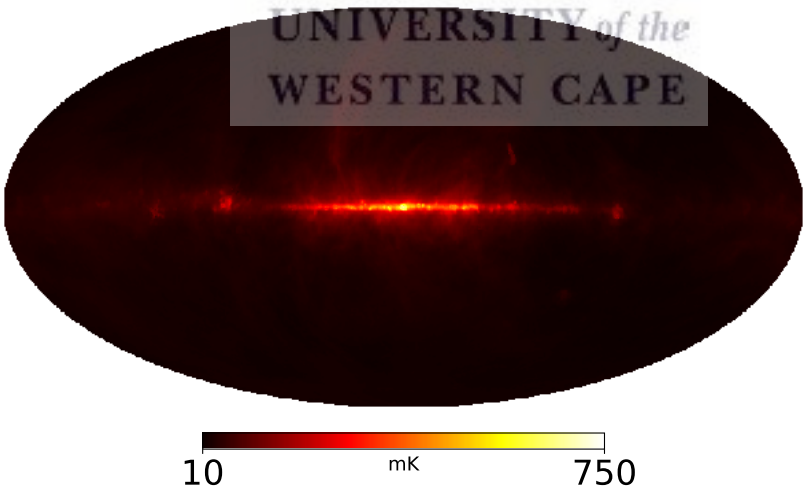


Fig. 2.7 Haslam 408MHz Galactic Synchrotron All Sky Map. The colour bar shows the temperature range in mk.

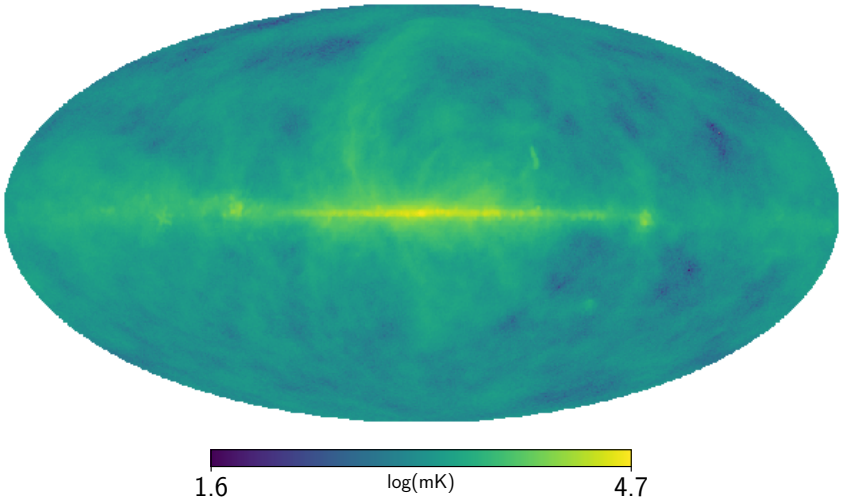


Fig. 2.8 Example of the Galactic synchrotron emission created in the simulation at $\nu =$ 1011.24 MHz – 1012.74 MHz.

Free-free emission are simulated for galactic and extra-galactic cases by respectively combining the information in Table (2.4) with Equation (2.29), producing Gaussian realizations.

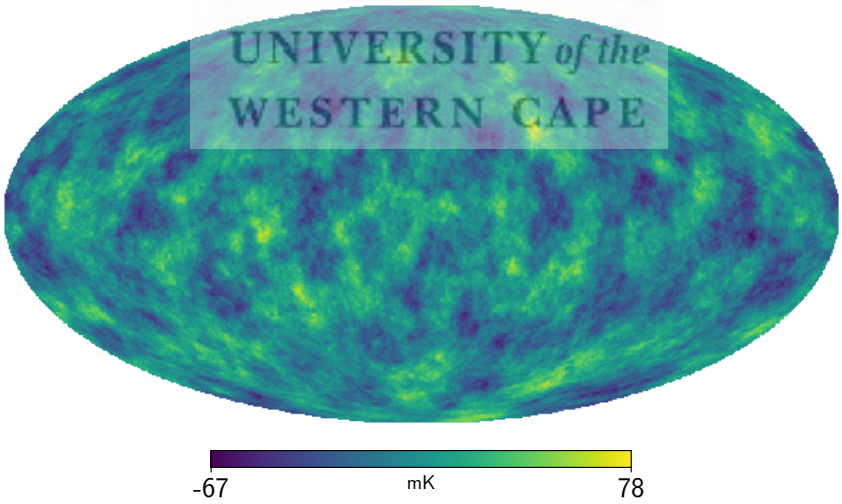


Fig. 2.9 Example of the Galactic free-free emission foreground created in the simulation at $\nu = 1011.24 - 1012.74$MHz.

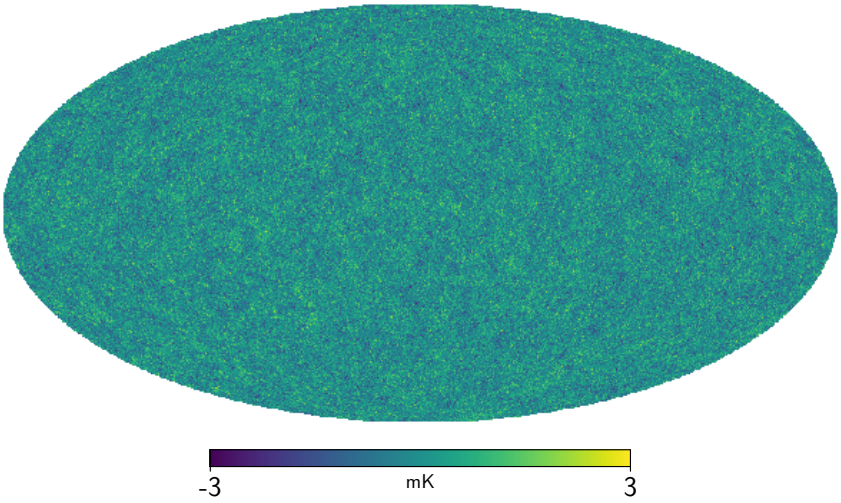


Fig. 2.10 Example of the Extra-galactic free-free foreground created in the simulation at $\nu = 1011.24 - 1012.74$MHz.

Point source maps are created by applying Eq 2.29 with Table 2.4, a similar process as with free-free emission.

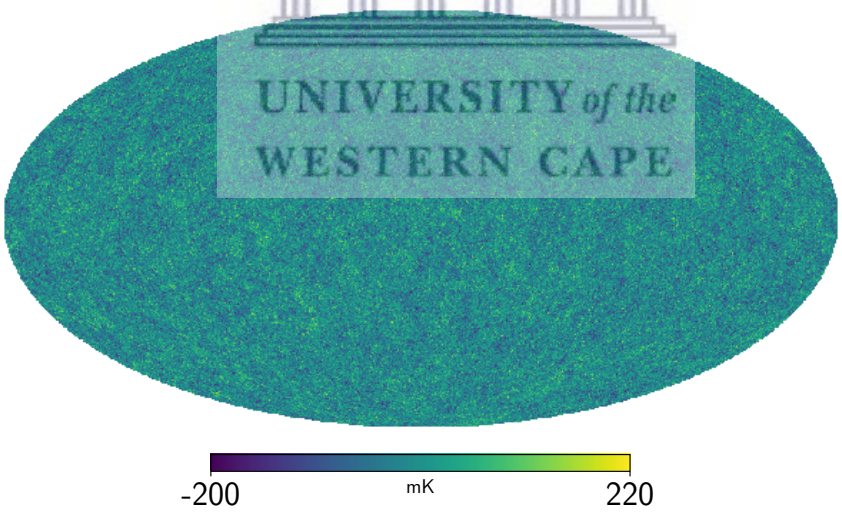


Fig. 2.11 Example of Point sources created in the simulation at $\nu = 1011.24$ MHz $- 1012.74$ MHz.

Mask

Various sky masks are applied in the simulation pipeline, initially we applied a mask which covered the MeerKLASS sky area [38] (Figure 2.12). We obtained the Healpix survey area map from [66]. We then proceeded to apply masks for different sky fractions to the pipeline as well.

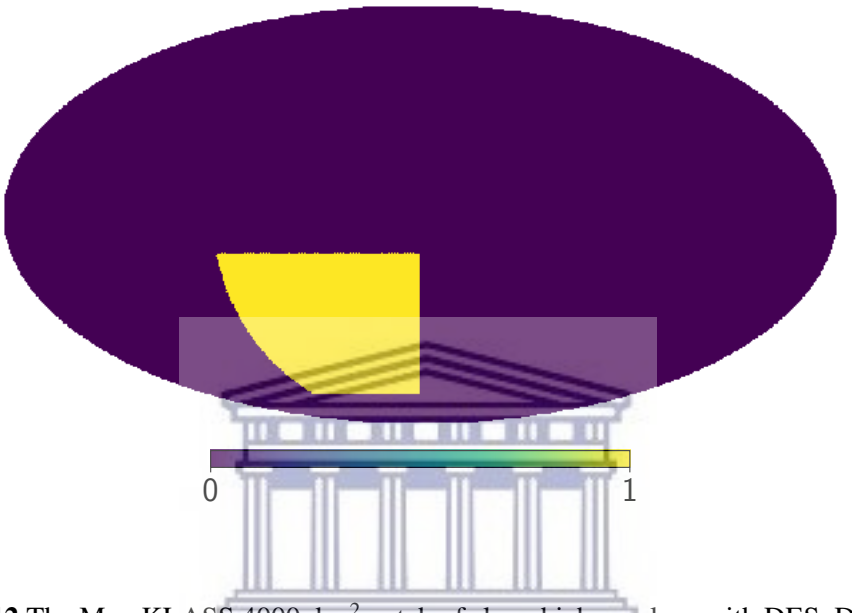


Fig. 2.12 The MeerKLASS 4000 deg^2 patch of sky which overlaps with DES. DES is the Dark Energy Survey, which is a photometric optical galaxy survey. The 1s represent the window of observation.The patch corresponds to a $f_{sky} = 0.088$.

Instrument

The instrumental parameters of MeerKAT contribute to both the systematic noise and the telescope beam that the instrument induces on to the sky. The addition of these two factors, contribute towards a realistic IM experiment in single dish mode (auto-correlation) [36, 44, 57]. Noise maps are generated from Gaussian realizations with a rms described by

$$\sigma_N^2(\nu) = T_{sys}^2(\nu)\frac{4\pi f_{sky}}{N_{dish}t_{tot}\Delta\nu}. \tag{2.30}$$

The system temperature is defined as $T_{sys} \approx T_{inst} + (60K)/(\nu/300MHz)^{-2.5}$ [62] and $T_{inst} = 20K$ [37]. t_{tot} is the total time of the observation; δ_ν is the frequency width and f_{sky} is the sky fraction of how our observation.

A Gaussian beam is also applied to each map following

$$\theta_{FWHM} \sim \frac{\lambda}{D_{dish}}, \tag{2.31}$$

where D_{dish} is the diameter of the telescope antenna.

In order to produce the final observed temperature maps, the *JoinT* command was applied with the following two parameters:

Parameters	
D_{dish}	13.5m
N_{dish}	64
T_{sys}	20K

Table 2.5 Instrumental parameters applied in the simulation.

Observational Maps

The observational maps which are produced in *CRIME* represent the final intensity sky maps, these maps are a combination of cosmological HI signal, foregrounds, systematic noise and the telescope beam effects. *CRIME* is able to produce sky maps of different combination from the list of inputs. Once the combination is chosen, the maps are combined into the total signal. These maps are then degraded from a pixel resolution $\theta_{pix} \approx 0.22 \rightarrow 0.92$ which corresponds to an $n_{side} = 256 \rightarrow 64$ during the *JoinT* execution.

Foreground Cleaning

The foreground cleaning method is done through *fg_rm* [17] [60] code. We select the number of foreground modes (n_{fg}) based on the work done in [44, 66]. We removed 5,6,7 and 8 modes of foregrounds; generally one sets n_{fg} as the same amount of foregrounds given to the simulation. We do not only recover the cleaned maps but also the radial 21cm power spectrum as well. Since we used a simulation box that is larger than L-band, we can retrieve the power spectrum for the entire band without concerning ourselves with boundary conditions which is introduced by foreground cleaning. We also do not look at $z < 0.2$, due to concern over RFI. Therefore we extract the radial power spectrum in the redshift range of, $0.2 < z < 0.58$.

[17] https://github.com/damonge/fg_rm

2.7 Summary

In this chapter we reviewed neutral hydrogen intensity mapping and showed the 21cm power spectrum. We looked at a few experiments which would probe the 21cm signal. We discussed the effects that the telescope beam for MeerKAT has on the BAO wiggles in regards to measuring the 21cm power spectrum. We highlighted the issue that foreground contamination as with HI IM experiments as well as a means to remove them. Lastly we explain the simulation pipeline which we used and developed in order to create a mock HI IM experiment for MeerKAT.

In figure 2.13 which shows the pipeline that has been created with the *CRIME* simulation. We initially start with GetHI-producing HI maps and from there we can go into any direction or do all three in a single simulation run. We are also able to loop this process as many times as required while being able to vary the seed parameters where need be, for instance: GetHI-HI map realizations, ForGet-foreground realizations and JoinT-noise realizations. The variation of the seed parameters ensures we have different universes, the seed parameters can be randomized or ordered.

The observational maps which are produced in *CRIME* represent the final intensity sky maps, these maps are a combination of cosmological HI signal, foregrounds, systematic noise and the telescope beam effects. *CRIME* is able to produce sky maps of different combinations from the list of inputs. In Figure 2.14, we show the various intensity maps extracted at different points in the simulation for the proposed MeerKAT HI IM survey area. Each map has temperature in units of mK. The *top left* is the HI signal only, *top right* is the total foreground contribution, in the simulation. The next four plots from *left to right* and *top to bottom* show the number of foreground modes removed.

Pipeline

This pipeline is essentially combing the individual codes for *CRIME* and *fg_rm*, we are then able to repeat the process n number of the repetitions. Depending on the end result, the pipeline can reproduce various HI radial power spectrums based on the users request. This can be seen in Foreground removal boxes shown in Figure 2.13.

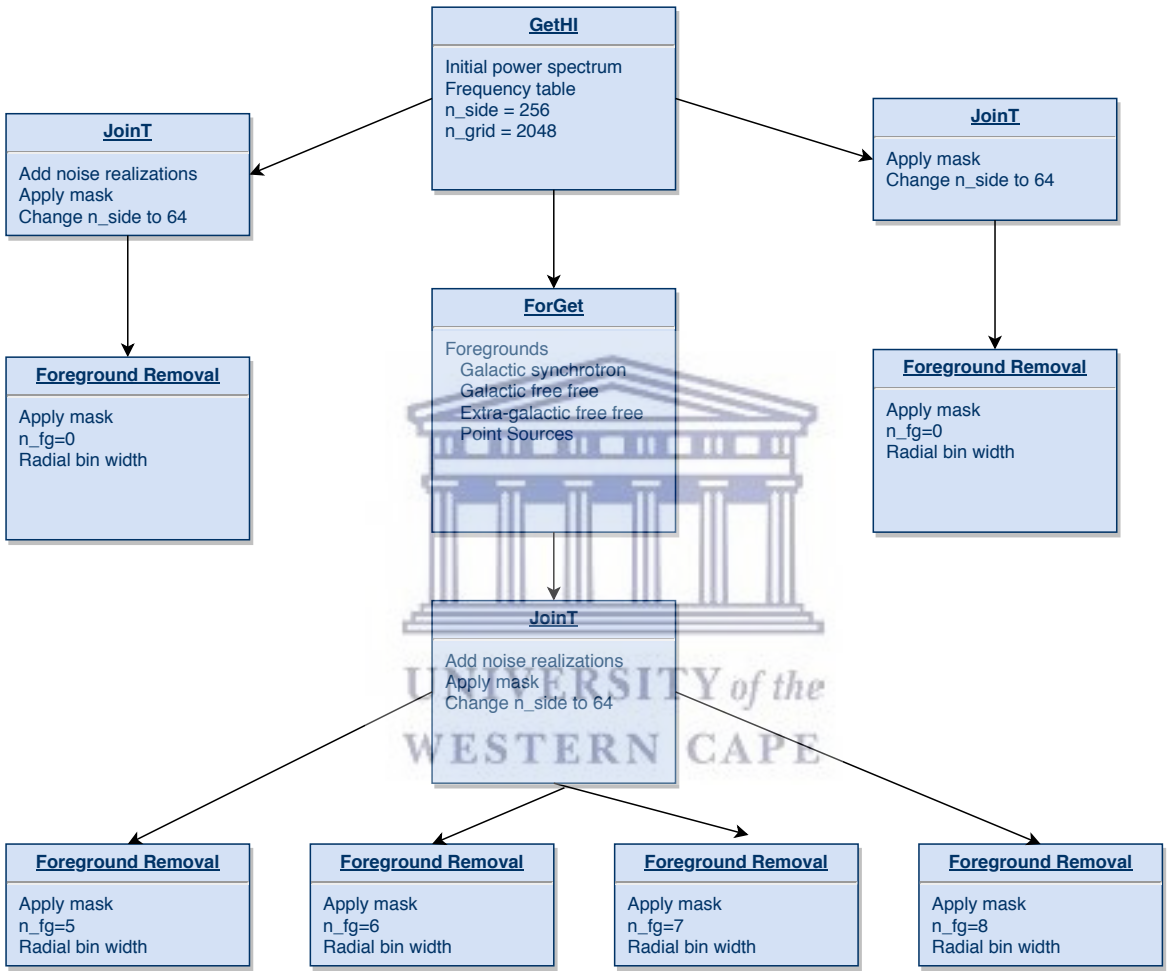


Fig. 2.13 *CRIME* simulation pipeline created to produce multiple realizations for a single experiment.

Simulation Outputs

Figure 2.14 shows the results of CS as well as the foreground residuals for the frequency bin of $900.2 \leq \nu \leq 901.61$MHz which corresponds to a redshift bin $0.575 \leq z \leq 0.578$.

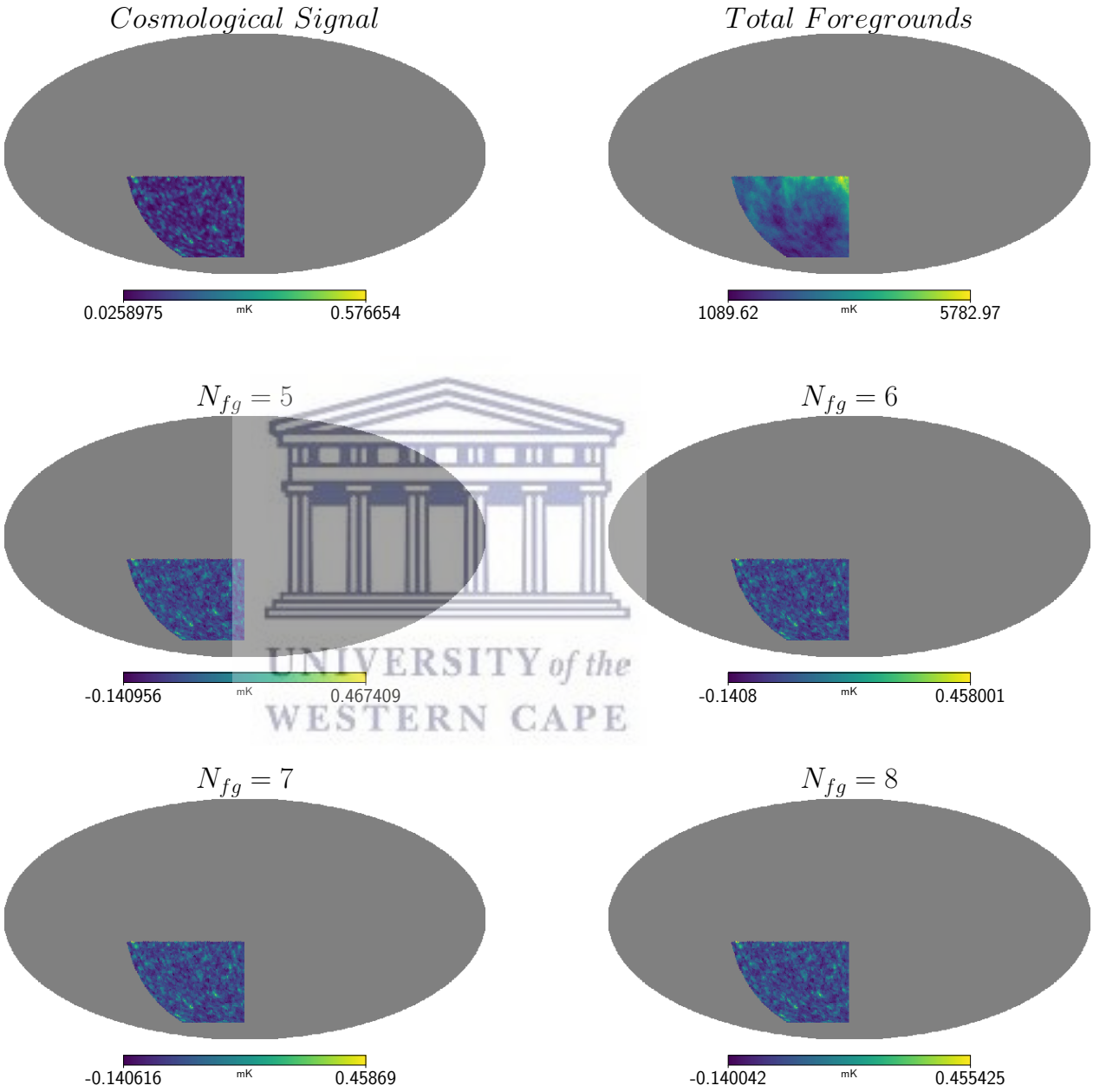


Fig. 2.14 Intensity maps at a $900.2 \leq \nu \leq 901.61$MHz, which correspond to a redshift of $0.575 \leq z \leq 0.578$ are shown in galactic coordinates, for the proposed survey in MeerKLASS [38].

Chapter 3

Extracting the BAO along the line-of-sight

In this section we explain the procedure undertaken to analyze the radial 21cm power spectrum, for the different cases: (a) Cosmological HI signal, (b) Cosmological HI signal plus systematic noise (c) Cosmological HI signal plus systematic noise plus foregrounds that we have simulated. Initially we investigate a proposed MeerKAT Intensity Mapping survey which would cover a sky area (A_{sky}) of 4000 deg^2 over a period of 4000 hours. This survey will be known as *Survey M* throughout this chapter. We explain the approach in dealing with cosmic variance as well as our method of extraction used to detect the BAO wiggle features. We test the signal-to-noise ratio for the BAO wiggle feature for the various cases. We also look at how varying both the sky area or observational time affects the detection of the BAO features.

3.1 Methodology

Testing Cosmic Variance

Cosmic variance is the statistical uncertainty one gets by doing observations on the Universe at large scales, since we can only achieve one realization of the observable Universe. In regards to the simulations, we test the effects of cosmic variance by running a suite of 100 simulations. Each simulation in a suite is unique due to the value of the *seed* parameter which is set to match the current Unix time stamp[1], which is the number of seconds past from the Unix epoch. From Figure 3.1 we see the effects that the MeerKAT telescope beam (top, middle and bottom) has on the measured value of the cosmological HI signal.

[1] https://www.unixtimestamp.com/

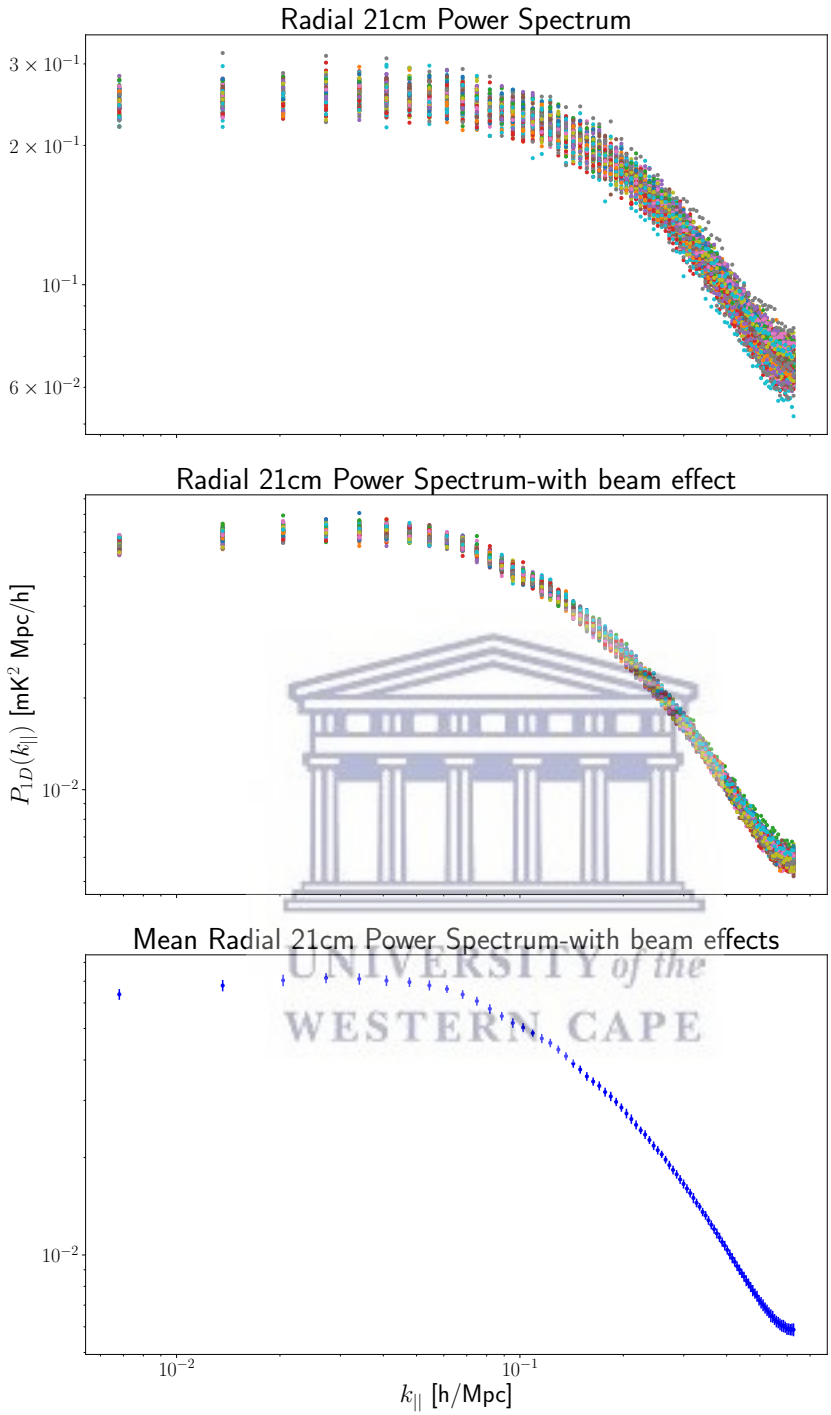


Fig. 3.1 A suite of 100 realizations, each colour represents a different simulation: *Top*: cosmological HI signal only; *Middle*: with MeerKAT Gaussian beam included; *Bottom*: the average signal with Gaussian beam and error bars. The redshift ranges from $0.2 \leq z \leq 0.58$.

From the 100 simulations, we determine the mean power spectrum and the root-mean-square (rms) of the mean as well. If we let $\bar{P}(k_{||})$ represent the mean power spectrum then the rms of the power spectrum ($\sigma(P)$) can be represented by:

$$\sigma(P(k_{||})) = \sqrt{\frac{\sum_{i=1}^{N}(\bar{P}(k_{||}) - P_i(k_{||}))^2}{N}}, \tag{3.1}$$

where $N = 100$, the number of simulations used. The signal-to-noise ratio is calculated with the following equation:

$$S/N(k_{||}) = \frac{\bar{P}(k_{||})}{\sigma(P(k_{||}))}. \tag{3.2}$$

In Figure 3.2 we see the results of the signal-to-noise for the cosmological HI signal only. Around the BAO scales ($k \approx 0.02$ h/Mpc $-$ 0.45 h/Mpc) the S/N is increasing and has a larger ratio, resulting in a favourable signal detection of the BAO.

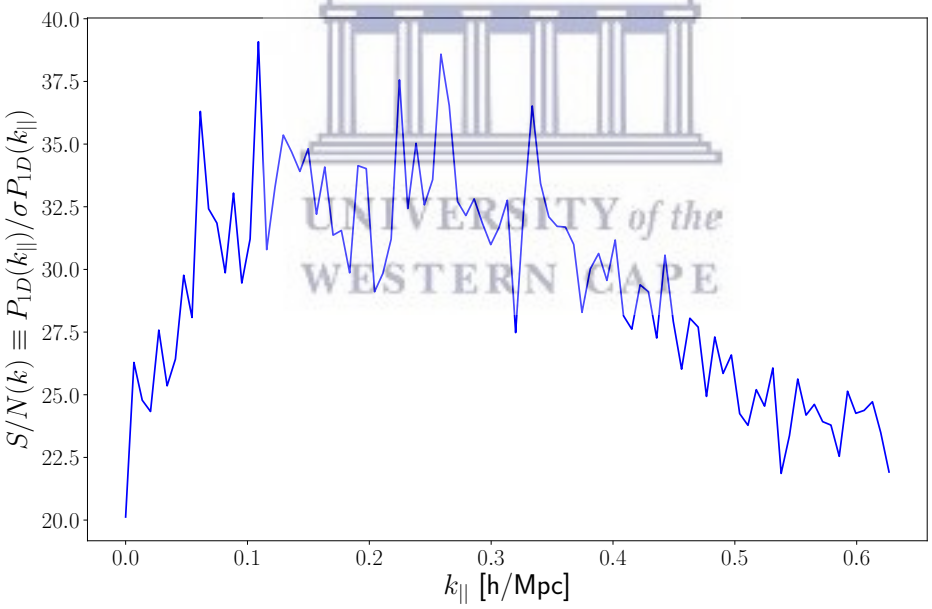


Fig. 3.2 signal-to-noise ratio for the cosmological HI signal.

Detecting the BAO signature

In order to extract the BAO signature from the power spectrum we fit a smooth component to the 21cm radial power spectrum. We assumed that the line-of-sight cosmological power spectrum can be described by a smooth component and a BAO (wiggle) component,

$$P_{total} = P_{smooth}(1 + w_{BAO}) + P_N + P_f. \tag{3.3}$$

Where P_{smooth} is the smooth component of the power spectrum, w_{BAO} is the BAO wiggle feature [14], P_N is the noise power spectrum originating from the instrumental effects and P_f is the residual (post cleaning) foreground power spectrum. In our assumption we assume that the smooth component the noise component and the residual cleaning can be described by a 7^{th} order polynomial ($P_{smooth} + P_N + P_f \equiv P_{polynomial}$), which is fitted in linear (not log) space across the k wavenumber. We chose 7^{th} order since it was the lowest order value that could visually extract the wiggle component. After subtraction we retrieve

$$w_{BAO} \times P_{smooth} = P_{total} - P_{polynomial}, \tag{3.4}$$

which we believe is a good estimator for the BAO wiggles,since the position of the wiggles are of importance in cosmology. We test later in this chapter the effectiveness of the smooth component in retrieving $w_{BAO} \times P_{smooth}$, for the sake of shorthand from here on out we will call

$$BAO_{wiggles} = w_{BAO} \times P_{smooth}. \tag{3.5}$$

When fitting the polynomial on P_{total}, we initially fit over a large $k_{||}$ range from $0.02 < k_{||} < 0.45$ following [36]. This fit extends further than the BAO range; the true fit for the BAO wiggles is in the range of $0.04 < k_{||} < 0.25$. This can be seen in Figure 3.3 for the HI signal only. Each radial power spectrum (shown as different coloured dots in Figure 3.1) are fitted and subtracted by a smooth component unique to that power spectrum. This process is done individually to each power spectrum. The residuals are then averaged together and we obtain the *top* plot of Figure 3.4.

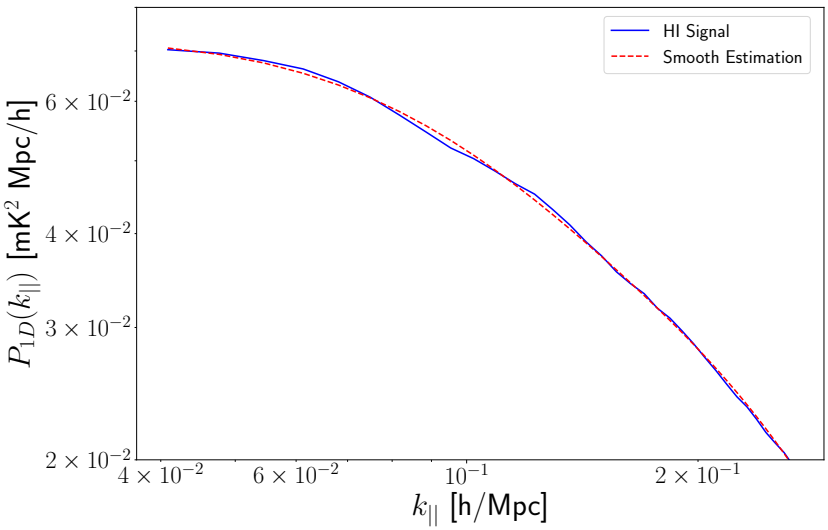


Fig. 3.3 The smooth $P(k_{||})$ component shown in the red dashed line is fitted over the radial power spectrum for the cosmological HI signal. The fit for the smooth component covers the BAO wiggles.

After subtracting the smooth component and averaging the suite of simulations, the total signal-to-noise can now be determined as shown in Figure 3.4 with the following equation:

$$(S/N)_{\text{total}} = \sqrt{\sum_{i=1}^{n}((\bar{P}_i/\sigma(P_i))^2)}. \tag{3.6}$$

Keeping true with our previous assumption (Eq 3.3), we assumed that we can retrieve the true BAO signal (multiplied by the smooth component). An indication of the BAO signal (and the smooth component) can be seen as fluctuations around zero in Figure 3.4 (*bottom*). Applying Eq. 3.6 for the HI signal only; we obtain a $(S/N)_{\text{total}} = 2.121$, resulting in a detection of the BAO.

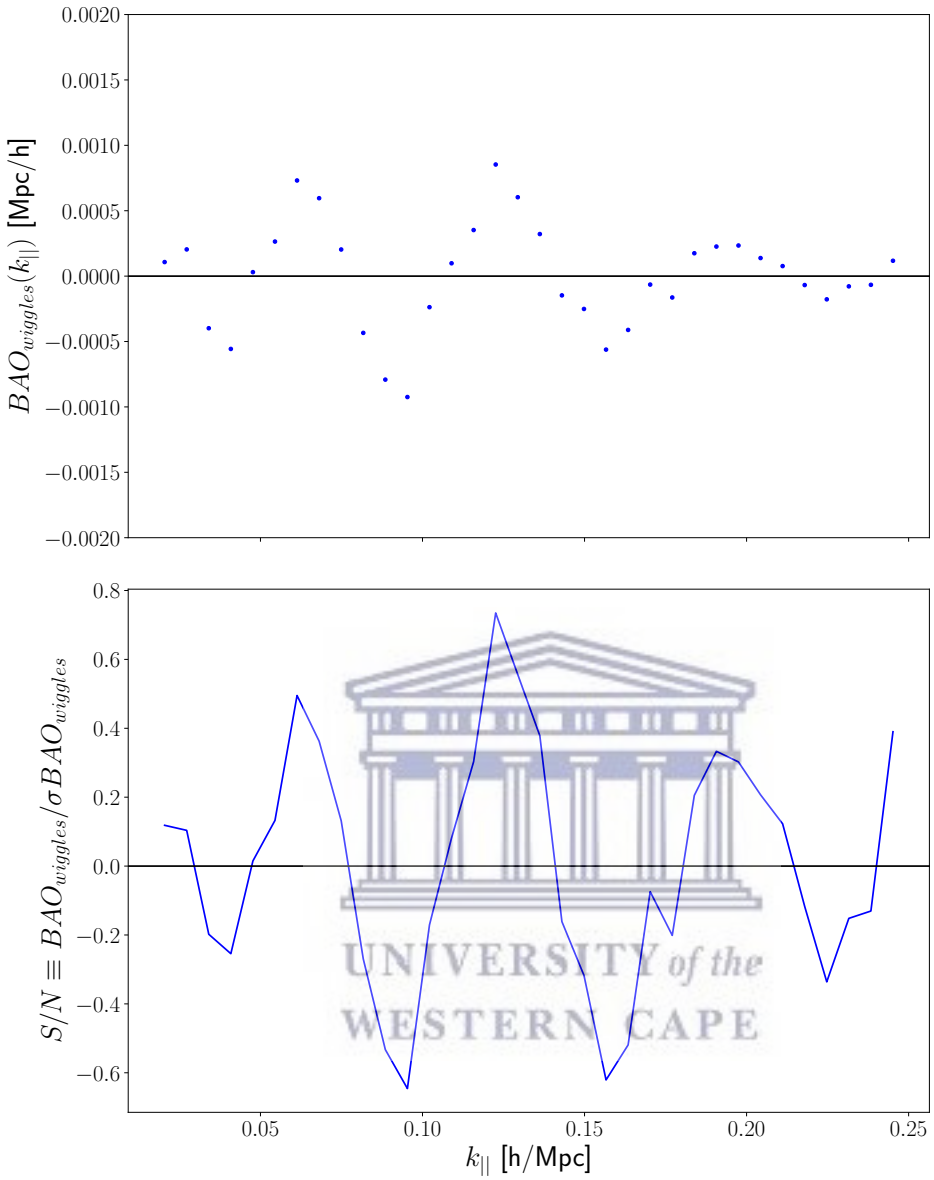


Fig. 3.4 The averaged BAO wiggles can be visually seen after the subtraction of the smooth component (top) and the respective averaged S/N plot for the wiggles (bottom)

3.2 Realistic Sky

We now consider a more realistic scenario with the inclusion of noise and foregrounds on top of the cosmological signal and test the BAO extraction. The simulation now produces random white Gaussian noise maps for each frequency bin, see Chapter 2 for instrumental effects. Each realization will have a noise variation that is different to the next realization. This is due to the different seed values corresponding to Unix time.

From here on we adopt the following naming convention; the cosmological HI signal is referred to as *CS*. The addition of systematic noise is referred to as *CSN*. The addition of cleaned foreground maps is given as *CSNF(n_{fg})*, where n_{fg} refers to the amount of foreground modes removed.

Foreground maps are constructed from the first realization of the suite of 100 simulations, therefore each realization has the same foreground map per frequency as the previous one. This was done under the assumption that the foregrounds do not vary greatly from one realization to the next. In order to remove the foreground contaminants, we employ the PCA cleaning method and vary the foreground degree of freedom (n_{fg}) from 5 to 8. In Figure 3.5, we observe that after the application of foreground removal, there is an overlap between the foreground clean power spectra and the power spectrum containing only the cosmological HI signal plus noise. The power spectra show that the least aggressive cleaning option, i.e; n_{fg}=5, is able to effectively remove the contaminants and increasing this value does not show any substantial improvement in the results.

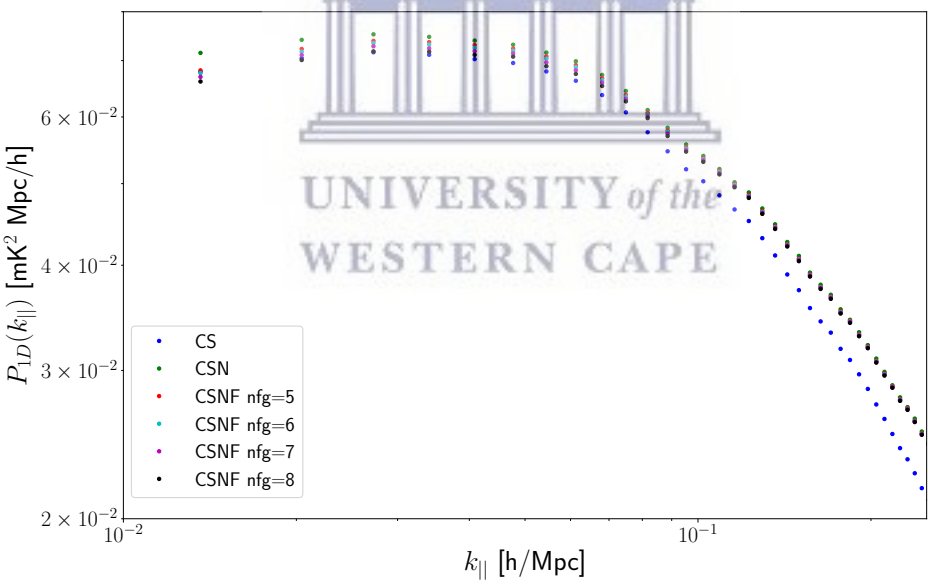


Fig. 3.5 100 averaged radial power spectra for the various cases of Survey M. The power spectra for the HI signal only is shown in blue and green represents the HI signal with systematic noise. The remaining power spectra represent the HI signal with noise and foreground, the number of foreground modes removed (5,6,7,8) are given by (red, cyan, magenta, black) respectively.

The next step is to fit the 7th-order polynomial to each power spectrum in order to obtain the respective smooth components (follows the same method as in the previous section). We then

individually subtract each smooth component from their respective wiggle counterpart (for all 100 simulations) leaving the oscillating residuals behind. Figure 3.6 shows the average residuals, i.e the average wiggles. After this step we can now calculate the mean S/N for the $\mathrm{BAO_{wiggles}}$, this is defined similarly to Eq. 3.6,

$$S/N_{\mathrm{BAO_{wiggles}}}(k_{||}) = \frac{\mathrm{BAO_{wiggles}}(k_{||})}{\sigma(\mathrm{BAO_{wiggles}}(k_{||}))}. \tag{3.7}$$

Where $\sigma(BAO_{\mathrm{wiggles}}(k_{||}))$ is the rms determined after the estimation subtraction from the suite of realizations.

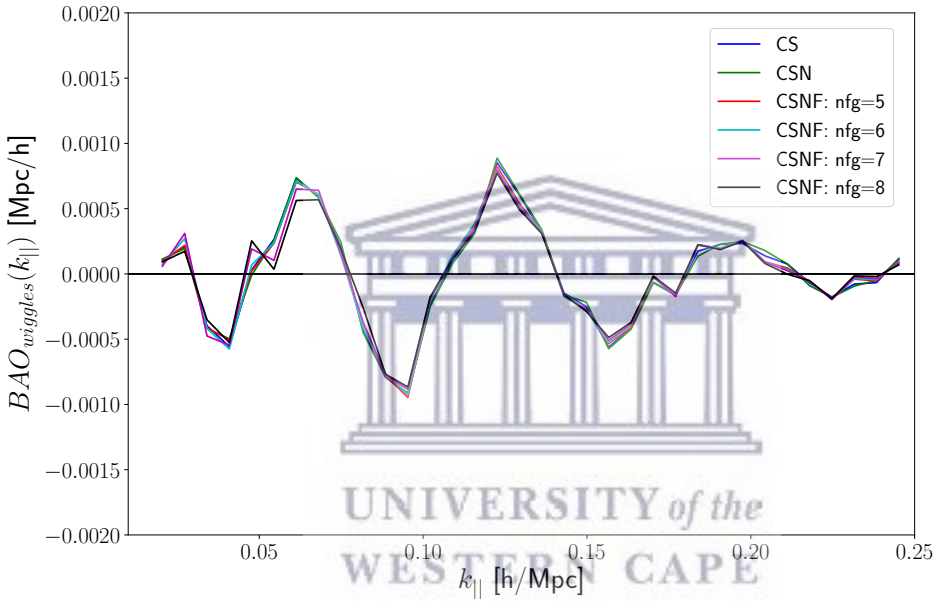


Fig. 3.6 The average wiggles after subtracting the smooth component individually from each radial power spectrum for the different cases of Survey M as shown in Figure 3.3. We follow the same legend as the previous plots.

The S/N for the BAO feature shows a general overlap with each case and also expresses noise as being sub-dominant over cosmic variance. The different cosmological signal plus noise plus foregrounds $\mathrm{n_{fg}}$ closely overlaps with that of the cosmological signals and the cosmological signal plus noise cases. This implies that for the survey of 4000 deg^2 and 4000 hours we would be able to detect a signal-to-noise ratio that is similar when considering the presence of foregrounds, unlike when we ignore foregrounds. Table 3.1 shows the total signal-to-noise ratio for Survey M, which can also be seen in Figure 3.7.

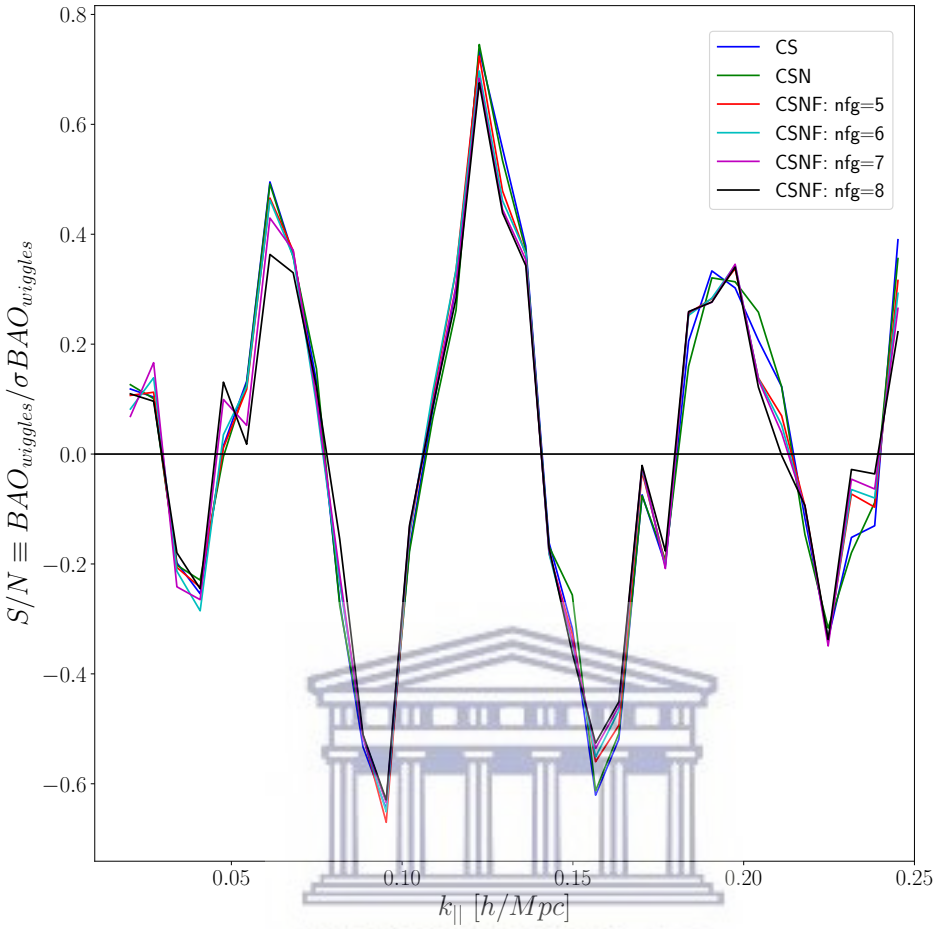


Fig. 3.7 The average S/N of the wiggles for the various cases. We follow the same legend as the previous plots

Case	CS	CSN	CSNF5	CSNF6	CSNF7	CSNF8
S/N	2.121	2.056	2.102	2.079	2.017	1.957

Table 3.1 Total S/N for Survey M

Table 3.1 shows that applying 5 degrees of freedom when removing foregrounds is sufficient in retaining the BAO signal. Moving forward, we will only consider the least aggressive cleaning mode, $n_{fg} = 5$.

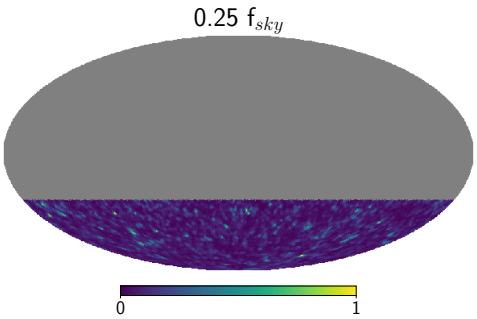
3.3 Dependence on observation design

In the previous section, we showed that one can approximately achieve the same S/N values for the various cases (CS,CSN and CSNF) of Survey M (see Table 3.1). However we wanted to investigate the conditions under which the different cases would have distinct S/N values. The conditions that we could vary to test this would be the sky area and the observational time. In terms of the sky area we compare the results for $f_{sky} \approx [0.1, 0.2, 0.25 \text{ and } 0.5] \approx [4000, 8000, 10000, 20000]$ deg^2 of the total sky. The total observational time varied from 1000, 2000 and 4000 hours. Based on results of the previous section with regards to the number of foregrounds removed, we set $n_{fg} = 5$ for the remainder of this section.

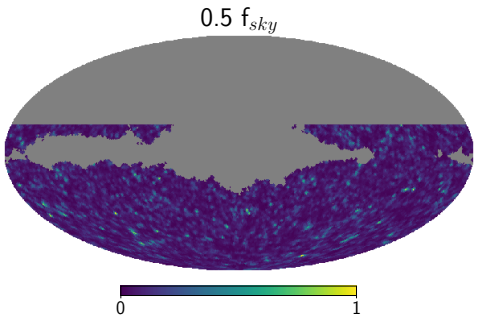
Shown below in Figure 3.8, are the different sky areas that were probed, the grey regions are parts of the sky that we have masked out. It should be pointed out that the results are not dependent on the location of the sky, but rather on the sky area covered; the presence of the galactic centre would be masked out. In the case of Figure 3.8d (bottom right) we masked

(a) ≈10% sky fraction

(b) ≈20% sky fraction



(c) ≈25% sky fraction



(d) ≈50% sky fraction

Fig. 3.8 The various sky areas that were probed in this analysis. The grey areas represent the masked regions.

out the galactic centre as well, this was done by removing all pixels in the Haslam 408MHz all sky map with synchrotron emission above 40 mK (following the work of [44]). In doing so we expect to see different results when calculating the total signal-to-noise for the BAO wiggles.

Radial Power Spectrum

From Figure 3.9, we see that an increase in the sky area results in an increase of the radial power spectrum amplitude. This can be seen in Figure 3.9 which covers 10% up to 50% of the sky. As one probes larger sky fractions, the amplitudes of the power spectrum increases due to the contribution of the noise. For each sky fraction case, we consider the line-of-sight power spectra that include cosmological signal plus noise only and the case of cosmological signal plus noise plus foregrounds ($n_{fg} = 5$). At the BAO scales, the foreground cleaned

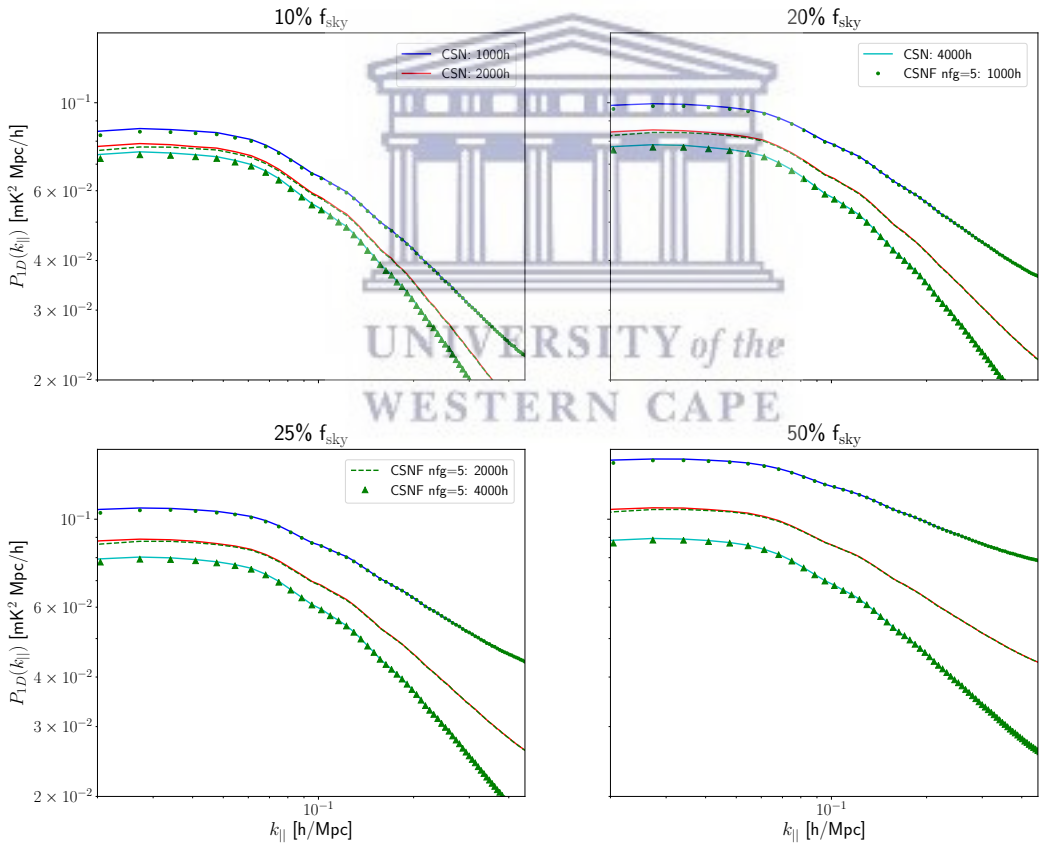


Fig. 3.9 The average Radial Power Spectrum for various sky fractions, fsky $= 0.1, 0.2, 0.25, 0.5$. The solid lines represent the HI signal plus noise with blue, red, cyan being 1000, 2000, 4000 observational respectively. The green circles, dashed line and triangles represent 1000, 2000, 4000 observational hours respectively for the case of HI signal plus systematic noise and foreground with $n_{fg} = 5$.

power spectrum is able to overlap with its cosmological signal plus noise counterpart, this implies that we are able to remove the foreground contaminants and recover the CSN spectrum. At large scales however, the effects of foreground cleaning on the cosmological signal or noise component can be seen. A slight dip in power can be noticed at this point. The average rms shown in Figure 3.10, describes the effect that the sky fraction has on the variation of the line-of-sight power spectrum (with noise included). An increase in the sky fraction reduces the rms amplitude for each radial power spectrum; implying that a larger area size constrains the variation in the radial power spectrum. The rms for CSNF5 follows the rms of the CSN for all observational hours.

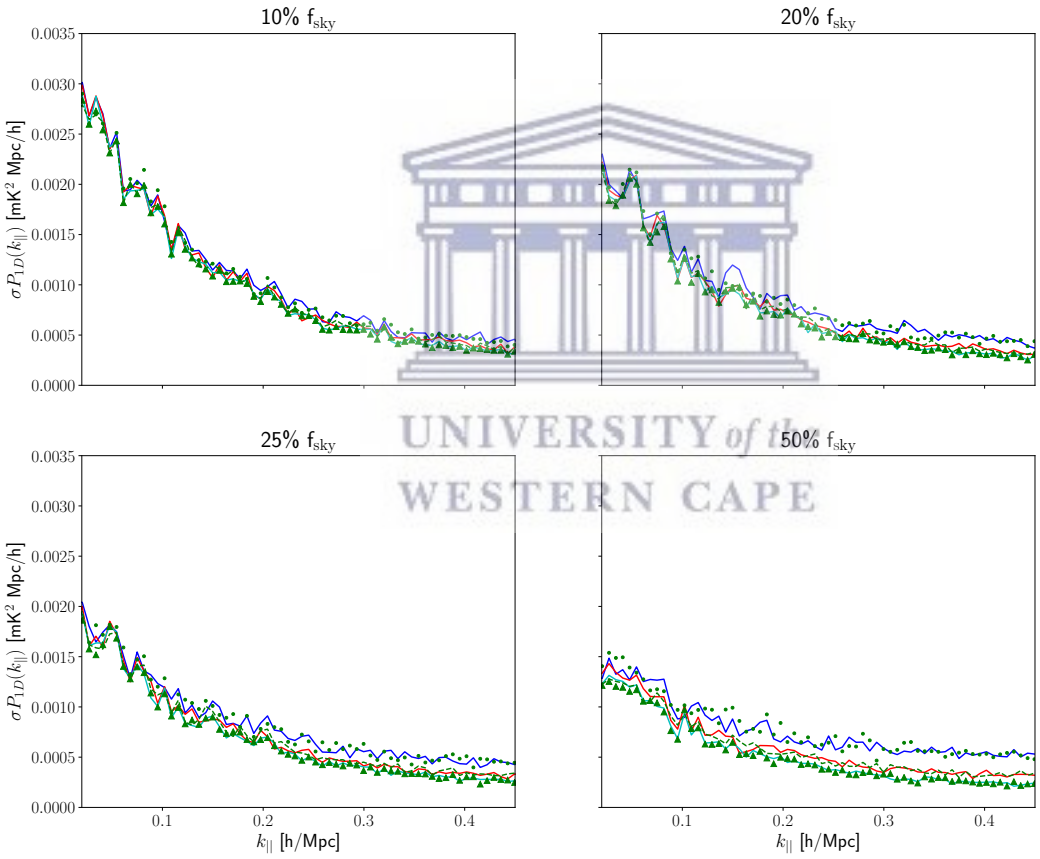


Fig. 3.10 The average error on the radial power spectrum for various sky fractions, fsky $= 0.1, 0.2, 0.25, 0.5$. The solid lines represent the HI signal plus noise with blue, red, cyan being 1000, 2000, 4000 observational respectively. The green circles, dashed line and triangles represent 1000, 2000, 4000 observational hours respectively for the case of HI signal plus systematic noise and foreground with $n_{fg} = 5$.

Extracting the BAO Signature

The polynomial smooth component (as done for the previous result in Figure 3.6) is applied to each CSNF5 radial power spectrum; this is done for each of the simulations individually. After the subtraction, the averaged residual wiggles are shown below (Figure 3.11) for the different sky fractions and observational hours. We are able to extract the BAO feature (visually) across the various sky fraction and observational hours.

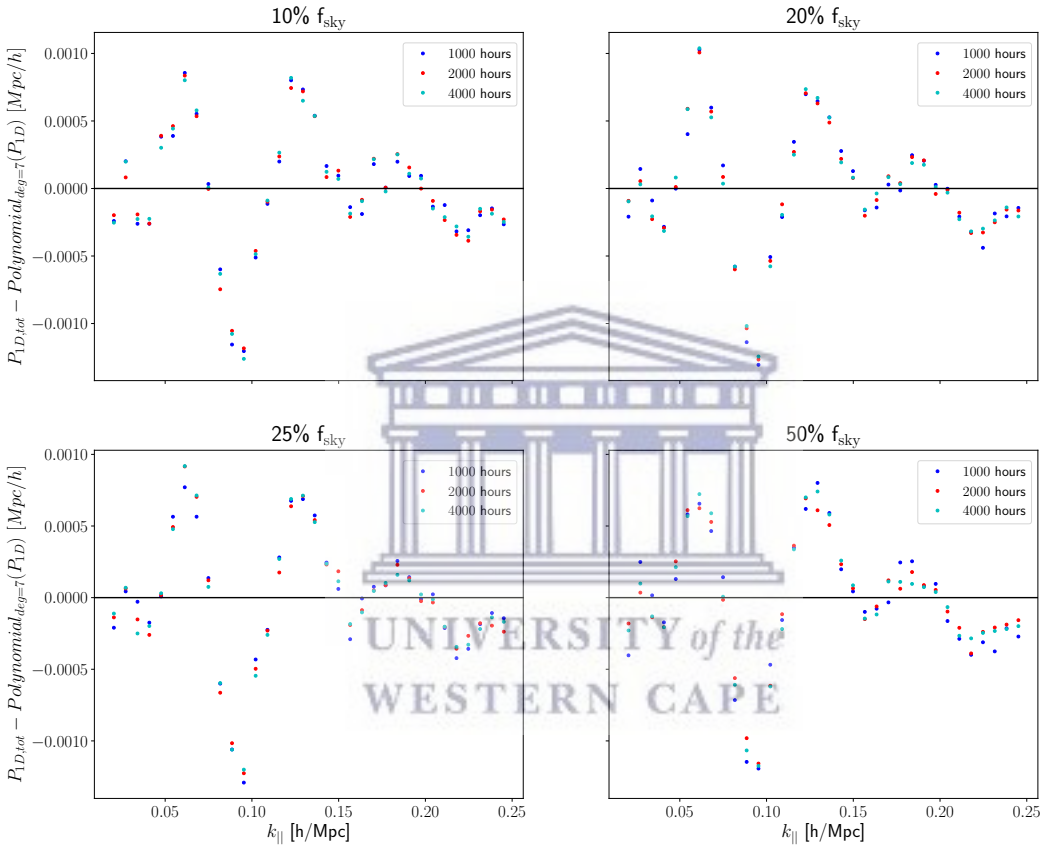


Fig. 3.11 The average BAO signal for 100 simulations, across different sky fractions sky fractions f_{sky} 0.1, 0.2, 0.25, 0.5 or an equivalent sky area of 4000, 8000, 10000, 20000 deg^2. The blue, red, cyan dots represent 1000, 2000, 4000 observational hours respectively.

In Figure 3.13 the rms is calculated using Eq.(3.1) except we are applying it to the BAO wiggles and not the full power spectrum. The rms increases at larger scales of the BAO feature and gradually decreases as we move to smaller scale of the BAO. The amplitude of the rms also decreases across the increasing sky fraction, this follows the same trend as in Figure 3.10. A larger sky area constrains better the variation on the BAO feature. Increasing the number of observational hours causes a slight decreases in the rms as well. This result is

evident in the larger sky area's (20000 deg^2), however in the smaller areas it is difficult to distinguish between the various observational hours.

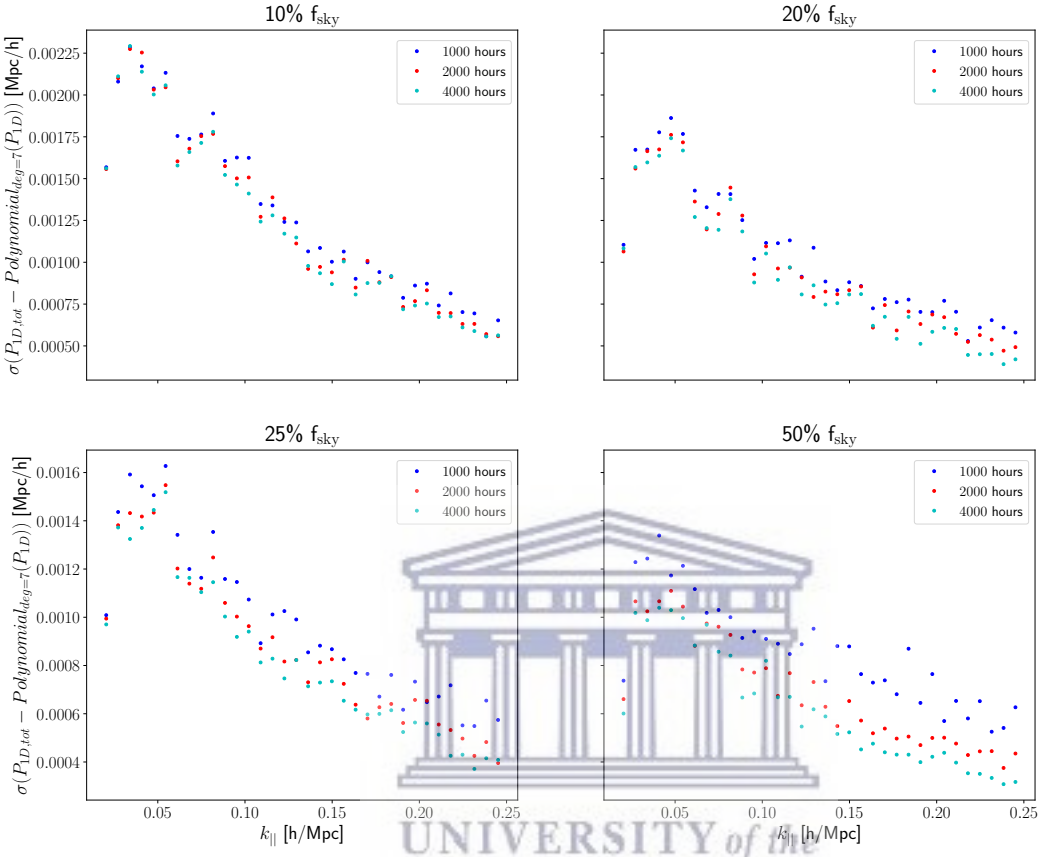


Fig. 3.12 The rms error of the BAO signal for different sky fractions f$_{sky}$ 0.1, 0.2, 0.25, 0.5 or an equivalent sky area of 4000, 8000, 10000, 20000 deg^2. The blue, red, cyan dots represent 1000, 2000, 4000 observational hours respectively.

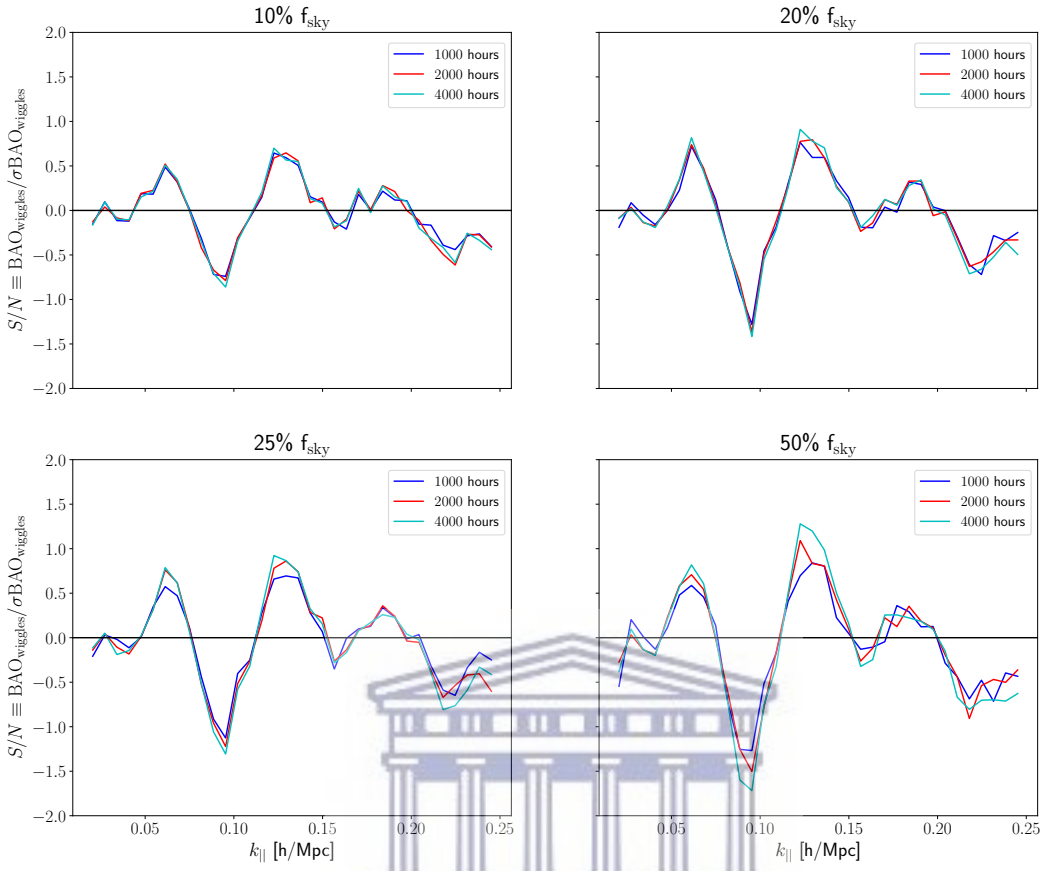


Fig. 3.13 The BAO S/N for different sky fractions f_{sky} 0.1, 0.2, 0.25, 0.5 or an equivalent sky area of 4000, 8000, 10000, 20000 deg^2. The blue, red, cyan lines represent 1000, 2000, 4000 observational hours respectively.

In Figure 3.13 we see that the observational time, $t_{obs} = 1000$ to 4000 hours in the case of the 10% sky area overlaps. This is in regards to a signal contribution of the cosmological plus noise plus foregrounds: $n_{fg} = 5$ (shown in Table 3.2). Moving to larger sky areas $f_{sky} \geq 0.2$ (8000 deg^2) we see that there is now a visible difference between the observational hour. In regards to the 20000 deg^2(50%) sky area there is a significant difference in the S/N ratio. From these results we can also make the claim that a larger surface contribute more to the S/N than the observational hours. This can be seen when comparing the 10% sky area 4000 hours case to the 25% or 50% sky area 2000 hour case.

In order to compare the results quantitatively, we have calculated the total S/N across the BAO scale with the following equation;

$$(S/N)^{\mathrm{BAO}}_{\mathrm{total}} = \sqrt{\sum_{i}^{n} (S/N)_i^2}. \tag{3.8}$$

Hours	CS	CSN	CSNF5	CSNF6	CSNF7	CSNF8
			10% $\mathbf{f}_{sky}$			
1000	2.121	1.879	1.923	1.896	1.861	1.773
2000	2.121	2.053	2.064	2.022	1.976	1.914
4000	2.121	2.056	2.102	2.079	2.017	1.957
			20% $\mathbf{f}_{sky}$			
1000	2.945	2.454	2.598	2.572	2.508	2.441
2000	2.945	2.706	2.688	2.671	2.602	2.519
4000	2.945	2.817	2.894	2.881	2.794	2.709
			25% $\mathbf{f}_{sky}$			
1000	3.147	2.654	2.487	2.469	2.424	2.371
2000	3.147	2.718	2.85	2.844	2.802	2.765
4000	3.147	3.037	3.061	3.055	3.028	3.005
			50% $\mathbf{f}_{sky}$			
1000	4.315	3.051	3.054	3.014	2.893	2.758
2000	4.315	3.38	3.371	3.32	3.313	3.289
4000	4.316	3.726	4.036	4.005	3.982	3.932

Table 3.2 Total S/N for different sky areas and varying observational hours.

In Table 3.2 we can see the effect that varying the sky area and the observational time would have on the total signal-to-noise for the BAO wiggles. In the second column we have the cosmological signal (CS) which is the HI signal only. Since there is no noise or foregrounds this value is constant over the different observational hours for the same sky area, this is to be expected and serves as a good check to ensure that the procedure is correct. Therein the third column includes the systematic noise which is introduced onto the power spectrum, here the effects of sky area on observational time can be seen in the variation of the S/N. The remaining columns include the foregrounds residuals and n_{fg} represents the number of

modes removed. As we increase n_{fg}, we observe a decrease in the S/N, this can be related to over cleaning of the total signal. Once the S/N value becomes lower than the CSN column we can say that we are removing the HI signal as well. From Table 3.2 we see that increasing the sky area and the observation hours, offer a better S/N result. The increase in sky area contributes more to the S/N, compared to when different observational times are considered.

3.4 Polynomial bias test

In the previous section "Dependence on observation design" we have tested how well our polynomial (smooth) estimator performs with regards to the signal-to-noise. In this section we wish to investigate if our estimator is unbiased. In order to test this we re-ran the *CRIME* simulation for two different scenarios, one scenario had an initial power spectrum with wiggles and the second had an initial power spectrum with no-wiggles (smooth). We obtained the no-wiggle power spectrum by applying the method described in [14] to the power spectrum used in the first scenario. All other simulation parameters were kept constant between the two scenarios. Each scenario contains a suite of 300 simulations, when generating the cosmological HI maps, the seed parameter is incremented in steps of 1 from 1000 to 1299, this ensures that each scenario has one wiggle and one non-wiggle simulation pair. The seed parameter for the noise realization is unique and set to UNIX time. Each scenario produces one realization of foreground maps and these maps are then combined with the other realizations for the respective scenario.

For our investigation we selected survey parameters for MeerKAT based on the results of Table 3.2. We selected the 20000 deg^2 (50% f_{sky}) and 2000 hours. We opted for 2000 hours, since doubling the observational hours did not considerably improve the S/N result even though 4000 hours offer a higher S/N. We will then compare the subtraction using the polynomial estimator (referred to as polynomial) to that of the smooth simulation (referred to as no-wiggle) in order to obtain the BAO wiggles. The subtraction is done per simulation for both cases and the residuals are then averaged together. Meaning that each polynomial fit and each no-wiggle fit is unique to one simulation in the suite.

In Figure 3.14 we see the results of the two subtraction methods. The polynomial subtraction (left) and the no-wiggle subtraction (right). We have already discussed the method behind the polynomial subtraction in the previous section, the no-wiggle subtraction follows a different route. In the case of the no-wiggle subtraction we have ran our simulation in order to obtain no-wiggle radial HI power spectra for each of the different cases (signal

only, with noise, with noise and foregrounds). This is possible in the simulation however in reality we would only be able to subtract the smooth cosmological HI signal; both the noise and the foreground residuals follow an unknown polynomial. Therefore our approach of using a polynomial to estimate the smooth component in the radial power spectrum is model independent.

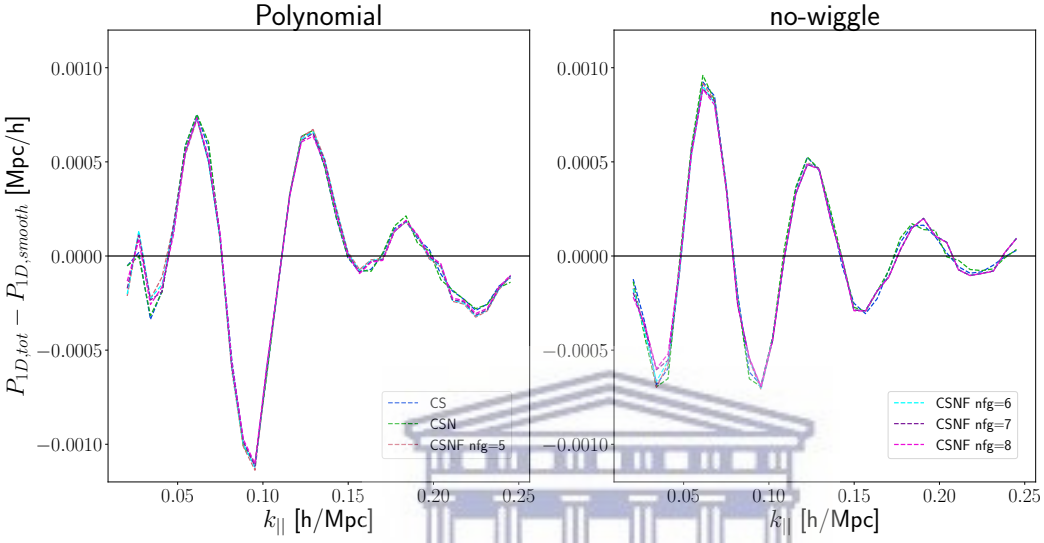


Fig. 3.14 Comparison of the BAO wiggles between the polynomial and the no-wiggle subtraction methods. The colours correspond to the different cases: cosmological HI signal (blue), cosmological HI signal plus noise (green), CSNF5 (rose), CSNF6 (cyan), CSNF7 (purple) and CSNF8 (fuchsia).

Overlapping the results of Figure 3.14 we see that in Figure 3.15 our polynomial estimator follows the no-wiggle subtraction. Accordingly; our estimator is able to identify the BAO features in the various radial power spectra and therefore unbiased in detecting the wiggles. Then, Figure 3.16 compares how our polynomial estimator performs on detecting the BAO wiggles using the foreground residual power spectrum (CSNF5), which is depicted using a blue dash curve. The blue dash curve represents the no-wiggle CS for the BAO wiggles, in reality this curve has been approximated (eg: [44, 67]) and becomes our standard of test.

As a consequence we conclude that our polynomial method can detect the BAO wiggles from the foreground residual case (CSNF5) and it shows it can identify the BAO peaks. This is corroborated by the blue dash curve, since the positions of the peaks are aligned. The no-wiggle CS curve is also contained within our polynomial variation as well.

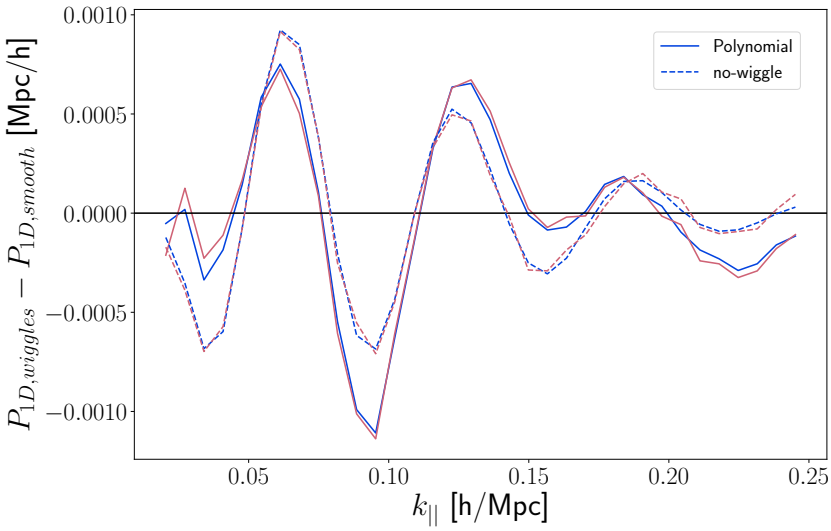


Fig. 3.15 Combining the results from Figure 3.14. The polynomial subtraction is shown with a line and the no-wiggle subtraction is shown with a dashed line. The colours represent the cosmological HI signal only (blue) and CSNF5 (rose).

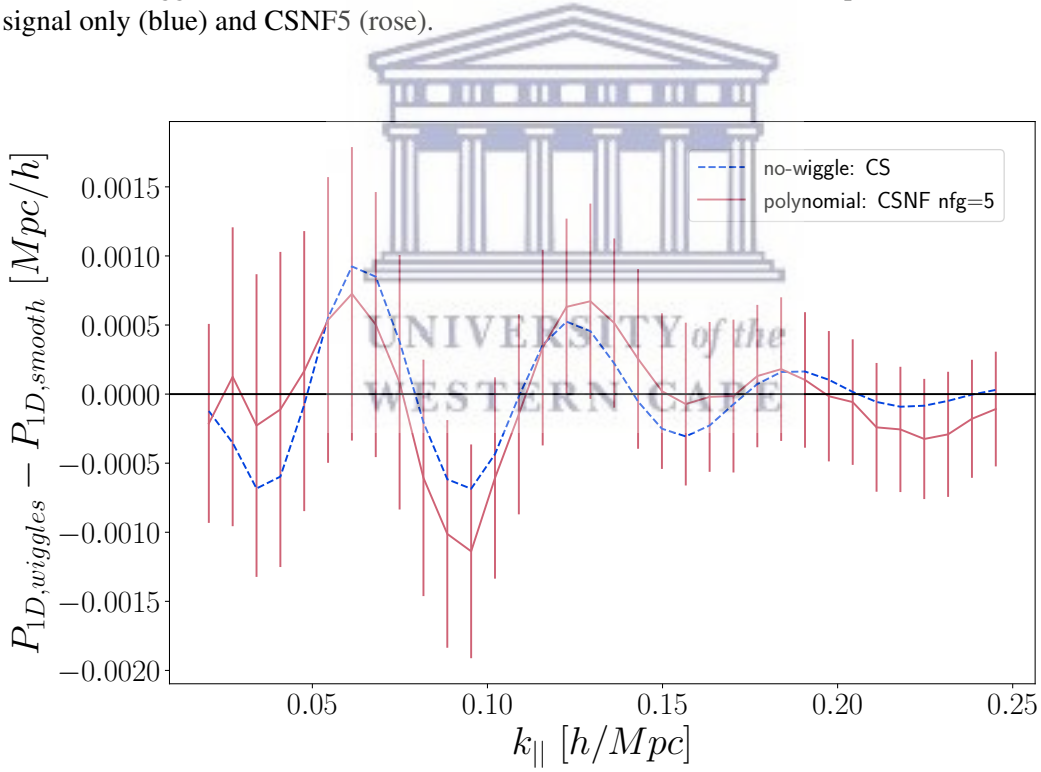
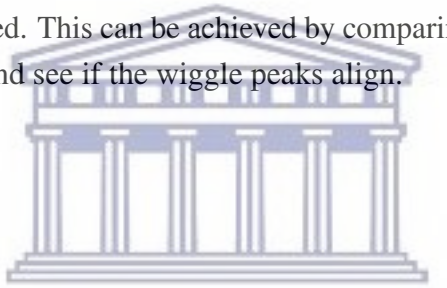


Fig. 3.16 The comparison between no-wiggle method (CS case) shown in the blue dash line and our polynomial estimator (CSNF5) shown in rose curve with error bars.

3.5 Summary

In this chapter we explain the methodology and results of the thesis. We study how we are able to extract the BAO wiggle feature using smooth polynomial fit for an ideal HI signal only case and calculate the signal-to-noise ratio. We find a signal-to-noise ratio be greater than 1. Next we study how the adding systematic noise, foreground contamination and foreground cleaning with various degree of cleaning applied effects the extraction of the BAO wiggle feature. This was to test how effective the extraction for the BAO wiggle feature is in a realistic case. We also find that aggressive cleaning $n_{fg} > 5$, does not improve results for extraction and therefore focused on $n_{fg} = 5$ for the next section. Based on those results we looked at how changing the survey parameters,i.e: sky area and observational hours would effect the results for the signal-to-noise ratio in regards to the BAO wiggle feature. We find that it is possible to reduce the observation time by increasing the survey area without compromising the results. The last section for this chapter is to test if our polynomial estimator is unbiased. This can be achieved by comparing our polynomial against a no-wiggle power spectrum and see if the wiggle peaks align.



Chapter 4

Conclusion

The aim of the research was to test the capability of MeerKAT, in detecting the Baryonic Acoustic Oscillations using the radial 21cm power spectrum. We employ numerical simulations in order to reconstruct intensity maps of the sky, which MeerKAT will observe in single-dish mode. We simulate the effects of a Gaussian telescope beam onto the signal. We employ methods to remove instrumental noise and foreground contamination.

From the literature: "Baryonic acoustic oscillations from 21cm intensity mapping: the Square Kilometre Array case" [44], we confirmed the effects that the instrumental beam has on the 21cm signal power spectrum in both the isotropic and the radial case. Mimicking the MeerKAT beam with a simple Gaussian function, we have shown that the BAO signature diminishes in the isotropic power spectrum as the angular smoothing scale increases. However, in the radial case, an increase in the angular smoothing scale enhances the BAO wiggle signature.

With the aid of publicly available simulations called *CRIME* and *fg_rm* we are able to construct both the 21cm intensity maps and the 21cm radial power spectrum. We follow the survey specifications set out in the proposed MeerKAT HI IM survey[38], which consists of a 4000 $\deg^2$ sky area and an observational time of 4000 hours. The effects of instrumental noise and foreground contamination were also taken into account by combining these signals to the HI intensity maps creating more realistic intensity maps. In order to remove the cosmological foreground contamination, the PCA method was used as this seemed to be the best technique available to us, based on literature [44] and previous work [66] for various degrees of freedom n_{fg}= 5,6,7,8. After the foreground cleaning method was applied we were able to recover the cosmological HI signal with noise in the radial 21cm power spectrum. We showed that it is possible to remove the foreground contaminants with $n_{fg} = 5$ (the least

aggressive cleaning mode) and increasing this number shows little improvement (see Fig 3.5) to the expected HI signal plus noise.

Once we were satisfied with the foreground removal (cleaning), we created a model-independent smooth component; the smooth component took the form of a 7^{th} order polynomial. This polynomial would fit the cosmological smooth power spectrum as well as both the noise and the foreground residuals. Such a polynomial was fitted over the $k_{||}$ region where the BAO resides. This is based on results obtained from [36, 44, 68]. We then subtracted our polynomial fit from the actual results for the various cases considered; the subtraction was done per simulation and then averaged the resultant together to obtain the wiggles. We are then also able to determine the rms and the total signal to noise for the BAO wiggles. This is done for the various cases, refer to Table 3.2. We achieve a signal to noise ratio > 1, which implies signal detection of the BAO wiggles in the case of a MeerKAT survey, surveying with a sky area of 4000 deg^2 and an observational time of 4000 hours.

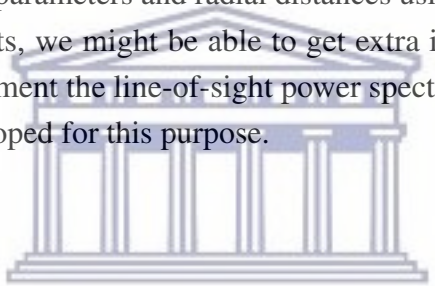
We followed up our investigation by looking at the effects of increasing the sky area from 4000 deg^2 to 20000 deg^2. Increasing the sky area results in an increase in the signal-to-noise ratio. We also investigated the variation of the observational time t_{obs} = 1000, 2000 and 4000 hours, would have on detecting the BAO wiggles. The sky area was not determined based on location but rather on the area size, in doing so we avoided (masked out) the harsh galactic synchrotron emission. Using the same estimator technique and method; we obtain the signal to noise for the BAO wiggles, refer to Table 3.8. From our results, we see that a telescope such as MeerKAT can achieve a better signal to noise ratio for the BAO wiggles, if the sky area increases. With an increase in sky area, we are able to reduce the amount of observational hours, without compromising the results. We see this when comparing CSNF5 results for the $A_{sky} = 4000$ deg^2 and $t_{obs} = 4000$ hours specifications and $A_{sky} = 20000$ deg^2 and $t_{obs} = 2000$ hours. This suggests that if one is only concerned about measuring the BAO wiggles, then a larger sky area reduces the amount of observational time for the experiment.

Lastly we investigate the biased nature of our estimator in identifying the wiggles position. To do this we ran the simulation in order to produce a suite of no-wiggle 21cm radial power spectra, that would have a simulated wiggle counterpart. We used the optimal survey based on previous results; $A_{sky} = 20000$ deg^2 and $t_{obs} = 2000$ hours case. We then applied our polynomial method and our no-wiggle method. The results are shown in Figure 3.15 and Figure 3.16, since our polynomial method does not cause a shift in the position of BAO

wiggles when compared to no-wiggle subtraction method we can conclude that our estimator is unbiased.

Overall, 21cm intensity mapping with MeerKAT will be a revolutionary survey that can produce concrete detection of the Baryonic Acoustic Oscillations. A MeerKAT proposed survey covering $A_{sky} = 4000$ deg^2 and $t_{obs} = 4000$ hours, the survey area will be able to detect the BAO signature, however we find that it is not optimally suited. If one wishes only to detect the BAO signature along the line-of-sight, we could essentially accomplish this with 2000 hours by increasing the sky area covered to 20000 deg^2.

I believe that further study can be done in regards to Detecting the BAO using HI IM with MeerKAT; similar to those set in [44] for SKA1-MID. We have only shown that the BAO signal can be detected with MeerKAT, we have not shown how MeerKAT can better constrain the cosmological parameters and radial distances using the radial power spectrum. Moreover, at lower redshifts, we might be able to get extra information from the perpendicular direction to complement the line-of-sight power spectrum and therefore fast robust estimators need to be developed for this purpose.



References

[1] Scott Dodelson. *"Modern Cosmology"*. Academic Press, Amsterdam, 2003.

[2] Daniel Baumann. Cosmology Part III Mathematical Tripos (Lectures). URL: http://www.damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf. Last visited on 2017/07/01.

[3] David W. Hogg. Distance measures in cosmology. *arXiv e-prints*, pages astro–ph/9905116, May 1999.

[4] W. Hu and M. White. The Cosmic Symphony. *Scientific American*, 290(2):44–53, February 2004.

[5] Patrick Peter and Jean-Philippe Uzan. *Primordial cosmology*. Oxford Graduate Texts. Oxford Univ. Press, Oxford, 2009.

[6] Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D: Particles, Fields, Gravitation and Cosmology*, vol. 23:pages 347–356, Jan 1981.

[7] Nick Kaiser. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, vol. 227:pages 1–21, Jul 1987.

[8] David H. Weinberg, Michael J. Mortonson, Daniel J. Eisenstein, Christopher Hirata, Adam G. Riess, and Eduardo Rozo. Observational probes of cosmic acceleration. *Physics Reports*, vol. 530:pages 87–255, Sep 2013.

[9] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. , 116:1009–1038, September 1998.

[10] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and T. S. C. Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. , 517:565–586, June 1999.

[11] Bruce Bassett and Renée Hlozek. Baryon acoustic oscillations. In Pilar Ruiz-Lapuente, editor, *Dark Energy: Observational and Theoretical Approaches, Edited by Pilar Ruiz-Lapuente. Published: Cambridge, UK ; New York by Cambridge University Press, 2010. ISBN 9780521518888*, page 246, 2010.

[12] P. J. E. Peebles and J. T. Yu. Primeval Adiabatic Perturbation in an Expanding Universe. *The Astrophysical Journal*, vol. 162:pages 815, Dec 1970.

[13] Daniel J. Eisenstein, Hee-Jong Seo, and Martin White. On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter. *The Astrophysical Journal*, vol. 664:pages 660–674, Aug 2007.

[14] P. Bull, S. Camera, A. Raccanelli, C. Blake, P. Ferreira, M. Santos, and D. J. Schwarz. Measuring baryon acoustic oscillations with future SKA surveys. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 24, Apr 2015.

[15] Daniel Eisenstein and Martin White. Theoretical uncertainty in baryon oscillations. *Physical Review D: Particles, Fields, Gravitation and Cosmology*, vol. 70:page 103523, Nov 2004.

[16] Shaun Cole, Will J. Percival, John A. Peacock, Peder Norberg, Carlton M. Baugh, Carlos S. Frenk, Ivan Baldry, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Matthew Colless, Chris Collins, Warrick Couch, Nicholas J. G. Cross, Gavin Dalton, Vincent R. Eke, Roberto De Propris, Simon P. Driver, George Efstathiou, Richard S. Ellis, Karl Glazebrook, Carole Jackson, Adrian Jenkins, Ofer Lahav, Ian Lewis, Stuart Lumsden, Steve Maddox, Darren Madgwick, Bruce A. Peterson, Will Sutherland, and Keith Taylor. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *Monthly Notices of the Royal Astronomical Society*, vol. 362:pages 505–534, Sep 2005.

[17] Daniel J. Eisenstein, Idit Zehavi, David W. Hogg, Roman Scoccimarro, Michael R. Blanton, Robert C. Nichol, Ryan Scranton, Hee-Jong Seo, Max Tegmark, Zheng Zheng, Scott F. Anderson, Jim Annis, Neta Bahcall, Jon Brinkmann, Scott Burles, Francisco J. Castand er, Andrew Connolly, Istvan Csabai, Mamoru Doi, Masataka Fukugita, Joshua A. Frieman, Karl Glazebrook, James E. Gunn, John S. Hendry, Gregory Hennessy, Zeljko Ivezić, Stephen Kent, Gillian R. Knapp, Huan Lin, Yeong-Shang Loh, Robert H. Lupton, Bruce Margon, Timothy A. McKay, Avery Meiksin, Jeffery A. Munn, Adrian Pope, Michael W. Richmond, David Schlegel, Donald P. Schneider, Kazuhiro Shimasaku, Christopher Stoughton, Michael A. Strauss, Mark SubbaRao, Alexander S. Szalay, István Szapudi, Douglas L. Tucker, Brian Yanny, and Donald G. York. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *The Astrophysical Journal*, vol. 633:pages 560–574, Nov 2005.

[18] Karl jansky and the discovery of cosmic radio waves, December 2018. [online] https://www.nrao.edu/whatisra/hist_jansky.shtml.

[19] Grote reber and his radio telescope, December 2018. [online] https://www.nrao.edu/whatisra/hist_reber.shtml.

[20] Radio telescope-astronomical instrument, December 2018. [online] https://www.britannica.com/science/radio-telescope.

[21] The ska project, December 2018. [online] https://www.skatelescope.org/the-ska-project/.

[22] Square Kilometre Array Cosmology Science Working Group, David J. Bacon, Richard A. Battye, Philip Bull, Stefano Camera, Pedro G. Ferreira, Ian Harrison, David Parkinson, Alkistis Pourtsidou, Mario G. Santos, Laura Wolz, Filipe Abdalla, Yashar Akrami, David Alonso, Sambatra Andrianomena, Mario Ballardini, Jose Luis Bernal, Daniele Bertacca, Carlos A. P. Bengaly, Anna Bonaldi, Camille Bonvin, Michael L. Brown, Emma Chapman, Song Chen, Xuelei Chen, Steven Cunnington, Tamara M. Davis, Clive Dickinson, Jose Fonseca, Keith Grainge, Stuart Harper, Matt J. Jarvis, Roy Maartens, Natasha Maddox, Hamsa Padmanabhan, Jonathan R. Pritchard, Alvise Raccanelli, Marzia Rivi, Sambit Roychowdhury, Martin Sahlen, Dominik J. Schwarz, Thilo M. Siewert, Matteo Viel, Francisco Villaescusa-Navarro, Yidong Xu, Daisuke Yamauchi, and Joe Zuntz. Cosmology with Phase 1 of the Square Kilometre Array; Red Book 2018: Technical specifications and performance forecasts. *arXiv e-prints*, page arXiv:1811.02743, Nov 2018.

[23] South africa's meerkat radio telescope technical fact sheet, December 2018. [online] http://www.ska.ac.za/wp-content/uploads/2016/07/meerkat-fact-sheet-2016.pdf.

[24] M. Bailes, E. Barr, N. D. R. Bhat, J. Brink, S. Buchner, M. Burgay, F. Camilo, D. J. Champion, J. Hessels, G. H. Janssen, A. Jameson, S. Johnston, A. Karastergiou, R. Karuppusamy, V. Kaspi, M. J. Keith, M. Kramer, M. A. McLaughlin, K. Moodley, S. Oslowski, A. Possenti, S. M. Ransom, F. A. Rasio, J. Sievers, M. Serylak, B. W. Stappers, I. H. Stairs, G. Theureau, W. van Straten, P. Weltevrede, and N. Wex. MeerTime - the MeerKAT Key Science Program on Pulsar Timing. *arXiv e-prints*, page arXiv:1803.07424, Mar 2018.

[25] B. W. Holwerda, S. L. Blyth, and A. J. Baker. Looking at the distant universe with the MeerKAT Array (LADUMA). In Richard J. Tuffs and Cristina C. Popescu, editors, *The Spectral Energy Distribution of Galaxies - SED 2011*, volume 284 of *IAU Symposium*, pages 496–499, August 2012.

[26] A. J. Baker, S. Blyth, B. W. Holwerda, and LADUMA Team. LADUMA: Looking At the Distant Universe with the MeerKAT Array. In *American Astronomical Society Meeting Abstracts #231*, volume 231 of *American Astronomical Society Meeting Abstracts*, page 231.07, January 2018.

[27] N. Gupta, R. Srianand, W. Baan, A. J. Baker, R. J. Beswick, S. Bhatnagar, D. Bhattacharya, A. Bosma, C. Carilli, M. Cluver, F. Combes, C. Cress, R. Dutta, J. Fynbo, G. Heald, M. Hilton, T. Hussain, M. Jarvis, G. Jozsa, P. Kamphuis, A. Kembhavi, J. Kerp, H. R. Kloeckner, J. Krogager, V. P. Kulkarni, C. Ledoux, A. Mahabal, T. Mauch, K. Moodley, E. Momjian, R. Morganti, P. Noterdaeme, T. Oosterloo, P. Petitjean, A. Schroeder, P. Serra, J. Sievers, K. Spekkens, P. Vaisanen, T. van der Hulst, M. Vivek, J. Wang, O. I. Wong, and A. R. Zungu. The MeerKAT Absorption Line Survey (MALS). In *Proceedings of MeerKAT Science: On the Pathway to the SKA. 25-27 May*, page 14, January 2016.

[28] W. J. G. de Blok, E. A. K. Adams, P. Amram, E. Athanassoula, I. Bagetakos, C. Balkowski, M. A. Bershady, R. J. Beswick, F. Bigiel, S. L. Blyth, A. Bosma, R. S. Booth, A. Bouchard, E. Brinks, C. Carignan, L. Chemin, F. Combes, J. Conway, E. C. Elson, J. English, B. Epinat, B. S. Frank, J. Fiege, F. Fraternali, J. S. Gallagher, B. K. Gibson, G. Heald, P. A. Henning, B. W. Holwerda, T. H. Jarrett, H. Jerjen, G. I. Józsa, M. Kapala, H. R. Klöckner, B. S. Koribalski, R. C. Kraan-Korteweg, S. Leon, A. Leroy, S. I. Loubser, D. M. Lucero, S. S. McGaugh, G. R. Meurer, M. Meyer, M. Mogotsi, B. Namumba, S. H. Oh, T. A. Oosterloo, D. J. Pisano, A. Popping, S. Ratcliffe, J. A. Sellwood, E. Schinnerer, A. C. Schröder, K. Sheth, M. W. L. Smith, A. Sorgho, K. Spekkens, S. Stanimirovic, K. J. van der Heyden, L. Verdes- Montenegro, W. van Driel, F. Walter, T. Westmeier, E. Wilcots, T. Williams, O. I. Wong, P. A. Woudt, and A. Zijlstra. An Overview of the MHONGOOSE Survey: Observing Nearby Galaxies with MeerKAT. In *Proceedings of MeerKAT Science: On the Pathway to the SKA. 25-27 May*, page 7, January 2016.

[29] The mhongoose large survey project, December 2018. [online] https://mhongoose.astron.nl/.

[30] Transients and pulsars with meerkat, December 2018. [online] http://www.trapum.org/.

[31] P. Serra, W. J. G. de Blok, G. L. Bryan, S. Colafrancesco, R. J. Dettmar, B. S. Frank, F. Govoni, G. I. G. Jozsa, R. C. Kraan-Korteweg, F. M. Maccagni, S. I. Loubser, M. Murgia, T. A. Oosterloo, R. F. Peletier, R. Pizzo, L. Richter, M. Ramatsoku, M. W. L. Smith, S. C. Trager, J. H. van Gorkom, and M. A. W. Verheijen. The MeerKAT Fornax Survey. In *Proceedings of MeerKAT Science: On the Pathway to the SKA. 25-27 May*, page 8, January 2016.

[32] M. Jarvis, R. Taylor, I. Agudo, J. R. Allison, R. P. Deane, B. Frank, N. Gupta, I. Heywood, N. Maddox, K. McAlpine, M. Santos, A. M. M. Scaife, M. Vaccari, J. T. L. Zwart, E. Adams, D. J. Bacon, A. J. Baker, B. A. Bassett, P. N. Best, R. Beswick, S. Blyth, M. L. Brown, M. Bruggen, M. Cluver, S. Colafrancesco, G. Cotter, C. Cress, R. Davé, C. Ferrari, M. J. Hardcastle, C. L. Hale, I. Harrison, P. W. Hatfield, H. R. Klockner, S. Kolwa, E. Malefahlo, T. Marubini, T. Mauch, K. Moodley, R. Morganti, R. P. Norris, J. A. Peters, I. Prand oni, M. Prescott, S. Oliver, N. Oozeer, H. J. A. Rottgering, N. Seymour, C. Simpson, O. Smirnov, and D. J. B. Smith. The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey. In *Proceedings of MeerKAT Science: On the Pathway to the SKA. 25-27 May*, page 6, Jan 2016.

[33] R. Fender, P. A. Woudt, R. Armstrong, P. Groot, V. McBride, J. Miller-Jones, K. Mooley, B. Stappers, R. Wijers, M. Bietenholz, S. Blyth, M. Bottcher, D. Buckley, P. Charles, L. Chomiuk, D. Coppejans, S. Corbel, M. Coriat, F. Daigne, W. J. G. de Blok, H. Falcke, J. Girard, I. Heywood, A. Horesh, J. Horrell, P. Jonker, T. Joseph, A. Kamble, C. Knigge, E. Koerding, M. Kotze, C. Kouveliotou, C. Lynch, T. Maccarone, P. Meintjes, S. Migliari, T. Murphy, T. Nagayama, G. Nelemans, G. Nicholson, T. O'Brien, A. Oodendaal, N. Oozeer, J. Osborne, M. Perez-Torres, S. Ratcliffe, V. Ribeiro, E. Rol, A. Rushton, A. Scaife, M. Schurch, G. Sivakoff, T. Staley, D. Steeghs, I. Stewart, J. Swinbank, K. van der Heyden, A. van der Horst, B. van Soelen, S. Vergani, B. Warner, and K. Wiersema. ThunderKAT: The MeerKAT Large Survey Project for Image-Plane Radio Transients. *arXiv e-prints*, page arXiv:1711.04132, Nov 2017.

[34] M. Jarvis, D. Bacon, C. Blake, M. Brown, S. Lindsay, A. Raccanelli, M. Santos, and D. J. Schwarz. Cosmology with SKA Radio Continuum Surveys. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 18, Apr 2015.

[35] F. B. Abdalla, P. Bull, S. Camera, A. Benoit-Lévy, B. Joachimi, D. Kirk, H. R. Klöckner, R. Maartens, A. Raccanelli, M. G. Santos, and G. B. Zhao. Cosmology from HI galaxy surveys with the SKA. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 17, Apr 2015.

[36] Philip Bull, Pedro G. Ferreira, Prina Patel, and Mário G. Santos. Late-time Cosmology with 21 cm Intensity Mapping Experiments. *The Astrophysical Journal*, vol. 803:page 21, Apr 2015.

[37] M. Santos, P. Bull, D. Alonso, S. Camera, P. Ferreira, G. Bernardi, R. Maartens, M. Viel, F. Villaescusa-Navarro, F. B. Abdalla, M. Jarvis, R. B. Metcalf, A. Pourtsidou, and L. Wolz. Cosmology from a SKA HI intensity mapping survey. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 19, Apr 2015.

[38] Mario G. Santos, Michelle Cluver, Matt Hilton, Matt Jarvis, Gyula I. G. Jozsa, Lerothodi Leeuw, Oleg Smirnov, Russ Taylor, Filipe Abdalla, Jose Afonso, David Alonso, David Bacon, Bruce A. Bassett, Gianni Bernardi, Philip Bull, Stefano Camera, H. Cynthia Chiang, Sergio Colafrancesco, Pedro G. Ferreira, Jose Fonseca, Kurt van der Heyden, Ian Heywood, Kenda Knowles, Michelle Lochner, Yin-Zhe Ma, Roy Maartens, Sphesihle Makhathini, Kavilan Moodley, Alkistis Pourtsidou, Matthew Prescott, Jonathan Sievers, Kristine Spekkens, Mattia Vaccari, Amanda Weltman, Imogen Whittam, Amadeus Witzemann, Laura Wolz, and Jonathan T. L. Zwart. MeerKLASS: MeerKAT Large Area Synoptic Survey. *arXiv e-prints*, page arXiv:1709.06099, Sep 2017.

[39] Euclid consortium: A space mission to map the dark universe, July 2019. [online] https://www.euclid-ec.org/.

[40] The dark energy spectroscopic instrument (desi), July 2019. [online] https://www.desi.lbl.gov/.

[41] Boss: Baryon oscillation spectroscopic survey, July 2019. [online] https://cosmology.lbl.gov/BOSS/.

[42] Hetdex illuminating the darkness, July 2019. [online] http://hetdex.org/.

[43] Cosmology with the wfirst high latitude survey, July 2019. [online] https://www.wfirst-hls-cosmology.org/.

[44] Francisco Villaescusa-Navarro, David Alonso, and Matteo Viel. Baryonic acoustic oscillations from 21 cm intensity mapping: the Square Kilometre Array case. *Mon. Not. Roy. Astron. Soc.*, vol. 466(3):pages 2736–2751, 2017.

[45] R. A. Battye, I. W. A. Browne, C. Dickinson, G. Heron, B. Maffei, and A. Pourtsidou. H I intensity mapping: a single dish approach. *Monthly Notices of the Royal Astronomical Society*, vol. 434:pages 1239–1256, Sep 2013.

[46] Alkistis Pourtsidou. HI Intensity Mapping with MeerKAT. *arXiv e-prints*, page arXiv:1709.07316, Sep 2017.

[47] Tzu-Ching Chang, Ue-Li Pen, Kevin Bandura, and Jeffrey B. Peterson. Hydrogen 21-cm Intensity Mapping at redshift 0.8. *arXiv e-prints*, page arXiv:1007.3709, Jul 2010.

[48] K. W. Masui, E. R. Switzer, N. Banavar, K. Bandura, C. Blake, L. M. Calin, T. C. Chang, X. Chen, Y. C. Li, Y. W. Liao, A. Natarajan, U. L. Pen, J. B. Peterson, J. R. Shaw, and T. C. Voytek. Measurement of 21 cm Brightness Fluctuations at z ~0.8 in Cross-correlation. , vol. 763:page L20, Jan 2013.

[49] What Is Radio Astronomy?, December 2018. [online] https://www.skatelescope.org/radio-astronomy/.

[50] Ravi K. Sheth and Giuseppe Tormen. Large-scale bias and the peak background split. , 308(1):119–126, Sep 1999.

[51] Francisco Villaescusa-Navarro, Susana Planelles, Stefano Borgani, Matteo Viel, Elena Rasia, Giuseppe Murante, Klaus Dolag, Lisa K. Steinborn, Veronica Biffi, Alexander M. Beck, and Cinthia Ragone-Figueroa. Neutral hydrogen in galaxy clusters: impact of AGN feedback and implications for intensity mapping. *Monthly Notices of the Royal Astronomical Society*, vol. 456:pages 3553–3570, Mar 2016.

[52] Francisco Villaescusa-Navarro, Matteo Viel, Kanan K. Datta, and T. Roy Choudhury. Modeling the neutral hydrogen distribution in the post-reionization Universe: intensity mapping. *Journal of Cosmology and Astro-Particle Physics*, vol. 2014:page 050, Sep 2014.

[53] Ravi K. Sheth, H. J. Mo, and Giuseppe Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. , 323(1):1–12, May 2001.

[54] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient computation of CMB anisotropies in closed FRW models. , 538:473–476, 2000.

[55] Francisco Villaescusa-Navarro, Shy Genel, Emanuele Castorina, Andrej Obuljen, David N. Spergel, Lars Hernquist, Dylan Nelson, Isabella P. Carucci, Annalisa Pillepich, Federico Marinacci, Benedikt Diemer, Mark Vogelsberger, Rainer Weinberger, and Rüdiger Pakmor. Ingredients for 21 cm Intensity Mapping. *The Astrophysical Journal*, vol. 866:page 135, Oct 2018.

[56] George B. Rybicki and Alan P. Lightman. Radiative processes in astrophysics. In *A Wiley-Interscience Publication, New York: Wiley, 1979*. 1979.

[57] David Alonso, Philip Bull, Pedro G. Ferreira, and Mario G. Santos. Blind foreground subtraction for intensity mapping experiments. *Mon. Not. Roy. Astron. Soc.*, vol. 447:p. 400, 2015.

[58] L. C. Olivari, C. Dickinson, R. A. Battye, Y-Z. Ma, A. A. Costa, M. Remazeilles, and S. Harper. Cosmological parameter forecasts for HI intensity mapping experiments using the angular power spectrum. *Mon. Not. Roy. Astron. Soc.*, vol. 473(3):pages 4242–4256, 2018.

[59] A. Witzemann, D. Alonso, J. Fonseca, and M. G. Santos. Simulated multi-tracer analyses with HI intensity mapping. *Monthly Notices of the Royal Astronomical Society*, page 747, Mar 2019.

[60] Cosmological realizations for intensity mapping experiments, July 2017. [online] http://intensitymapping.physics.ox.ac.uk/CRIME.html.

[61] Karl Pearson F.R.S. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2(11):pages 559–572, 1901.

[62] David Alonso, Pedro G. Ferreira, and Mário G. Santos. Fast simulations for intensity mapping experiments. *Mon. Not. Roy. Astron. Soc.*, vol. 444(4):pages 3183–3197, 2014.

[63] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astronomy Astrophysics*, vol. 594:page A13, Sep 2016.

[64] Mário G. Santos, Asantha Cooray, and Lloyd Knox. Multifrequency Analysis of 21 Centimeter Fluctuations from the Era of Reionization. *Astrophysical Journal*, vol. 625:pages 575–587, Jun 2005.

[65] C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408 MHz all-sky continuum survey. II. The atlas of contour maps. *Astronomy and Astrophysics Supplement Series*, vol. 47:pages 1–143, Jan 1982.

[66] Sibonelo Ngobese. Point source simulations and foreground cleaning techniques for hi intensity mapping. Master's thesis, University of the Western Cape, Robert Sobukwe Rd, Bellville, Cape Town, South Africa, 2018.

[67] Daniel J. Eisenstein and Wayne Hu. Baryonic features in the matter transfer function. *Astrophys. J.*, 496:605, 1998.

[68] Alkistis Pourtsidou, David Bacon, and Robert Crittenden. HI and cosmological constraints from intensity mapping, optical and CMB surveys. *Mon. Not. Roy. Astron. Soc.*, vol. 470(4):pages 4251–4260, 2017.

Appendix A

A.1 $R_{\mu\nu}$ - Ricci Tensor

The Ricci tensor is expressed in terms of the Christoffel symbol

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\beta\alpha}\Gamma^{\beta}_{\mu\nu} - \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\mu\alpha} \tag{A.1}$$

There are two sets are of non-vanishing terms in a FRW Universe: if $\mu = \nu = 0$ and the other if $\mu = \nu = i$

A.2 $\mathscr{R}$ - Ricci Scaler

$$\mathscr{R} \equiv g^{\mu\nu} R_{\mu\nu} \tag{A.2}$$

A.3 Γ - Christoffel symbol

$$\Gamma^{\mu}_{\alpha\beta} = \frac{g^{\mu\nu}}{2}\left[\frac{\partial g_{\alpha\nu}}{\partial x^{\beta}} + \frac{\partial g_{\beta\nu}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\nu}}\right] \tag{A.3}$$