Cosmology with HI intensity mapping: effect of higher order corrections

UNIVERSITY *of the*
WESTERN CAPE

Liantsoa Finaritra Randrianjanahary

Department of Physics and Astronomy

University of the Western Cape

Supervisor

Prof. M Santos

Co-supervisor

Dr. A Pénin

In partial fulfillment of the requirements for the degree of

Master of Science

May 9, 2020

Declaration

I, Liantsoa Finaritra Randrianjanahary, declare that this thesis titled, "Cosmology with HI intensity mapping: effect of higher order corrections" 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.

Liantsoa Finaritra Randrianjanahary.

Acknowledgements

First and foremost, I would like to extend my sincere gratitude to my Supervisor Prof Mario Santos for giving me a chance to work with him and to be part of the Centre for Radio Cosmology (CRC) team. He has gave his time to offer me valuable comments toward improving my work. I would like to thank him for the patient guidance, encouragement and advice he has provided throughout my time as his student.

There are no proper words to convey my deep gratitude and respect for my Co-Supervisor, Dr Aurélie Pénin. She has given her time and effort to generously assist me. She also provided me constructive criticism which helped me to develop a broader perspective to my thesis. I have learnt extensively from her and her advice on both research as well as on my career have been invaluable.

I am grateful to my fellows and to all of those with whom I have had the pleasure to work during this thesis. I am also thankful to the Astrophysics group at the University of the Western Cape. I am indebted to them for their help in numerous ways.

I am thankful to the South African Radio Astronomy Observatory (SARAO), the Centre for Radio Cosmology (CRC) and the National Research Foundation (NRF) for their financial support throughout my study, without them I would not have gotten to this point.

Finally, I must express my very profound gratitude to my family for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of completing this thesis. Thank you for standing with me in every phase of life.

I dedicate this thesis to my late mother, Hanitra.

Abstract

One of the main challenges of cosmology is to unveil the nature of dark energy and dark matter. They can be constrained with baryonic acoustic oscillations (BAO) and redshift space distortions, amongst others. Both have characteristic signatures in the dark matter power spectrum. Biased tracers of dark matter, such as neutral hydrogen, are used to quantify the underlying dark matter density field. It is generally assumed that on large scales the bias of the tracer is linear. However, there is a coupling between small and large scales of the biased tracer which gives rise to a significant non-linear contribution on linear scales in the power spectrum of the biased tracer. The Hydrogen Intensity and Real-time eXperiment (HIRAX) will map the brightness temperature of neutral hydrogen (HI) over BAO scales thanks to the intensity mapping technique. We forecasted cosmological parameters for HIRAX taking into account non-linear corrections to the HI power spectrum and compared them to the linear case. We used methods based on Fisher matrices. We found values for the bias to error ratio of the cosmological parameters as high as 1 or 7, depending on the noise level. We also investigated the change in peaks location on the baryonic acoustic oscillations signal. The value of the shift goes up to $\Delta k = 10^{-2}$h/Mpc with a reduction of amplitude of the BAO features from 16.33% to 0.33%, depending on the scales.

Keywords: Baryonic acoustic oscillation, Fisher matrices, Cosmology.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Relativistic Cosmology aims to determine the structure of the Universe using results from astronomical observations and knowledge derived from physical experiment (Ellis **1975**). To implement an experiment on how to constrain or how to make a measurement of some quantities, we need an observable. Neutral hydrogen, which is the most abundant baryonic element of the Universe, is an obvious choice for such studies.

We can use a radio-telescope to collect the signal from the sky and then retrieve information from it to conduct our studies. Due to the spin-flip transition, neutral hydrogen emits radiation which lies in the radio range of electromagnetic waves. This is called the 21cm line. Observations can be deformed, biased or even incorrect due to several facts like cosmological lensing (Er et al. **2013**), the Sunyaev Zel'dovich effect (Carlstrom et al. **2002**) or redshift space distortions (Nock et al. **2010**).

This project is based on cosmology using baryonic acoustic oscillations which leaves an imprint in the matter density field. We will consider non-linear corrections to the matter power spectrum and the HI power spectrum. Why do we care about non-linearity? The nonlinear domain appears to be a gold mine of cosmological information (Tegmark **1997**). Some perturbation techniques will be required at

that point. By taking into account non-linearities on the power spectrum of neutral hydrogen we are going to use a forecasting method to estimate cosmological parameters, which are our ultimate aim.

1.1 Background Cosmology

In this chapter, we list some of the major concepts of cosmology and describe the formation and evolution of the Large Scale Structure. Lastly, we describe our target tracer of the large scale structure, neutral hydrogen, along with one dedicated technique of observation and several experiments either already observing or being built.

1.1.1 The Cosmological Principle

The basic hypothesis of a Standard Cosmological theory is that all the points of the Universe have to be essentially equivalent: this hypothesis is required in order to avoid any privileged observer. This assumption is called the Cosmological Principle and relates to the concept of Homogeneity and Isotropy. The spatial distribution of visible matter in the Universe on large scales should be statistically homogeneous. By homogeneous, we mean that any spot in the Universe has the same conditions as any other. The Cosmological Principle also implies the condition of spherical symmetry with respect to every position, so the Universe is isotropic. The distribution of light in the Universe is statistically isotropic with respect to the barycentre of the Solar system (Schwarz 2010). This Isotropy, which is a symmetry under rotation, implies that there is no preferred direction in the Universe.Homogeneity and isotropy are complementary concepts. A homogeneous medium is not necessarily isotropic nor is the impression of isotropy at one location an evidence for the overall homogeneity of the Universe as shown in Figure 1.1. Observations confirm that we can see that

the distribution of matter is homogeneous and isotropic on large spatial scales on the order of $\sim$ 100 Mpc/h (Park et al., Scrimgeour et al. **2017**, **2012**).

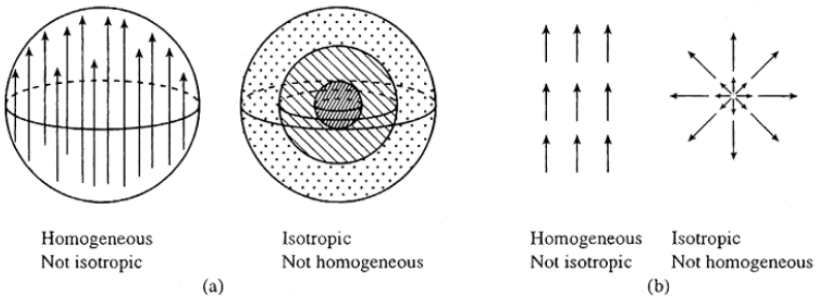


Figure 1.1 Illustration of how homogeneity and isotropy are equivalent in (a) three dimensions and in (b) two dimensions. In the first example of each, unique direction is picked out but translation invariance is maintained. In the second example of each, all directions are the same (rotation invariance) but a radial gradient exists. Image credits: James Schombert. From the website of Oregon University [2]

The extended Cosmological Principle includes two additional points. Firstly, if the Hubble expansion were not uniform (the Universe grows with the same rate at every location in every direction) it would be in conflict with the statement that the Universe's state would be everywhere and in each direction the same. Secondly, if the physical laws are not the same throughout the Universe or throughout cosmic history, we have a problem in formulating a cosmological theory.

In summary, all physical quantities measured by a comoving observer are spatially homogeneous and isotropic. This formulation leads us to the Friedmann models, in section 1.1.4, which are successful in describing the expansion of the Universe.

1.1.2 Einstein field equations

Einstein field equations describe gravity as a result of spacetime being curved by mass and energy. It is the pillar of general relativity which is a theory of gravity as geometry. Einstein field equations map the geometry of spacetime to the matter

and energy in the Universe. The distribution of mass and energy on the curvature of spacetime is given by the following Einstein field equations (i.e. equation 1.1):

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

where $G_{\mu\nu}$ is called the Einstein Tensor and is derived from the Riemann tensor $R_{\alpha\beta\mu\nu}$. $G_{\mu\nu}$ contains information about the geometry of spacetime. Mathematically, the Einstein tensor, $G_{\mu\nu}$, is given by:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \tag{1.2}$$

where $R_{\mu\nu}$ and R are respectively the Ricci tensor and Ricci scalar, such that $R = g^{\mu\nu} R_{\mu\nu}$ and $g_{\mu\nu}$ is the metric. The stress-energy tensor $T_{\mu\nu}$ describes the density and flux of energy and momentum in spacetime. To balance the expansion of the universe Einstein also added a term on the left hand side of equation 1.1. This term is called the cosmological constant and is symbolized by Λ. Thus the full Einstein field equations take the form:

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

Equation (1.3) describes actually how the presence of mass (energy) determines the geometry of space and conversely how the geometry of space determines the motion of mass (energy). The solutions of these equations are the components of the metric tensor $g_{\mu\nu}$ which describes the spacetime geometry. From equation (1.3), which governs the Universe, we get different equations for its components. One of these are the Friedmann equations which describe the expansion of an isotropic and homogeneous Universe.

1.1.3 The Friedmann-Lemaitre-Robertson-Walker metric

On large spatial scales, where the Universe is homogeneous and isotropic, it can be described by the Friedmann-Lemaitre-Robertson-Walker metric (Scrimgeour et al. **2012**). The metric is given by $g_{\mu\nu}$, so that the line element is $ds^2 = g_{\mu\nu}dx^\nu dx^\mu$. The dimensionless expansion factor $a(t)$ specifies the growth of the Universe with respect to time. It is defined such that at present time $a_0 = 1$. In other words, it specifies the distance $s(t)$ between two fundamental observers (a set of observers in different locations, all of whom are at rest with respect to the matter in their vicinity [3]), with the present-day distance s_0 :

$$s(t) = a(t)s_0\,. \tag{1.4}$$

With a time-dependent expansion factor $a(t)$ the spacetime interval is

$$ds^2 = c^2dt^2 - g_{ij}dx^i dx^j = c^2dt^2 - a^2(t)h_{ij}dx^i dx^j\,. \tag{1.5}$$

where $g_{ij} = a^2(t)h_{ij}$. In a curved space time, we have

$$dx^2 = \frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2sin^2\theta d\Phi^2\,, \tag{1.6}$$

such that the spacetime interval is written as:

$$ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2sin^2\theta d\Phi^2\right]\,. \tag{1.7}$$

The line element ds in equation (1.7) is a geodesic line joining two configurations (events) in the Universe within the framework of Robertson-Walker metric. k (may take one of these values: -1, 0, 1) is the curvature constant which sets the geometry

[3] https://ned.ipac.caltech.edu/level5/Peacock/Peacock3_1.html

of the Universe (k = -1 corresponds to an open Universe, k = 0 corresponds to a flat Universe, k = +1 corresponds to a closed Universe).

1.1.4 Friedmann equations

In general, the stress-energy tensor $T_{\mu\nu}$ could be complicated, but we will focus on the simple case of an ideal perfect fluid, which seems to describe the background Universe (assumed to be nearly homogeneous and isotropic on large scales) quite well (Das et al. **2018**). For an ideal perfect fluid,

$$T_{\mu\nu} = diag\left(-\rho, p, p, p\right), \tag{1.8}$$

where ρ is the energy density of the fluid and p is the pressure. So we now have everything we need to describe the background Universe in the framework of general relativity. The components of the Ricci tensor $R_{\mu\nu}$ and the Ricci scalar are thus given as:

$$R_{00} = -\frac{3\ddot{a}}{a}, \tag{1.9}$$

$$R_{ij} = \left(\frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} + \frac{2k}{a^2}\right)\delta_{ij}, \tag{1.10}$$

$$R = 6\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right), \tag{1.11}$$

where δ_{ij} is the Kronecker delta, which is defined to be 1 for $i = j$ and zero otherwise. By taking the temporal components of equation 1.3 and demanding component by component equality, we have:

$$R_{00} - \frac{1}{2}g_{00}R + \Lambda g_{00} = \frac{8\pi G}{c^4}\rho u_0 u_0, \tag{1.12}$$

$$-3\frac{\ddot{a}}{a} + 3\frac{\ddot{a}}{a} + 3\left(\frac{\dot{a}}{a}\right)^2 + 3\frac{k}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G}{c^2}\rho\,. \tag{1.13}$$

We get another equation 1.14 from R_{ij}, since all of the R_{ij} are the same,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3\frac{p}{c^2}\right) + \frac{\Lambda c^2}{3}\,. \tag{1.14}$$

By introducing the Hubble parameter $H = \frac{\dot{a}}{a}$, as a function of the component of the Universe,

$$\frac{H^2}{H_0^2} = \frac{\Omega_{r,0}}{a^4} + \frac{\Omega_{m,0}}{a^3} + \Omega_{\Lambda,0} + \frac{\Omega_{k,0}}{a^2}\,, \tag{1.15}$$

where

$$\Omega_{\mathrm{k},0} = 1 - (\Omega_{\mathrm{r},0} + \Omega_{\mathrm{m},0} + \Omega_{\Lambda,0})\,.$$

The densities Ω_{r}, Ω_{m}, Ω_{Λ} and Ω_{k} , which respectively are defined as radiation-density, matter-density, dark energy density and curvatures density as follows,

$$\Omega_{\mathrm{m}} = \frac{8\pi G\rho_m}{3H(t)^2}\,,$$

$$\Omega_{\mathrm{r}} = \frac{8\pi G\rho_r}{3H(t)^2}\,,$$

$$\Omega_{\Lambda} = \frac{\Lambda}{3H(t)^2}\,,$$

$$\Omega_{\mathrm{k}} = -\frac{k}{a(t)^2H(t)^2}\,,$$

such that at the initial time ($t = 0$),

$$\Omega_{\mathrm{m},0} = \Omega_{\mathrm{m}}|_{\mathrm{t}=0}, \Omega_{\mathrm{r},0} = \Omega_{\mathrm{r}}|_{\mathrm{t}=0}, \Omega_{\mathrm{k},0} = \Omega_{k}|_{\mathrm{t}=0}, \Omega_{\Lambda,0} = \Omega_{\Lambda}|_{t=0}\,.$$

Equation 1.13 and Equation 1.14, are known as Friedmann equations, and can be

written as,

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k\,c^2}{a^2} + \frac{\Lambda\,c^2}{3}, \tag{1.16}$$

$$\dot{H} + H^2 = -\frac{4\pi G}{3}\left(\rho + 3\frac{p}{c^2}\right) + \frac{\Lambda c^2}{3}\,. \tag{1.17}$$

Friedmann equations lead to the equations of the expansion of the Universe at different epochs : radiation dominated, matter-dominated, and cosmological constant dominated.

1.1.5 Distances

The most used distances in cosmology and radio astronomy are comoving distance, angular diameter distance and luminosity distance. The comoving distance is the distance between two objects that remains constant in time if these objects have the speed of the Hubble flow. On the line of sight, we have $D_{\parallel}$ such that,

$$D_{\parallel} = \frac{c}{H_0}\int_0^z \frac{dz}{E(z)}, \tag{1.18}$$

where $E^2(z) = \dfrac{H^2}{H_0^2} = \dfrac{\Omega_{m,0}}{a^3} + \Omega_{\Lambda,0} + \dfrac{\Omega_{K,0}}{a^2}$. On the transverse direction to the line of sight where two events are at the same redshift but separated by some angle $\delta\theta$, the distance between them is equal to $D_{\perp}\delta\theta$ where the transverse comoving distance $D_{\perp}$ is given by,

$$D_{\perp} = \begin{cases} \frac{c}{H_0}\frac{1}{\sqrt{\Omega_K}} sinh[\sqrt{\Omega_K} D_c \frac{H_0}{c}],\ \Omega_K > 0 \\ D_c,\ \Omega_K = 0 \\ \frac{c}{H_0}\frac{1}{\sqrt{|\Omega_K|}} sin[\sqrt{|\Omega_K|} D_c \frac{H_0}{c}],\ \Omega_K < 0\,. \end{cases} \tag{1.19}$$

The angular diameter distance D_A is defined as the ratio of an object's physical transverse size to its angular size. If we know a object's size, its angular width will tell us its distance from the observer. However, D_A can also be calculated at a given redshift z using the following equation,

$$D_A = \frac{D_\perp}{1+z}\,. \tag{1.20}$$

Another type of distance, the luminosity distance, D_L, is related to the intrinsic luminosity L, is defined as the total energy radiated from an object per second. Luminosity distance is given as follows,

$$D_L = \sqrt{\frac{L}{4\pi S}}\,. \tag{1.21}$$

The relation between luminosity distance D_L and angular diameter distance D_A is given as follows,

$$D_L = (1+z)^2 D_A\,. \tag{1.22}$$

1.1.6 The energy content of the Universe

The Universe contains non-relativistic matter in the form of ordinary, baryonic matter (i.e protons, neutrons, and electrons) as well as dark matter, which is practically pressureless, weakly interacting and cold[4]. Dark matter does interact gravitationally with other forms of matter. It is called dark matter because it does not interact in the electromagnetic window.

The Universe is homogeneous, isotropic and accelerating, so general relativity (GR) is unambiguous about the need for some sort of dark energy source [5]. The Dark energy can be considered as a negative pressure as its equation of state is given

[4] https://www.hindawi.com/journals/aa/2011/604898/

[5] http://www.astro.caltech.edu/~george/ay21/readings/carroll.pdf

by $p = \omega\rho$ where $\omega = -1$. Actually, a number of two-parameter descriptions of ω can give the dark energy parameters as function of redshift $\omega = \omega_0 + \omega_a\,(1 - a) = \omega_0 + \omega_a z/(1 + z)$. The evolution of dark energy density ρ_Λ with the expansion of the Universe for this equation of state parameters is then given by,

$$\frac{\rho_\Lambda}{\rho_{\Lambda,0}} = (1 + z)^{(1+\omega_0+\omega_a)} \exp\left[-\frac{3\omega_a z}{(1 + z)}\right]. \tag{1.23}$$

By substituting the equation of state (with $\omega = -1$) in the second Friedmann equation 1.14 with k= 0 shows that it produces a constant positive acceleration,

$$\frac{\ddot{a}}{a} = \frac{8\pi G}{3}\rho_\Lambda > 0\,. \tag{1.24}$$

That is, the effect of the Cosmological Constant on the Universe expansion is opposite to the one caused by standard gravitating matter. Current observations suggest that the Cosmological Constant is the dominant component of energy content of the Universe, with $\Omega_\Lambda = 0.6889 \pm 0.0056$ (Planck Collaboration et al. **2018**). Although the first evidence for dark matter was discovered in the 1930s, it was not until the early 1980s that astronomers became convinced that most of the mass holding galaxies and clusters of galaxies together is invisible to our instruments. In the standard model of Cosmology, dubbed ΛCDM model, in which the Universe consists of ∼ 5% ordinary baryonic matter, ∼ 27% dark matter, and ∼ 68% dark energy (Planck Collaboration et al. **2018**). Following Klypin et al. **2011**, the Universe is spatially "flat" so its geometry is Euclidian, its expansion is not affected by curvature. The main foundations of the standard ΛCDM model of cosmology are that: the redshifts of the galaxies are due to the expansion of the Universe plus peculiar motions; the cosmic microwave background radiation (in section 1.2.3) and its anisotropies derive from the high energy primordial Universe when matter and radiation became decoupled; and the formation and evolution of galaxies can be explained also in

terms of gravitation within inflation, dark matter, dark energy scenario (López-Corredoira **2017**). In the framework of ΛCDM model, cosmological parameters, which are a set of parameters whose measured values characterize the observed Universe (Baumann et al. 2009), and their physical origin are given in table 1.1

Label	Definition	Physical Origin
Ω_b	Baryon Fraction	Baryogenesis
Ω_{CDM}	Dark Matter Fraction	TeV-Scale Physics
Ω_Λ	Cosmological Constant	Unknown
τ	Optical Depth	First Stars
h	Hubble Parameter	Cosmological Epoch
A_s	Scalar Amplitude	Inflation
n_s	Scalar Index	Inflation

Table 1.1: Following Baumann et al. 2009, the parameters of the concordance cosmology model are listed. We assume a flat Universe, i.e. $\Omega_b + \Omega_{\mathrm{CDM}} + \Omega_\Lambda = 1$; if not, we must include a curvature contribution Ω_k. Likewise, the conventional cosmology includes the microwave background and the neutrino sector. Both these quantities contribute to Ω_{total}, but at a (present-day) level well below Ω_b, the smallest of the three components listed above. The number and energy density of photons is fixed by the observed black body temperature of the microwave background. The parameter h describes the expansion rate of the Universe today,$H_0 = 100$ h Km s^{-1} Mpc^{-1}.

1.2 Structure formation

In the framework of ΛCDM, the structure of the Universe began with small density fluctuations, which are a very small initial deviations from the homogeneous FLRW model, and grew under the action of gravity leading to the Large Scale Structure at the present day. In this section, we develop the Newtonian theory of structure formation to explain how do structures grow in the universe and how can we describe them. At epochs when these deviations are very small, they can be treated as perturbations around the smooth background, while we keep only terms of first

order in perturbation quantities. This is called "linear theory" and the regime where this approach is valid is called "linear regime". In Newtonian perturbation theory, which works provided we consider density variations on length scales smaller than the Hubble length c/H(t) and weak gravitational field, a universe filled with self-gravitating fluid is governed by the basic hydrodynamical equations of Newtonian physics (Dodelson 2003):

$$\partial_t \rho + \nabla \cdot (\rho \bar{v}) = 0\,, \tag{1.25}$$

$$\partial_t \bar{v} + (\bar{v} \cdot \nabla)\bar{v} = -\frac{1}{\rho}\nabla P - \nabla \phi\,, \tag{1.26}$$

$$\nabla^2 \phi = 4\pi \mathrm{G}\rho\,. \tag{1.27}$$

Where $\rho(t, \boldsymbol{r})$ is the matter density, $v(t, \boldsymbol{r})$ is the velocity field, $P(t, \boldsymbol{r})$ is the pressure and $\phi(t, \boldsymbol{r})$ represents the gravitational potential. The first equation 1.25 is called the continuity equation which describes the conservation of mass. The second equation 1.26 is the Euler equation, which follows from Newton's second law applied to the fluid. The third equation 1.27 is the Poisson equation which relates the gravitational potential ϕ, in Euler equation, to the distribution of mass. We use the homogeneous solutions to equations 1.25, 1.27, $\bar{\rho}(t)$, $\bar{v}(t)$, $\bar{P}(t)$, $\bar{\phi}(t)$ and then add small, spatially dependent perturbations,

$$\begin{aligned} \rho &= \bar{\rho} + \delta\rho\,, \\ v &= \bar{v} + \delta v\,, \\ P &= \bar{P} + \delta P\,, \\ \phi &= \bar{\phi} + \delta\phi\,. \end{aligned} \tag{1.28}$$

We start with a static space while neglecting the effect of gravity ($v = 0$ and $\phi = 0$) so the background terms, $\bar{\rho}$ and $\bar{P}$, are constant. For first order perturbations, the continuity and Euler equations become respectively,

$$\partial_t \delta\rho + \nabla \cdot (\bar{\rho}\delta v) = 0\,, \tag{1.29}$$

and

$$\bar{\rho}\partial_t \delta v + \nabla \delta P = 0. \tag{1.30}$$

Combining the partial time derivative of continuity equation and the gradient for the Euler equation:

$$\begin{aligned} \partial_t^2 \delta\rho &= -\partial_t \nabla \cdot (\bar{\rho}\delta v)\,, \\ \partial_t \nabla \cdot (\bar{\rho}\delta v) + \nabla^2 \delta P &= 0 \\ \implies \partial_t^2 \delta\rho &= \nabla^2 \delta P. \end{aligned} \tag{1.31}$$

Since we have four unknowns, but only three equations we must impose an extra condition for the system. We solve this by restricting to the adiabatic perturbations, for which the pressure perturbation is related to the density perturbation by the speed of sound c_s in the medium: $\delta P = c_s^2 \delta\rho$. Equation (1.31) becomes,

$$\partial_t^2 \delta\rho - c_s^2 \nabla^2 \delta\rho = 0. \tag{1.32}$$

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

$$\partial_t^2 \delta\rho - c_s^2 \nabla^2 \delta\rho = 4\pi G \bar{\rho} \delta\rho, \tag{1.33}$$

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

$$\mathbf{r} = a\mathbf{x}. \tag{1.34}$$

The velocity field can be described by

$$\mathbf{v} = H\mathbf{r} + \mathbf{u}, \tag{1.35}$$

with $H\mathbf{r}$ being the Hubble flow and **u** the proper velocity. The gradient in co-moving coordinates becomes

$$\nabla_{\mathbf{r}} = a^{-1}\nabla_{\mathbf{x}} \tag{1.36}$$

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

$$\left(\frac{\partial}{\partial t}\right)_{\mathbf{r}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} + \left(\frac{\partial \mathbf{x}}{\partial t}\right)_{\mathbf{r}} \cdot \nabla_{\mathbf{x}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} + \left(\frac{\partial a^{-1}\mathbf{r}}{\partial t}\right)_{\mathbf{r}} \cdot \nabla_{\mathbf{x}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} - H\mathbf{x} \cdot \nabla_x. \tag{1.37}$$

Substituting Eq. 1.36 and Eq. 1.37 into Eq. 1.25, we obtain

$$\left[\frac{\partial}{\partial t} - H \cdot \nabla\right][\bar{\rho}(1+\delta)] + \frac{1}{a}\nabla \cdot [\bar{\rho}(1+\delta)(Ha\mathbf{x} + \mathbf{v})] = 0 \tag{1.38}$$

where

$$\delta \equiv \frac{\delta\rho}{\bar{\rho}} \tag{1.39}$$

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

$$\frac{\partial \bar{\rho}}{\partial t} + 3H\bar{\rho} = 0, \tag{1.40}$$

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

$$\left[\frac{\partial}{\partial t} - H \cdot \nabla\right][\bar{\rho}\delta] + \frac{1}{a}\nabla \cdot [\bar{\rho}Ha\mathbf{x}\delta + \bar{\rho}\mathbf{v}] = 0, \tag{1.41}$$

which can be re-written as

$$\left[\frac{\partial \bar{\rho}}{\partial t} + 3H\bar{\rho}\right]\delta + \bar{\rho}\frac{\partial \delta}{\partial t} + \frac{\bar{\rho}}{a}\nabla \cdot \mathbf{v}. \tag{1.42}$$

The first term is zero by Eq. 1.40, therefore

$$\dot{\delta} = -\frac{1}{a}\nabla \cdot \mathbf{v}. \tag{1.43}$$

A similar process for the Euler equation (Eq. 1.26) gets us:

$$\dot{\mathbf{v}} + H\mathbf{v} = -\frac{1}{a\bar{\rho}}\nabla\delta P - \frac{1}{a}\nabla\delta\phi, \tag{1.44}$$

and the Poisson equation (Eq. 1.27) becomes

$$\nabla^2\delta\phi = 4\pi G a^2\bar{\rho}\delta. \tag{1.45}$$

Combining the partial time derivative of Eq. 1.43 with the gradient dot product of Eq. 1.44 and Eq. 1.45 we obtain the time evolution for perturbations,

$$\ddot{\delta} + 2H\dot{\delta} + \left(\frac{c_s^2}{a^2} - 4\pi G\bar{\rho}\right)\delta = 0\,. \tag{1.46}$$

In the comoving wave, $\boldsymbol{k}$, space equation 1.46 becomes:

$$\ddot{\delta_k} + 2H\dot{\delta_k} + \left(\frac{k^2c_s^2}{a^2} - 4\pi G\bar{\rho}\right)\delta_k = 0\,. \tag{1.47}$$

In the absence of expansion, $H = 0$, this equation tells us that the growth of δ is

governed by the perturbation in the gravitational potential (first term on the right-hand side, which comes from the Poisson equation for $\delta\phi$) and the pressure (the second term on the right-hand side, where the sound speed reveals that this term has its origin in the pressure perturbation). Gravity makes δ grow, while pressure tries to prevent it. One can show that the perturbations grow exponentially with time or oscillate as sound waves depending on whether their wave number is greater than or less than the Jeans wave number,

$$k_{\mathrm{J}} = a(t)\frac{\sqrt{4\pi G\bar{\rho}}}{\mathrm{c_s}}\,. \tag{1.48}$$

For $k > k_{\mathrm{J}}$ we have sound waves, for $k < k_{\mathrm{J}}$ we have collapse. The expansion adds a sort of friction term on the left-hand side: The expansion of the Universe slows the growth of perturbations down. In linear theory the growth of initial perturbations is self-similar. The overdense regions ($\delta > 0$) increase their density their density contrast over the course of time, while underdense regions ($\delta < 0$) decrease their density contrast.

In the framework of general relativity, we consider the gravitational potential $\phi(\vec{k}, a)$ that is solution of the Bardeen equation (Appendix B.1) and proportional to its initial value $\Phi_{\mathrm{p}}\vec{k}$ back to inflation period. It is written as from the book of Dodelson 2003 :

$$\phi(\vec{k}, a) = \phi_{\mathrm{p}}\vec{k} \times [\text{ Transfer function (k)}] \times [\text{ Growth function (a)}]\,, \tag{1.49}$$

where the evolution of the perturbations during the transition from radiation-dominated era to matter-dominated era given in figure 1.2 is described by the "Transfer function (k)" in equation (1.49) is given by :

$$\mathrm{T(k)} = \frac{\phi(k, a_{\mathrm{late}})}{\phi_{\mathrm{Large-scale}}(k, a_{\mathrm{late}})}\,. \tag{1.50}$$

Here a_{late} is related to the matter dominated epoch. And for the different values of mode k we have: T(k) $\sim$ 1 at very large scales (where $k \to 0$) and T(k) $\sim$ 0 at very small scales (where $k \to \infty$).

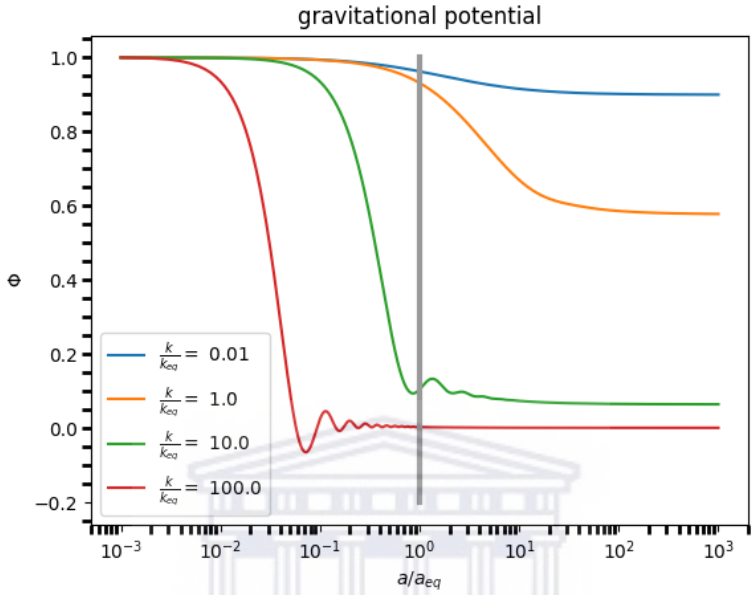


Figure 1.2 The evolution of gravitational potential Φ for 3 different modes. The wavenumber is indicated by the label and the epoch at which the mode enters the Hubble radius is indicated by a small arrow. The top most curve is for a mode which stays outside the Hubble radius for most of its evolution. The other two modes show the decay of ϕ after the mode has entered the Hubble radius in the radiation dominated epoch.

1.2.1 The dark matter power spectrum

In this section we outline the methods used to measure the power spectrum from the distribution of dark matter particles. We start from the definition of the power spectrum. Let $\rho(\boldsymbol{x})$ be the dark matter density field and $\overline{\rho}$ the mean density. Then

the density contrast $\delta(\boldsymbol{x})$ can be expressed as

$$\delta(\boldsymbol{x}) = \frac{\rho(\boldsymbol{x}) - \overline{\rho}}{\overline{\rho}} . \tag{1.51}$$

From statistical homogeneity, $\langle\rho(\boldsymbol{x})\rangle = \overline{\rho}$ and $\langle\delta\rangle = 0$. As we saw in section 1.2, we can use the linear theory to predict the evolution of an overdensity field. To compute the power spectrum, we introduce the two point correlation function $\xi(\boldsymbol{x_1}, \boldsymbol{x_2}) = \langle\delta(\boldsymbol{x_1})\delta(\boldsymbol{x2})\rangle$ that also allows one to compute the variance of any linear function of the random field. $\xi(\boldsymbol{x_1}, \boldsymbol{x_2})$ is positive if the density perturbation has the same sign at both $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$, and negative when there is overdensity at one and underdensity at the other. Thus, it probes how density perturbations at different locations are correlated with each other. Due to statistical homogeneity, $\xi(\boldsymbol{x_1}, \boldsymbol{x_2})$ can only depend on the difference $\boldsymbol{r} = \boldsymbol{x_2} - \boldsymbol{x_1}$, so we redefine ξ as

$$\xi(\boldsymbol{r}) = \langle\delta(\boldsymbol{x})\delta(\boldsymbol{x} + \boldsymbol{r})\rangle . \tag{1.52}$$

The Fourier transform of $\xi(\boldsymbol{r})$ is the power spectrum,

$$P(k) = \int d^3\boldsymbol{r} \exp(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r})\xi(\boldsymbol{r}), \tag{1.53}$$

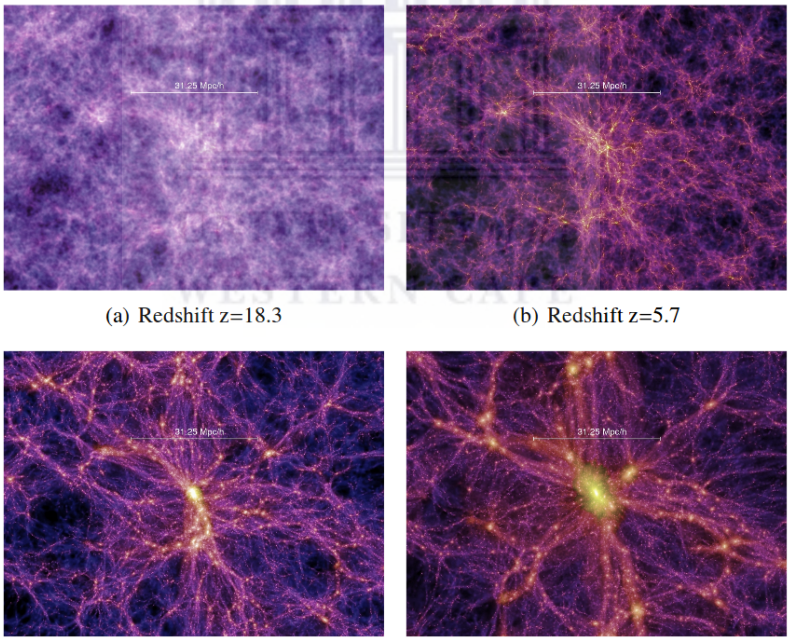
which is related to the density field in k-space by,

$$\langle\delta(\boldsymbol{k_1})\delta(\boldsymbol{k_2})\rangle = P(k_1)(2\pi)^3\delta_D(\boldsymbol{k_1} + \boldsymbol{k_2}), \tag{1.54}$$

where δ_D is the Dirac delta function.

1.2.2 The Large Scale Structure

The Large Scale Structure (LSS) of the Universe is the filamentary structure made by dark matter, also commonly called the cosmic web. Its origin arises from the CMB anisotropies: some places in the Universe were less dense while others were overdense. The latter accreted matter while the former were depleted. It underwent a hierarchical process that is shown in Figure 1.3: it shows the evolution of the LSS from redshift $z = 18$ to redshift $z = 0$. While the dark matter distribution is relatively homogeneous at redshift 18 it becomes much more heterogeneous as redshift decreases. In addition, we can see that the most dense region got more and more dense until forming an extremely dense knot that is fed by many filaments. On large scales, this evolution process is linear but on smaller scales, the scales of a cluster roughly, the evolution process becomes highly non-linear.



(a) Redshift z=18.3 (b) Redshift z=5.7

(c) Redshift z=1.4 (d) Redshift z=0

Figure 1.3 The cosmic web from the Millennium Simulation of the ΛCDM at several redshifts. It shows the evolution of the distribution of dark matter.

To probe the large-scale structure and to understand the patterns within it, one needs galaxy redshift surveys. This remains one of the key tools being used to get cosmological information and to reveal the structure of the observable Universe. We present in figure 1.4 a map of the large-scale structure of the Universe from the Sloan Digital Sky Survey (SDSS). In the image, each dot is a galaxy and the colour bar shows the local density.

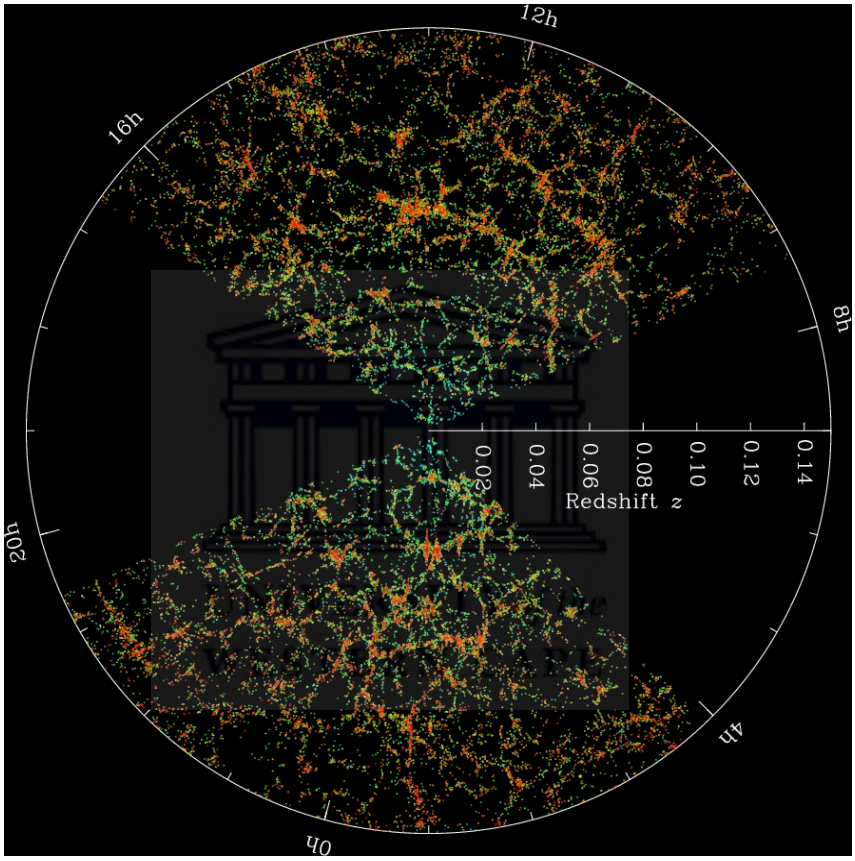


Figure 1.4 Slices through the SDSS 3−dimensional map of the distribution of galaxies. Earth is at the center, and each point represents a galaxy, typically containing about 100 billion stars. Galaxies are colored according to the ages of their stars, with the redder, more strongly clustered points showing galaxies that are made of older stars. The outer circle is at a distance of two billion light years. The region between the wedges was not mapped by the SDSS because dust in our own Galaxy obscures the view of the distant Universe in these directions. Image from SDSS website [6].

1.2.3 The Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is a blackbody radiation which is the afterglow of the Big bang. The early Universe consists of a hot plasma of photons, electrons and baryons: the primordial soup. It is so hot and dense that the mean free path of photons was extremely small, therefore the Universe was opaque. As the Universe expanded and cooled adiabatically, electrons combined with protons and formed hydrogen atoms. This recombination event happened when the temperature was around 3000 K and the Universe was approximately 379,000 years old. Photons began to travel freely through space, resulting in the decoupling of matter and radiation. These photons have continued cooling with the expansion of the Universe: the CMB frequency spectrum is now at a temperature $T = 2.725$ K. While it is almost a perfect blackbody, it displays anisotropies with $\Delta T/T \sim 10^{-5}$. These anisotropies are the seeds of the large scale structure (see Sect. 1.2.2). Those anisotropies have been measured by several telescopes over the past 30 years: the latest one is the Planck satellite which mapped the sky almost 5 times over 4 years from the far-infrared to millimeter wavelengths. The resulting map of the anisotropies is shown in figure 1.5. Those anisotropies contain information on the history of the Universe, therefore on the cosmological model. One of the main aims of the Planck satellite was to measure the power spectrum of the anisotropies which is shown in figure 1.6. This power spectrum is able to constrain many of the cosmological parameters that describe the formation and the evolution of Universe: for instance the fraction of baryonic matter, dark matter and dark energy, etc. Although, dark matter and dark energy is not well understood.

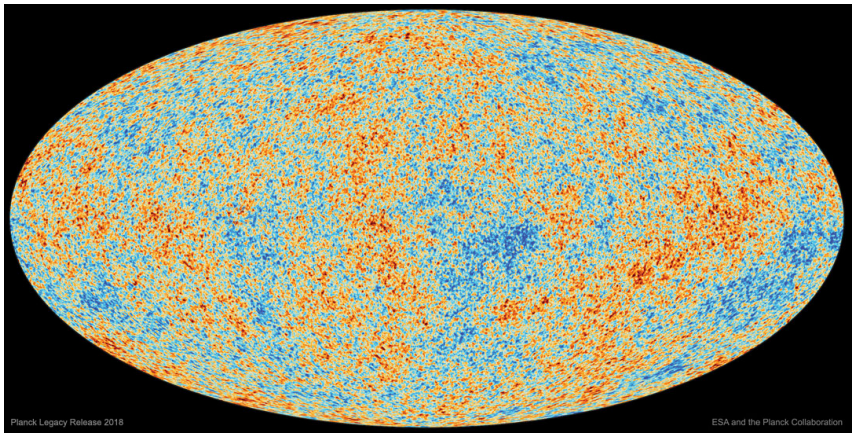


Figure 1.5 Map of the Cosmic Microwave Background with Planck satellite. Image credit: European Space Agency, Planck Collaboration. Taken from NASA website [7].

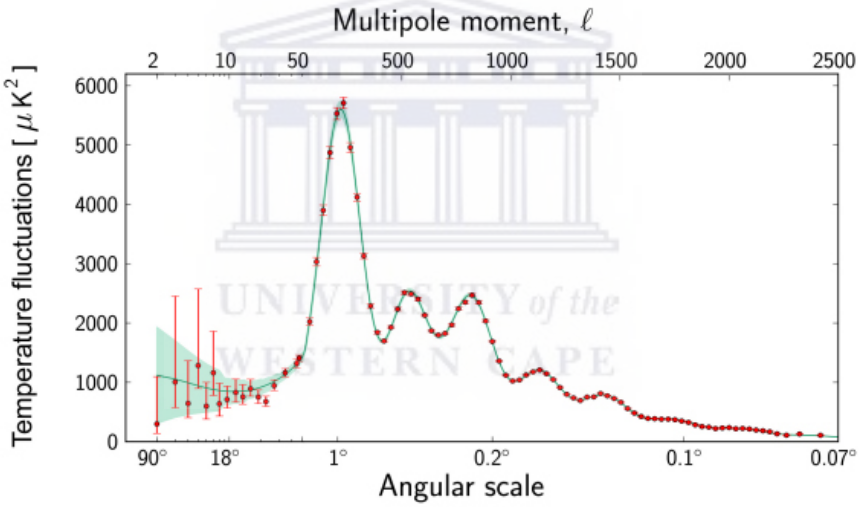


Figure 1.6 This figure depicts the power spectrum of temperature fluctuations in the Cosmic Microwave Background as a function of angular scale. Image credits: ESA and the Planck Collaboration[8].

1.2.4 Baryon Acoustic Oscillations

Before recombination, the Universe was made of a hot and dense plasma in which there were acoustic waves: the radiation pressure from the CMB prevents the ionized gas from clustering and this pressure leads to relativistic sound waves. At decoupling, those waves were frozen in the matter distribution, they are the wiggles seen in the power spectrum of the CMB. The first peak at 1 degree is the scale of the sound horizon $s \sim 150/(1+z)$ Mpc and it is a "standard ruler" of the Universe. The scale can be computed in the following way:

$$s = \int_0^{t_{rec}} c_s(1+z)dt = \int_{z_{rec}}^{\infty} \frac{c_s dz}{H(z)}, \tag{1.55}$$

where $c_s = [3(1+3\rho_b/\rho_\gamma)]^{-1/2}$ is the sound speed. Accurately measuring the scale of the first peak enables constraints on the Hubble parameter, H, and the knowledge of the redshift evolution of H leads to constraints on the acceleration of the expansion of the Universe, and therefore on dark energy. The BAO scale can also be measured in the 2-point correlation function of any biased tracer of matter, the real-space counterpart of the power spectrum. It has been measured for the first time in the SDSS data by Eisenstein et al. **2005** and is shown in Fig. 1.7. The BAO peak at a redshift z appears at an angular separation $\Delta\theta = r_d/[(1+z)d_A(z)]$ and at a redshift separation $\Delta z = r_d/d_H(z)$, where d_A and $d_H = c/H$ are the angular and Hubble distances, respectively, and r_d is the sound horizon at the drag epoch. For instance, if we respectively know the size of an object along the line of sight ($L_\parallel$) and across the sky ($L_\perp$) then we can measure the Hubble parameter H and the angular diameter distance d_A separately, such that:

$$L_\parallel = \frac{c\Delta z}{H(z)}, \tag{1.56}$$

$$L_{\perp} = (1+z)d_A(z)\Delta\theta\,. \tag{1.57}$$

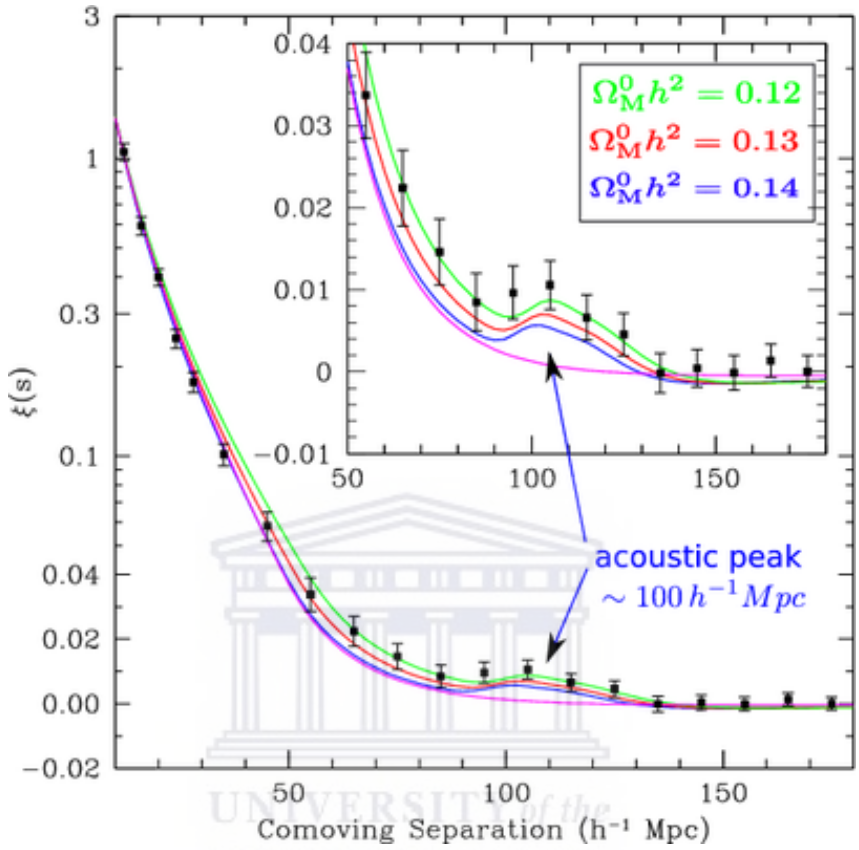


Figure 1.7 The Baryon Acoustic Peak (BAP) in the correlation function. The BAP is visible in the clustering of the SDSS LRG galaxy sample, and is sensitive to the matter density (shown are models with $\Omega_m h^2 = 0.12$ **(green)**, 0.13 **(red)** and 0.14 **(blue)**, all with $\Omega_b h^2 = 0.024$). The purple line without a BAP is the correlation function in the pure CDM model. From Eisenstein et al. **2005**.

Figure 1.8 shows that BAOs are a standard ruler, which allows measurements of cosmological distances as functions of redshift to be performed, because we can find both $d_A(z)$ and $H(z)$ separately.

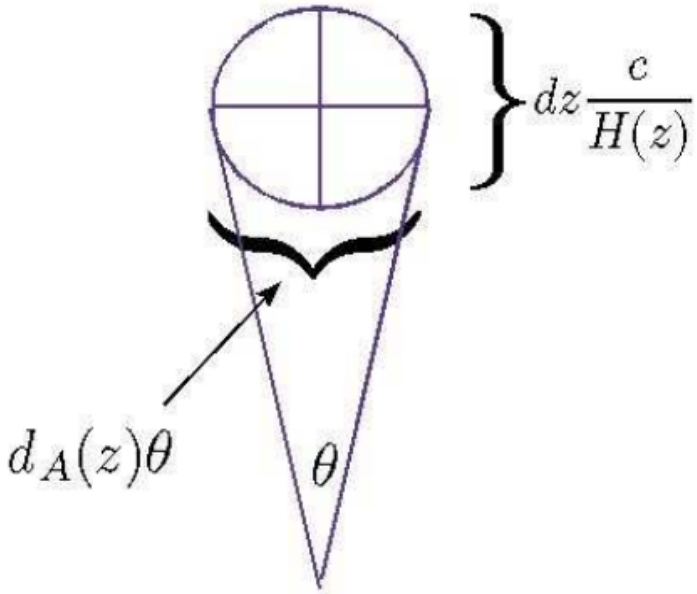


Figure 1.8 The radial length of an object is given by $c\frac{dz}{H(z)}$ where dz is the difference in redshift between the front and back of the object while the transverse size of the object is $dA(z)\theta$, where θ is its angular size. If, as in the case of BAO, one can theoretically determine the diameter, one has the bonus of finding $d_A(z)$ and $H(z)$ separately. From the paper of Bassett and Hlozek **2009**.

1.3 Cosmology with neutral hydrogen

1.3.1 Neutral hydrogen in the Universe

Hydrogen constitutes about 75% of the elemental mass of the Universe [9] which is about 16% of the total matter (dark matter and baryonic matter). One of the most promising ways to observe the Universe is by detecting the 21cm emission from cosmic neutral hydrogen (HI) through radio-telescopes (Zamudio-Fernandez

[9] http://www.elementsdatabase.com/Hydrogen-H-1-element/

et al. **2019**). It has been argued that HI can be used as a tracer of the large scale structure and observations can be used to constrain cosmological parameters with a special emphasis on observations of the baryon acoustic oscillations in the power spectrum (Bagla et al. **2010**, Wyithe and Loeb **2009**, Wyithe et al. **2009**). Following Zwaan et al. **2005**, the HI content can be estimated more directly through emission in the Hyperfine transition at low redshifts and the observations in the local Universe indicate a much lower neutral Hydrogen content than seen at $z > 1$.

1.3.2 Observing neutral hydrogen

Neutral hydrogen has some physical properties that enable its observation. The most important of its properties is the spin flip transition: neutral hydrogen consists of one proton and one electron that have spin, therefore they have magnetic moments. The spin of the electron can have two states, "up" or "down". Similarly, the proton also has its spin either "up" or "down". It yields to four dynamical states of the total spin. HI has a spin one-half which only allows two possible states, parallel and anti-parallel.

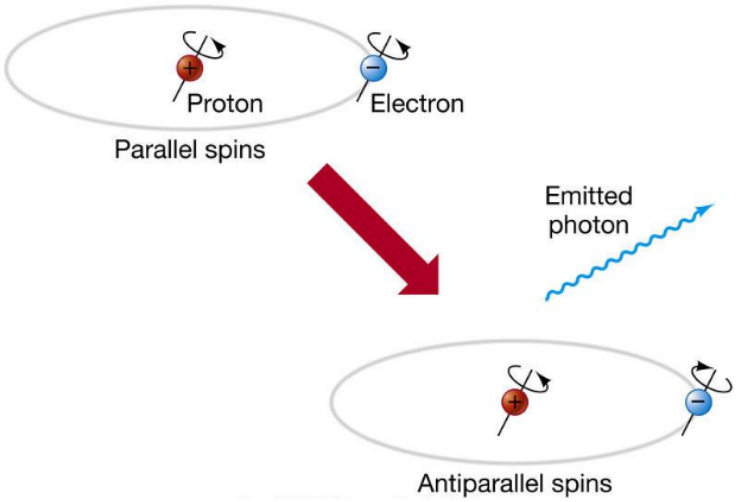


Figure 1.9 Spin-flip transition from parallel to anti-parallel which results in the emission of a photon at a wavelength of 21 cm. Image credits: Pearson Prentice Hall, Inc.

This is a hyperfine structure. The anti-parallel state is more stable than the parallel one and, thus, its energy is lower. So the transition from the parallel state to the anti-parallel one emits the following amount of energy $E \sim 10^{-6}$ eV (Bradt 2008). Its relaxation gives an emission line at the specific frequency $\nu \approx 1420.4$ MHz. Such a photon with radio wavelengths experiences minor absorption while travelling through the atmosphere. Therefore it can be observed on Earth.

HI 21 cm is generally seen in emission but it can also be seen in absorption against continuum radio sources. Let us consider a cloud of spin temperature T and opacity τ_ν placed in front of a radio source with temperature T_s and the brightness temperature of the antenna T_b,

$$T_b(\nu) = T_s e^{-\tau_\nu} + T(1 - e^{-\tau_\nu}) \,. \tag{1.58}$$

Then, the excess brightness temperature from the cloud is,

$$\Delta T_b(\nu) = T_b(\nu) - T_s = (T - T_s)(1 - e^{-\tau_\nu}) \,. \tag{1.59}$$

So, $\Delta T_b(\nu) = T - T_s$ if $\tau_\nu >> 1$ and $\Delta T_b(\nu) = (T - T_s)\tau_\nu$ if $\tau_\nu << 1$. Hence, we have an emission line if:

$$T >> T_s \,, \tag{1.60}$$

and we have an absorption line if :

$$T << T_s \,. \tag{1.61}$$

Here we will only consider HI in emission.

1.3.3 HI line Intensity Mapping

HI line intensity mapping (IM) is a novel observation technique capable of mapping the large-scale structure of the Universe in three dimensions by using HI as a biased tracer of the dark matter density field. This is achieved by measuring the intensity of the redshifted 21 cm line over the sky in a range of redshifts without the requirement to resolve individual galaxies. The observable is the HI integrated intensity from hundreds of galaxies in one single large voxel (3D pixel). Therefore, instead of getting galaxies as dots we get a map of fluctuations of intensity as shown in Figure 1.10. Such a technique allows the observations of volume scales of the Universe that have not been probed yet as shown in Figure 1.11. The IM survey of SKA1-mid will observe a volume 3 times larger than the one of the galaxy survey of the Euclid satellite. This volume gives access to extremely large scales. On those scales, inflation can be tested due to accurate measurements of spatial curvature and

primordial non-Gaussianity. On linear scales, the BAO peaks and the redshift space distortions can be measured to determine the expansion and growth of the Universe. 21 cm IM can also be used to probe the epoch of reionization which cannot be obtained by means other than the use of spectral emission lines (Morales and Wyithe **2010**). IM studies can be used in conjunction with optical or infrared galaxy surveys to study the physics of the late-time Universe. The idea is to combine different types of surveys so that their systematics can be mitigated and their constraining power augmented (e.g., Pourtsidou **2016**; Wolz et al. **2017**; Hall and Bonvin **2017**; Carucci 2018).

The advantages of HI IM compared to optical galaxy surveys are: (i) the redshift comes directly from the measurement of the redshifted 21 cm line; and (ii) HI is expected to be a good tracer of mass with minimal bias (e.g., Padmanabhan et al. **2015**; Pénin et al. **2018**).

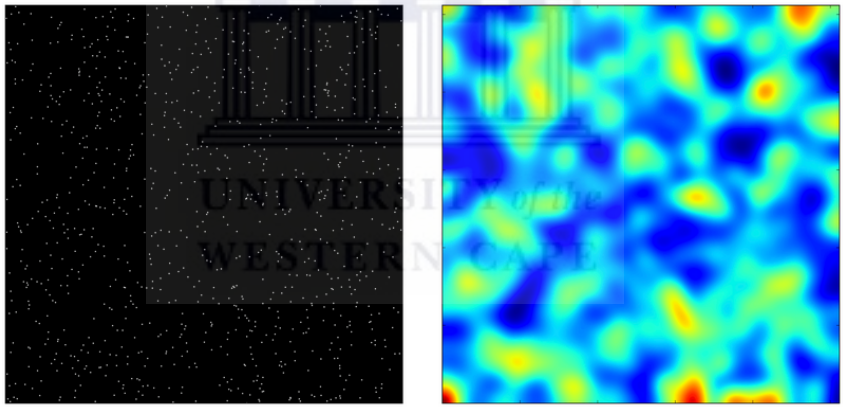


Figure 1.10 The left panel shows the distribution of galaxies, represented by white dots. While on the right panel we have the intensity map such that pixel intensity corresponds to joint emission from multiple galaxies. With the courtesy of Francisco Villaescusa https://franciscovillaescusa.github.io/im.html.

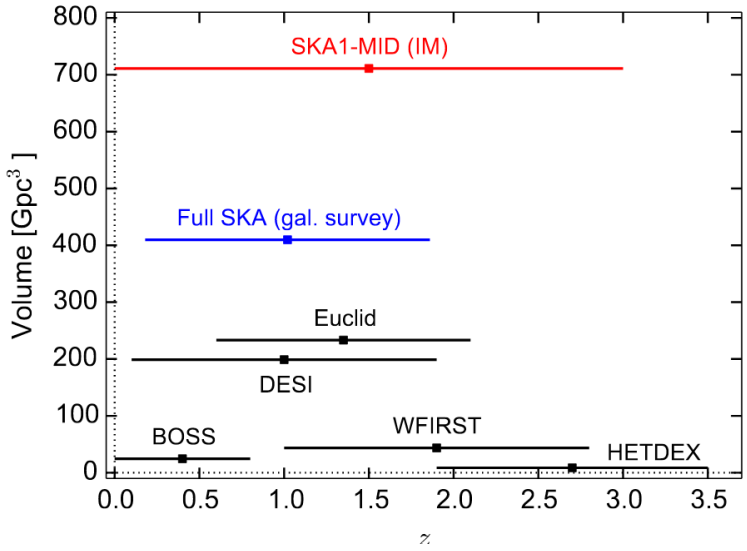


Figure 1.11 This plot shows survey volumes and redshift range for various current and future surveys. From the paper of of Santos et al. 2015.

1.3.4 HI intensity mapping experiments

We present in this section several experiments that are or will carry out HI intensity mapping surveys. The Square Kilometer Array (SKA) is a radio observatory built both in South Africa and Australia. It is comprises of two instruments, SKA-LOW at low frequencies in Australia and SKA-Mid in South Africa. SKA1-LOW with frequencies between 50 and 350 MHz is dedicated to the study of the Epoch of Reionization. The HI line beyond $z \sim 3$. SKA1-Mid is dedicated to the study of galaxies and the large scale structure up to that redshift. There are numerous science cases with the SKA, amongst others: detection of radio transients, heliospheric physics, cosmic magnetism, galaxy evolution, testing Einstein general relativity and cosmology. Its first light is planned for the 2020s.

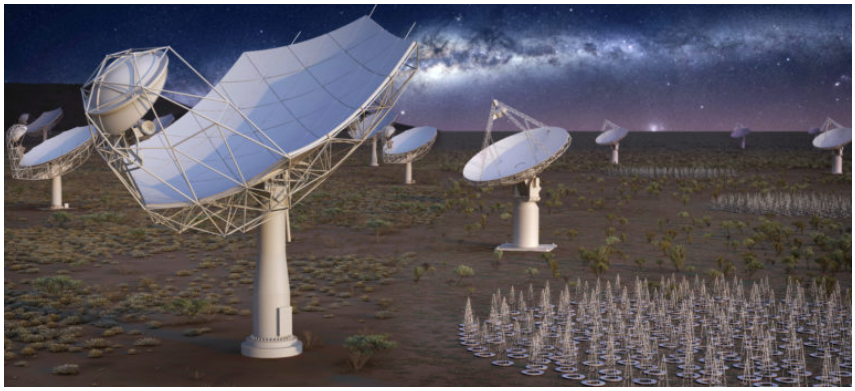


Figure 1.12 An artistic impression of SKA telescope. From the SKA website [10].

MeerKAT is an array of 64 13.5-meter dishes in the Karoo Desert in South Africa which observes in mid radio wavelength. It is the pathfinder of SKA-1 Mid and its first light was in 2018. An intensity mapping survey with MeerKAT, known as MeerKLASS is planned (Santos et al. 2017): 4000 $\deg^2$ over 4000 hours with the L-band receivers (900 - 1670 MHz). The main aim is measuring the BAO peaks with unprecedented accuracy.

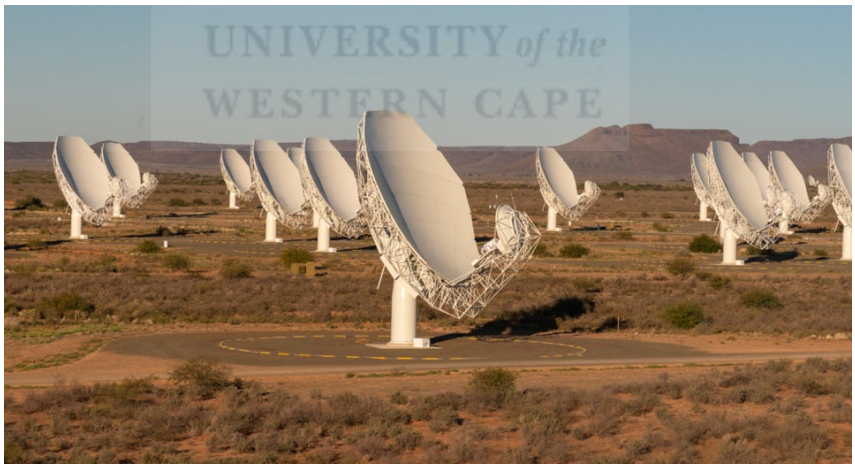


Figure 1.13 Picture of the Meerkat Radio Telescope. From reddit website[11].

The Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX) is a radio

telescope array that will map nearly all of the southern sky in radio continuum and neutral hydrogen line emission over a frequency range of 400 to 800 MHz. These wavelengths correspond to a redshift range of 0.8 to 2.5 and HIRAX's primary goal is to measure the BAO wiggles over this redshift range. It will be an interferometer that comprises roughly 1000 six-meter dishes placed in a close-packed, redundant configuration, and will be deployed at the SKA site in the Karoo desert. HIRAX is highly complementary to the Canadian Hydrogen Intensity Mapping Experiment (CHIME). Both have the same primary purpose but with different observing technologies. CHIME consists of four adjacent 20 m $\times$ 100 m cylindrical reflectors oriented north-south in the same wavelength range. The telescope detects and measures HI.

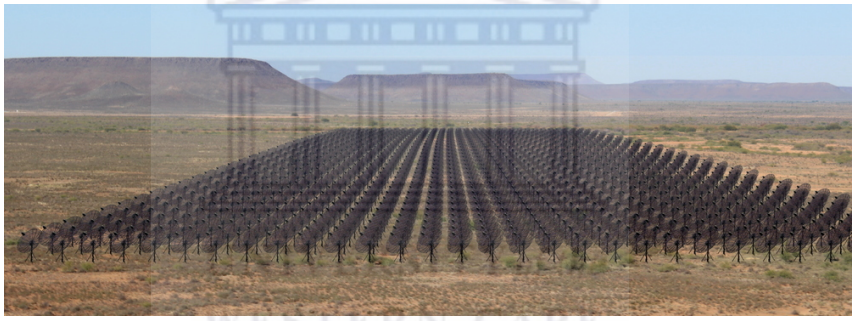


Figure 1.14 Image of the 1000-dish HIRAX telescope. From Quantum leap Africa website [12] .

Figure 1.15 Image of the Canadian Hydrogen Intensity Mapping Experiment (CHIME) which has been officially granted Square Kilometre Array (SKA) pathfinder status since 21 November 2018. From SKA telescope website [13] .

Lastly the BAO from Integrated Neutral Gas Observations (BINGO) telescope is a 40-meter single dish telescope which is currently in construction in Brazil. It will operate in the frequency range from 0.96 GHz to 1.26 GHz which is equivalent to redshifts 0.13 and 0.48. Its primary purpose is also the measurement of the BAO wiggles with intensity mapping. Its first light is planned in the 2020s.

Figure 1.16 On the left is a picture of the Castrillon quarry which is the proposed site of the BINGO telescope. On the right is an engineering drawing of the telescope with primary, secondary and focal plane array of horns in position. The black box in the right hand figure is approximately the plane of the picture on the left. From the paper of of Battye et al. **2016** about the update on the BINGO 21 cm intensity mapping experiment.

Chapter 2

The HI Power Spectrum

The power spectrum can be seen as a Fourier transform of the correlation function of density fluctuations. What is measured in HI intensity mapping is a combination of the cosmological 21cm brightness temperature power spectrum and the instrumental noise. In the Post-reionization era, one can make the assumption that all HI resides in the dark matter halos (Villaescusa-Navarro et al. **2015**). Under this assumption, and following- the spirit of the halo model in Cooray and Sheth **2002**, we can then predict the shape and amplitude of the HI power spectrum, in real space, if we have the following ingredients: the halo mass function, $n(M, z)$, the halo bias, $b(M, z)$, the linear matter power spectrum $P_{\rm lin,m}(k)$, and the HI mass function $M_{\rm HI}(M, z)$ which gives the average HI mass in a dark matter halo of mass M at redshift z. On large, linear, scales the HI power spectrum in real-space does not depend on the $M_{\rm HI}(r|M, z)$ function, but only on $M_{\rm HI}(M, z)$, and it is given by,

$$P_{\rm HI}(k, z) = T^2 b^2_{\rm HI}(z) P_{\rm m}(k, z), \tag{2.1}$$

where the HI bias, $b_{\rm HI}(z)$ is given by,

$$b_{\rm HI}(z) = \frac{\int_0^\infty n(M,z)b(M,z)M_{\rm HI}(M,z)dM}{\int_0^\infty n(M,z)M_{\rm HI}(M,z)dM}\,. \tag{2.2}$$

2.1 The linear HI power spectrum

The average signal on the sky is the HI brightness temperature that writes (Battye et al. **2013**)

$$\overline{T}_b(z) = 566\,h\left(\frac{H_0}{H(z)}\right)\left(\frac{\Omega_{\rm HI}(z)}{0.003}\right)(1+z)^2\ \ \mu{\rm K}, \tag{2.3}$$

where the HI density fraction is defined as $\Omega_{\rm HI} = \rho_{\rm HI}/\rho_{c,0}$ with $\rho_{c,0}$ is the critical density of the Universe today. The fluctuating part is

$$T(z,\mathbf{x}) = \overline{T}(z)\,(1+\delta_{\rm m}(\mathbf{x}))\,, \tag{2.4}$$

where $\delta_{\rm m}(\mathbf{x})$ is the matter density fluctuation at position $\mathbf{x}$, hence, in Fourier space,

$$\langle T(z,\mathbf{k})T^\star(z,\mathbf{k}')\rangle = (2\pi)^3 P_{\rm HI}(k,z)\delta^3(\mathbf{k}-\mathbf{k}'), \tag{2.5}$$

where $P_{\rm HI}(k,z)$ is the HI power spectrum. Following Kaiser **1987**, the linear term in redshift space is ,

$$P^{11}_{\rm HI}(k,\mu) = \overline{T}^2\left[b_1 + f\,\mu^2\right]^2\ P^{11}_{\rm m}(k), \tag{2.6}$$

with $\mu = k_\parallel/k$, b_1 the HI linear or first order bias, $P_{\rm m}(k)$ the linear power spectrum of matter, and f the linear growth rate. We compute the former using the transfer function of Eisenstein and Hu **1998** and we use $f(z) = \Omega_{\rm m}(z)^\gamma$ with $\gamma = 0.55$ for ΛCDM (Linder, Peebles **2005**, 1980).

The HI linear bias is

$$b_1^{\rm HI}(z) = \sqrt{\frac{P_{\rm HI}(z)}{P_{\rm lin}(z)}}\,. \tag{2.7}$$

It measures the clustering of HI with respect to the clustering of dark matter and it can be computed analytically in the framework of the halo model. The non-linear contributions to the power spectrum can be derived using perturbation theory.

2.2 The non-linear power spectrum of HI

With a full one-loop derivation of the HI brightness temperature in Perturbation Theory Bernardeau et al. **2002**, the power spectrum of HI in real space at redshift z is ,

$$P_{\rm HI}(z,k) = P_{\rm HI}^{11}(z,k) + P_{\rm HI}^{22}(z,k) + P_{\rm HI}^{13}(z,k)\,, \tag{2.8}$$

where $P_{\rm HI}^{11}(z,k)$ is the linear power spectrum (tree level) while $P_{\rm HI}^{22}(z,k)$ and $P_{\rm HI}^{13}(z,k)$ are the non-linear corrections. For clarity purposes we will not specify the redshift dependence in the following. From Umeh et al. **2016** and Umeh **2017** the three terms of $P_{\rm HI}(k)$ are

$$P_{\rm HI}^{11}(k) = \overline{T}^2\, b_1^2\, P_{\rm m}^{11}(k)\,, \tag{2.9}$$

$$P_{\rm HI}^{22}(k) = \frac{\overline{T}^2}{2}\int \frac{{\rm d}^3k_1}{(2\pi^3)}\Big[b_1\,F_2(\mathbf{k}_1,\mathbf{k}_2) + b_2\Big]^2 \times P_{\rm m}^{11}(k_2)\,P_{\rm m}^{11}(k_1)\,, \tag{2.10}$$

$$P_{\rm HI}^{13}(k) = \overline{T}^2\, b_1\Big\{\left(b_3 + \frac{68}{21}b_2\right)\sigma_\Lambda^2 P_{\rm m}^{11}(k) + b_1\,P_{\rm m}^{13}(k)\Big\}\,, \tag{2.11}$$

where $k_2 = |\mathbf{k}_1 - \mathbf{k}|$, b_1, b_2, and b_3 are the linear, second and third order HI biases, respectively. The latter are the higher order terms of the bias expanded in Taylor series, which means that the HI bias is local. F_2 is the non-linear density kernel

defined in equation 2.25. Finally σ_Λ, the variance of the dark matter field, is

$$\sigma_\Lambda^2 = \int_{k_{\rm min}}^{k_{\rm max}} \frac{{\rm d}^3 k}{(2\pi)^3} P_{\rm m}(k)\,. \tag{2.12}$$

For simplicity, we set $k_{\rm max}$ to the non-linear dispersion scale, $k_{\rm NL} = 0.2\,h(1+z)^{2/(2+n_s)}$ Mpc^{-1} with n_s the spectral index.

In redshift space, the 3D power spectrum of HI on linear and quasi-linear scales at scale k, and μ, the cosine of the angle between the line of sight and the separation vector $\mathbf{k}$, writes

$$P_{\rm HI}(k,\mu) = P_{\rm HI}^{11}(k,\mu) + P_{\rm HI}^{22}(k,\mu) + P_{\rm HI}^{13}(k,\mu)\,. \tag{2.13}$$

Following Kaiser **1987**, the linear term in redshift space is given by equation 2.6 Following Umeh et al. **2016** and Umeh **2017** the one-loop corrections are

$$\begin{aligned} P_{\rm HI}^{22}(k,\mu) &= \frac{\overline{T}^2}{2}\int \frac{{\rm d}^3 k_1}{(2\pi)^3}\Big[b_1\,F_2(\mathbf{k}_1,\mathbf{k}_2) + \mu^2 G_2(\mathbf{k}_1,\mathbf{k}_2) + b_2 + K_R(\mathbf{k}_1,\mathbf{k}_2)\Big]^2 \\ &\times\ P_{\rm m}^{11}(k_2)\,P_{\rm m}^{11}(k_1), \end{aligned} \tag{2.14}$$

$$\begin{aligned} P_{\rm HI}^{13}(k,\mu) &= \overline{T}^2\left(b_1 + \mu^2 f\right)\times\left\{\left[\left(b_3 + \frac{68}{21}b_2\right)\sigma_\Lambda^2 + I_R(k,\mu)\right]P_{\rm m}^{11}(k)\right. \\ &+\ \left.\Big[b_1\,P_{\rm m}^{13}(k) + \mu^2 f\,P_\theta^{13}(k)\Big]\right\}, \end{aligned} \tag{2.15}$$

where $P_{\rm m}^{13}(k)$ and $P_{\theta}^{13}(k)$ are the third order matter power spectrum and velocity field power spectrum, respectively. Their expressions along with that of $I_R(k,\mu)$ are

$$\begin{aligned} P_{\rm m}^{13}(k) &= \frac{1}{252}\frac{k^3}{4\pi^2}P_{\rm m}^{11}(k)\int_0^\infty {\rm d}r\, P_{\rm m}^{11}(kr)\Big[\frac{12}{r^2}-158 \\ &+ 100\,r^2-42\,r^4+\frac{3}{r^3}(r^2-1)^3(7r^2+2)\log\left|\frac{1+r}{1-r}\right|\Big], \end{aligned} \tag{2.16}$$

$$\begin{aligned} P_{\theta}^{13}(k) &= \frac{1}{84}\frac{k^3}{4\pi^2}P_{\rm m}^{11}(k)\int_0^\infty {\rm d}r\, P_{\rm m}^{11}(kr)\Big[\frac{12}{r^2}-82 \\ &+ 4\,r^2-6\,r^4+\frac{3}{r^3}(r^2-1)^3(r^2+2)\log\left|\frac{1+r}{1-r}\right|\Big]. \end{aligned} \tag{2.17}$$

The last component of the $P_{\rm HI}^{13}(k,\mu)$ is

$$\begin{aligned} I_R(k,\mu) &= \frac{k^3}{(2\pi)^2}\int {\rm d}r P^{11}(kr) \\ &\times \mu^2 f([b_2B_1(r)+b_1B_2(r)] \\ &+ \mu^2 f^2\left[b_1B_3(r)+B_4+\mu^2(b_1^2B_5(r)+fB_6(r))\right]), \end{aligned} \tag{2.18}$$

with

$$\begin{aligned} B_1(r) &= \frac{1}{6}, \\ B_2(r) &= \frac{1}{84}\left[-2(9r^4-24r^2+19)+\frac{9}{r}(r^2-1)\log\left(\frac{1+r}{|1-r|}\right)\right], \\ B_3(r) &= -\frac{1}{3}, \\ B_4(r) &= -\frac{1}{336\,r^3}[2(-9\,r^7+33\,r^5+33\,r^3-9\,r) \\ &+ 9(r^2-1)\log\left(\frac{1+r}{|1-r|}\right)], \\ B_5(r) &= \frac{1}{336\,r^3}[2\,r(-27\,r^6+63\,r^4-109\,r^2+9) \\ &+ 9(3\,r^2+1)(r^2-1)\log\left(\frac{1+r}{|1-r|}\right)]. \end{aligned}$$

Lastly, in this framework, the full 1-loop matter power spectrum in real space is

$$P_{\rm m}^{\rm NL}(k) = P_{\rm m}^{11}(k) + P_{\rm m}^{22}(k) + P_{\rm m}^{13}(k), \tag{2.19}$$

where the second order term writes

$$P_{\rm m}^{22}(k) = \frac{1}{2}\int \frac{{\rm d}^3k_1}{(2\pi)^3} F_2^2(\mathbf{k}_1, \mathbf{k}_2) P_{\rm m}^{11}(k_2)\, P_{\rm m}^{11}(k_1)\,. \tag{2.20}$$

In redshift space, the full 1-loop matter power spectrum is

$$\begin{aligned}
P_{\rm m}^{\rm NL}(k,\mu) &= P_{\rm m}^{11}(k,\mu) + P_{\rm m}^{22}(k,\mu) + P_{\rm m}^{13}(k,\mu), && (2.21)\\
P_{\rm m}^{11}(k,\mu) &= \left[1 + f\,\mu^2\right]^2\, P_{\rm m}^{11}(k), && (2.22)\\
P_{\rm m}^{22}(k,\mu) &= \frac{1}{2}\int \frac{{\rm d}^3k_1}{(2\pi)^3}\Big[F_2(\mathbf{k}_1,\mathbf{k}_2) + \mu^2 G_2(\mathbf{k}_1,\mathbf{k}_2) + K_R(\mathbf{k}_1,\mathbf{k}_2)\Big]^2 \\
&\times\ P_{\rm m}^{11}(k_2)\, P_{\rm m}^{11}(k_1), && (2.23)\\
P_{\rm m}^{13}(k,\mu) &= \Big(1 + \mu^2 f\Big)\Big\{I_R(k,\mu)\, P_{\rm m}^{11}(k) + \left[P_{\rm m}^{13}(k) + \mu^2\, f\, P_\theta^{13}(k)\right]\Big\}. && (2.24)
\end{aligned}$$

Finally, several kernels are involved in the computation of the $P_{\rm HI}^{22}(k,\mu)$ term : G_2 induced by peculiar velocities at second order and K_R arises from non-linear mode coupling velocity-velocity and velocity-density (Bernardeau et al. **2002**). Their expressions are:

$$\begin{aligned}
F_2(\mathbf{k}_1,\mathbf{k}_2) &= \frac{5}{7} + \frac{1}{2}\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\left[\frac{k_1}{k_2} + \frac{k_2}{k_1}\right] + \frac{2}{7}\left[\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right]^2, && (2.25)\\
G_2(\mathbf{k}_1,\mathbf{k}_2) &= \frac{3}{7} + \frac{1}{2}\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\left[\frac{k_1}{k_2} + \frac{k_2}{k_1}\right] + \frac{4}{7}\left[\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1k_2}\right]^2, && (2.26)
\end{aligned}$$

$$K_R(\mathbf{k}_1, \mathbf{k}_2) = f\, b_1 \mu_1^2 + f\, b_1\, \mu_2^2 + \mu_1 \mu_2 \left[f\, b_1 \frac{k_1}{k_2} + f\, b_1 \frac{k_2}{k_1} \right] + f^2 \left[2\mu_1^2\, \mu_2^2 + \mu_1 \mu_2 \left(\mu_1^2 \frac{k_1}{k_2} + \mu_2^2 \frac{k_2}{k_1} \right) \right] . \quad (2.27)$$

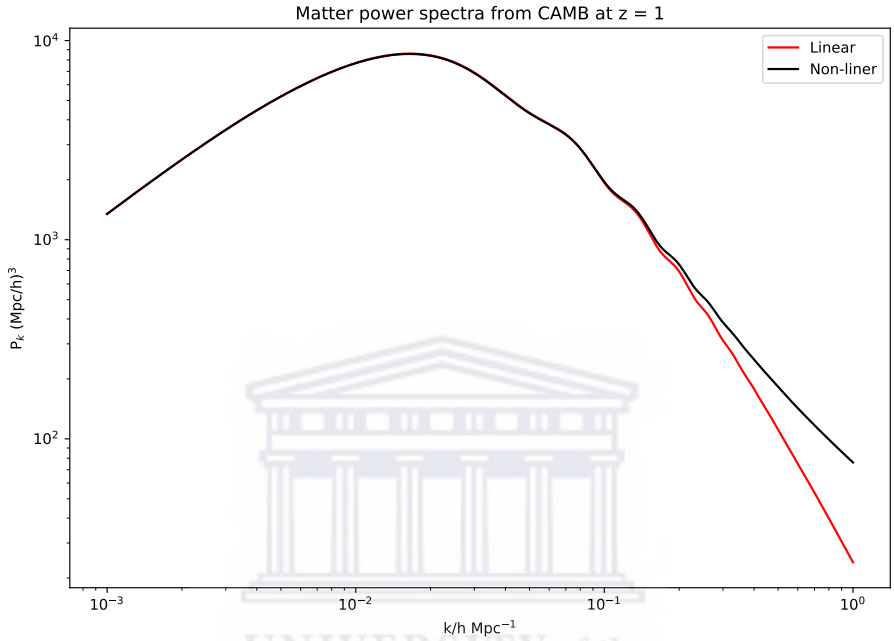


Figure 2.1 This figure shows the matter power spectrum from CAMB at $z = 1$. It shows that the non-linear contribution is only relevant at small scales. The non-linear scales are calculated using the CAMB Halofit option.

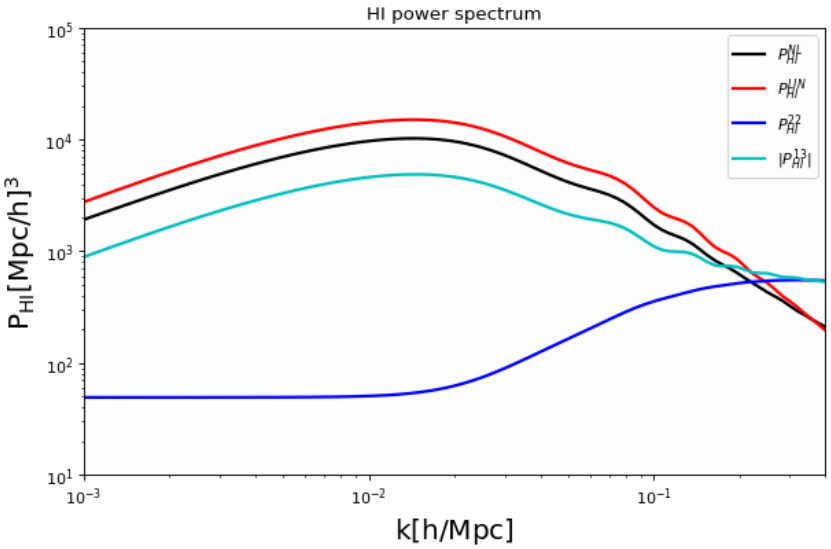


Figure 2.2 This figure shows the components of the HI power spectrum, at $z = 1$, using our perturbation theory calculations. The red solid line is the first order term, $P_{\rm HI}^{11}(k)$, which is the linear power spectrum. The blue solid line,$P_{\rm HI}^{22}(k)$, is the second order non-linear contribution to the total power spectrum. The light blue, $P_{\rm HI}^{13}(k)$, refers to the third order component of the power spectrum. It is negative and therefore removes the amplitude of the linear one. Therefore, the amplitude of the total power spectrum, $P_{\rm HI}^{TOT}(k)$, given by the black solid line is much smaller than the red one. $P_{\rm HI}^{TOT}(k) = P_{\rm HI}^{11}(k) + P_{\rm HI}^{22}(k) + P_{\rm HI}^{13}(k)$.

Fig. 2.1 shows the linear and non-linear power spectra of matter at $z = 1$. We recover that the non-linear contribution is significant only at small scales. Fig. 2.2 shows the linear and non-linear components of the HI power spectrum. Contrary to our expectations, both $P_{\rm HI}^{22}(k)$ and $P_{\rm HI}^{13}(k)$ terms have significant contributions on linear scales. In addition, in the case of HI, the component $P_{\rm HI}^{13}(k)$ is negative which removes power from the power spectrum. As shown by Pénin et al. **2018**, even on linear scales the bias is scale dependent. This depends on the HI biases: $b_n^{\rm HI}$ which can be computed analytically with the halo model associated to a halo occupation distribution. We also note that the scale dependence of the bias models is significant because the coupling between small and large scale modes.

2.3 Halo model and halo occupation distribution

The halo model is an analytical framework that describes the clustering of dark matter on both linear and non-linear scales. It relies on the assumption that all dark matter lies in halos which are spherical virialized objects (Cooray and Sheth **2002**). It relies on two main ingredients: the halo mass function $\mathrm{d}n/\mathrm{d}M$ which is the number of halos with respect to the halo mass and the associated $n-$th order halo biases $b_n^h(M)$, Fig. 2.3. Both are measured in N-body simulations and we use the prescriptions of Sheth and Tormen **1999**. The comoving density of HI writes

$$\rho_{\rm HI} = \int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} M_{\rm HI}(M)\,. \tag{2.28}$$

The n-th order HI biases are

$$b_n^{\rm HI} = \frac{1}{\rho_{\rm HI}} \int \mathrm{d}M \frac{\mathrm{d}n}{\mathrm{d}M} b_n^{\rm h}(M)\, M_{\rm HI}(M)\,, \tag{2.29}$$

where $M_{\rm HI}(M)$ is the relation between the HI mass and the halo mass. Fig. 2.3 shows an example of a set of biases. Note that only the first order bias is always positive while the two others change sign. All of them increase for high halo masses. We will use the terms *linear* and *first order* bias interchangeably.

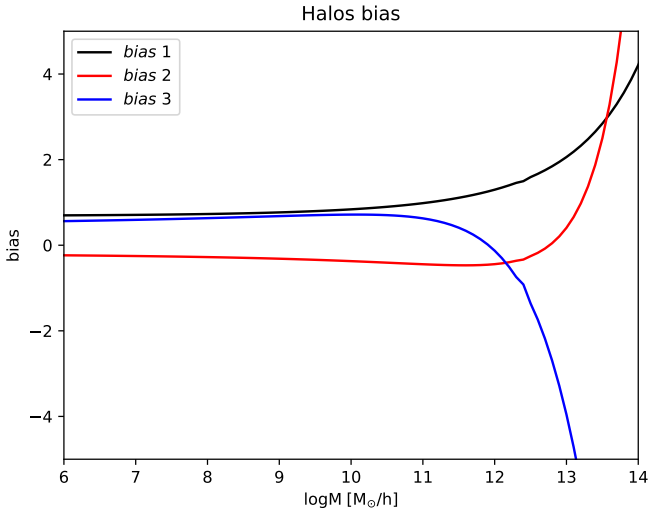


Figure 2.3 This figure shows a set of halo biases from the first order up to third order. We can see that only the first order bias is always positive while the two others change sign. The magnitude of the biases increase with the mass of the halo. There is also a small kink in each line, especially the blue line, at $\log \mathrm{M} \sim 12.5$ which can be considered as a cutoff. Beyond $\log \mathrm{M} \sim 12.5$ halo biases become higher and go to infinity.

Dark matter halos are filled with galaxies or HI through the halo occupation distribution (HOD) which describes the number of galaxies within a halo of a certain mass or, in our case, the relationship between the HI mass and the halo mass. The distribution of HI within the Large Scale Structure is rather unclear today. It is believed that in the post-reionization era most of HI lies within galaxies while only a negligible fraction is diffuse (Seehars et al. **2016**). It is often simply parametrised by relating the mass of HI to the mass of its host dark matter halo through a simple power law including, or not, a cut-off at small and high halo masses. We compile here several HI mass and halo mass relations that have been used or estimated using both hydrodynamical simulations and parametrised models fitted on data measurements. We also consider a DLA model.

1. **Bagla10**: One relation that has been widely used is that of Bagla et al. **2010**.

It has been inspired from quasar observations and assumes that there is no HI in high mass halos:

$$M_{\rm HI}(M) = \frac{f_3\, M}{1 + \frac{M}{M_{\rm max}}} \ , \tag{2.30}$$

where f_3 comes from the normalisation to $\Omega_{\rm HI}$. This prescription is commonly used for studies of 21 cm intensity mapping amongst others, Seehars et al., Sarkar et al., Villaescusa-Navarro et al. **2016**, **2016**, **2014**. $M_{\rm min}$ and $M_{\rm max}$ are the limits for a dark matter halo to host HI. They assume that only halos with 30 km/s $< v_{\rm circ} <$200 km/s host HI, which translates to lower and upper bounds, $M_{\rm min}$ and $M_{\rm max}$, through

$$v_{\rm circ} = 30\sqrt{1+z}\left(\frac{M}{10^{10} M_\odot}\right)^{1/3} \ {\rm km/s}\,. \tag{2.31}$$

2. **AGN**: Nevertheless, Villaescusa-Navarro et al. **2016** measured the $\rm M_{HI}M_h$ relation in hydrodynamical simulations including AGN feedback and show that there is HI in halos that have $v_{\rm circ}$ >200 km/s. They measured $M_{\rm HI}(M) = e^{\alpha}\, M^{\gamma}$ and fit α and γ up to redshift 2.

3. **DLA50**: A prescription adapted from DLA studies (Barnes and Haehnelt, Barnes and Haehnelt **2010**, **2014**) by Padmanabhan et al. **2016**

$$M_{\rm HI}(M) = \alpha\, f_{\rm H,c}\, M\, \exp\left[-\left(\frac{v_{c,0}}{v_c(M)}\right)^3\right] \exp\left[-\left(\frac{v_c(M)}{v_{c,1}}\right)^3\right], \tag{2.32}$$

where α is the ratio of HI within halos and cosmic HI, $f_{\rm H,c} = (1 - Y_p)\Omega_b/\Omega_m$ is the cosmic hydrogen fraction with Y_p the cosmological helium fraction by mass, and $v_c(M)$ is the virial velocity of a halo (Bullock et al. **2001**):

$$v_c(M) = 96.6\,{\rm km/s}\left(\frac{\Delta_v \Omega_m h^2}{24.4}\right)^{1/6}\left(\frac{1+z}{3.3}\right)^{1/2}\left(\frac{M}{10^{11} M_\odot}\right)^{1/3}, \tag{2.33}$$

where Δ_v, the mean overdensity of the halo, is taken to be 200. For DLAs, Padmanabhan et al. **2016** considered $v_{c,0} = 50$ km/s and an infinite $v_{c,1}$. They fitted α to measurements between redshift 0 and 4 (column density distributions, biases, $\Omega_{\rm HI}$ and the incidence rate).

4. **21cm**: Padmanabhan et al. **2016** adapted Eq. 2.32 to 21 cm IM observations using ad-hoc velocity cuts $v_{c,0} = 30$ km/s and $v_{c,1} = 200$ km/s. Similarly to the DLA50 model, Padmanabhan et al. **2016** fitted α on the same measurements. Note that in both latter cases the slope is fixed and equal to unity which is higher than what is measured in hydro-simulations.

5. **HOD A**: Padmanabhan and Kulkarni **2017** improved Eq. 2.32 by introducing a flexible slope, β, as well as the velocity cut-offs:

$$M_{\rm HI}(M) = \alpha\, f_{\rm H,c}\, M \left(\frac{M}{10^{11} h^{-1} M_\odot}\right)^{\beta} \exp\left[-\left(\frac{v_{c,0}}{v_c(M)}\right)^3\right] \times \exp\left[-\left(\frac{v_{c,1}}{v_c(M)}\right)^3\right], \tag{2.34}$$

where α, β, $v_{c,0}$, and $v_{c,1}$ are free parameters and fitted on data measurements.

6. **HOD B**: Lastly, Padmanabhan et al. **2017** fitted an updated version of Eq. 2.34,

$$M_{\rm HI}(M) = \alpha\, f_{\rm H,c}\, M \left(\frac{M}{10^{11} h^{-1} M_\odot}\right)^{\beta} \exp\left[-\left(\frac{v_{c,0}}{v_c(M)}\right)^3\right], \tag{2.35}$$

on all the available measurements including galaxy clustering. Their free parameters are β and α.

All these prescriptions are shown in Fig. 2.4 at $z = 1$. They vary in shape, amplitude, and slope. Clearly the DLA50 scheme favours high halo masses as compared to the

other models. We limit our analysis to $z = 1$, the values of the free parameters are given in Table 2.1.

Model	Parameters	b_1	b_2	b_3	$b_{\rm eff}$	$T_{\rm HI} \times 10^4$ K
Bagla10	None	0.93	-0.41	0.62	0.80	1.65
AGN	$\alpha = 0.73$, $\gamma = 2$	0.91	-0.27	0.41	0.82	12.14
21cm	$\alpha = 0.15$	0.96	-0.42	0.60	0.81	2.43
HOD A	$\log v_{c,0} = 1.58$, $\log v_{c,1} = 3.14$, $\alpha = 0.17$, $\beta = -0.5$	1.00	-0.35	0.38	0.82	4.38
HOD B	$\log_{10} v_{c,0} = 1.56$, $\alpha = 0.09$, $\beta = -0.58$	0.96	-0.37	0.49	0.85	2.55
DLA50	$\alpha = 0.13$	1.64	0.56	-1.27	1.74	49.94

Table 2.1: Free parameters of the $M_{HI}M_h$ prescriptions along with the associated HI biases and mean temperatures at $z = 1$. These models, except DLA50, fit to the GBT results (Masui et al. **2013**) which show that the linear bias of HI is expected to be ~ 0.65 to ~ 1 at $z \leq 1$.

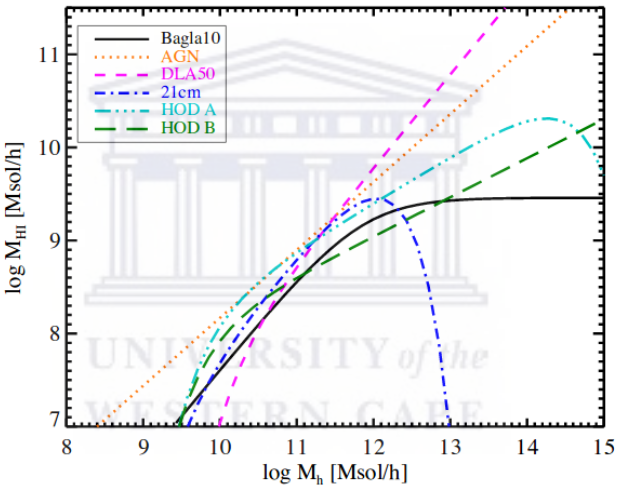


Figure 2.4 This figure shows the relations between the HI mass and the halo mass at $z = 1$. They vary in shape, amplitude, and slope. Clearly the DLA50 scheme favours high halo masses as compared to the other models. We limit our analysis to $z = 1$, the values of the free parameters are given in Table 2.1

Redshift dependencies of the HI biases of HOD B, $b1$ and $b2$, from equation 2.2 are presented in figure 2.5.

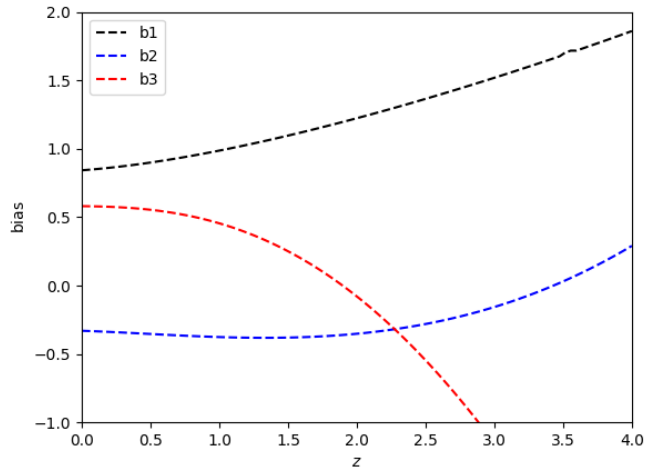


Figure 2.5 Evolution of the HI biases of HOD B with respect to the values of redshift which is given in equation 2.2. $n = 1, 2, 3$ respectively for first order b_1, second order b_2, third order b_3 HI biases. We can see that only the first order bias is always positive while the two others change sign.

Figure 2.6 shows the scale dependence of the HI bias.

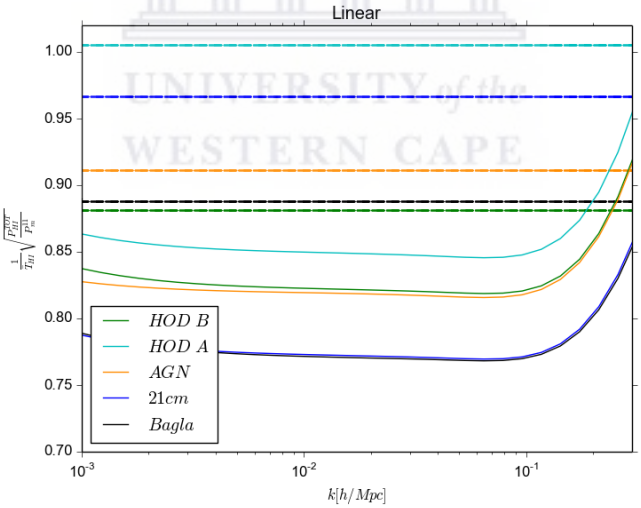


Figure 2.6 Solid line shows the scale dependence of the HI bias given by the ratio $\sqrt{P_{HI}^{TOT}/P_m^{11}}/\bar{T}_{HI}$. Horizontal dashed lines are the linear biases for each $M_{HI}M_h$ models computed with equation 2.7.

Chapter 3

Methodology

One of the main objectives of HI surveys and more widely, galaxy surveys, is to measure the matter power spectrum and in particular the BAO wiggles. Those measurements are carried out under the assumption that the power spectrum of the biased tracer is linear but we have seen in the previous chapter that non-linear contributions can affect linear scales. Therefore the estimations of the cosmological parameters might be biased. We will measure this bias for several cosmological parameters focusing on an intensity mapping survey which will be carried out by the planned HIRAX experiment, which is aimed at measuring the BAO around $z \sim 1$.
In this chapter, we will first explain the design of the HI survey as well as the measurements of the HI power spectrum using this survey. Lastly, we will explain the formalism of the Fisher matrix we use to compute the biases on the cosmological parameters, when assuming that non-linearities in the power spectrum are negligible. We will consider the following cosmological parameters: density of dark matter (Ω_m), baryon density (Ω_b), Hubble parameter (h), the spectral index (n_s), and the amplitude of mass fluctuations (σ_8).

3.1 HIRAX survey design

Number of dishes	1024
Dish diameter	6m
Survey area	10000 deg^2
Operating frequency	400MHz - 800MHz
redshift range	0.8 - 2.5
System temperature	50 K

Table 3.1: This table gives the HIRAX telescope characteristics. The observation time is 8760 hours which cover 10000 square degrees. These values serve as an input for our experimental design from which we obtain the characteritics of our survey given in table 3.2.

The establishment of the Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX) is proposed consisting of an array of 32×32 radio telescopes and requires an area of approximately 400 m × 400 m to install each 6 m diameter dishes with 3 m gaps in between dishes [1]. This makes HIRAX a compact array and therefore, to a good approximation, we can assume that the minimum baseline is close to the dish diameter. In this section we consider an HI intensity mapping survey for HIRAX. The specifications of HIRAX are listed in table 3.1. Since HIRAX is an interferometer, the largest scale it can probe is set by its minimum baseline and its instantaneous field of view ($FoV = \Delta\theta^2$) is set by its dish size, e.g. $\Delta\theta \sim (\lambda/D)$, where λ is the wavelength of the observation ($0.21(1+z)$ m for 21 cm observations). For HIRAX we have D=6m, which gives $\Delta\theta = 4.5°$ at $z = 1$ (or a FoV of 20.25deg^2). Although the HIRAX survey is planned to observe a large fraction of the sky (25% of the sky that HIRAX can cover, which is 10000 deg^2), we can divide that area into patches of the size of the FoV since HIRAX will not probe scales larger than that. It is then safe to use the flat sky approximation for those patches.

[1] https://www.sarao.ac.za/wp-content/uploads/2019/03/SKA-IEMP-Chapter-2.pdf

The angular diameter distance D_A at the same redshift ($z = 1$) are used to compute the transverse comoving size $L_\perp$,

$$L_\perp = (1 + z)D_A(z)\Delta\theta = D_c(z)\Delta\theta\,. \tag{3.1}$$

The comoving distance $D_c\,(z) = (1 + z)\,D_A\,(z)$ can be written as

$$D_c(z) = \chi(z) = c\int_0^z \frac{dz'}{H(z')}\,. \tag{3.2}$$

Along the line of sight, we have the comoving size $L_{||}$ which can be calculated within a redshift bin Δz centred at $z = 1$,

$$L_{||} = \chi(z + \Delta z/2) - \chi(z - \Delta z/2) = c\int_{z-\Delta z/2}^{z+\Delta z/2} \frac{dz'}{H(z')}\,. \tag{3.3}$$

For this survey, we will consider a redshift interval (frequency range) so that we have a square box in comoving coordinates, e.g. $L_\perp = L_{||}$. For small redshift intervals, we can assume H(z) to be constant, therefore:

$$L_{||} = \Delta\chi \approx c\frac{\Delta z}{H(z)}. \tag{3.4}$$

The bin size of the mode Δk is set by the length of the box:

$$\Delta k = \frac{2\pi}{L_{||}}\,. \tag{3.5}$$

Following Grasshorn Gebhardt et al. **2019**, the number of Fourier modes is given by N_k that is written as follow,

$$N_k = \frac{V_{\rm box}}{(2\pi)^3}\Delta V_k\,. \tag{3.6}$$

With $\Delta V_k\,(k, \Delta k)$ being the volume in Fourier space contributing to the estimation of the power spectrum and the Fourier bin can be also written as $\Delta k = 2\pi/(V_{\rm box})^{\frac{1}{3}}$. In this study, we divide our survey into patches of $N_{\rm boxes}$ given by the following equation,

$$N_{\rm boxes} = \frac{\text{Survey area}}{(\Delta\theta)^2}. \tag{3.7}$$

We add instrumental noise, in section 3.2, from the HIRAX telescope with the characteristics given in table 3.1 . Table 3.2 sums up the characteristics of the survey calculated from these equations at $z = 1$,

H(z)	123.248 Km s^{-1} Mpc^{-1}
Comoving distance $D_c(z)$	3303.83 Mpc
$\Delta\theta$	0.0785 rad= 4.5° (small angle approximation)
Comoving size $L_{\|\|} = L_\perp$	259 Mpc = 174 Mpc/h
bin size $\Delta k \approx k_{\rm min}$	0.02425 Mpc^{-1} = 0.036132h Mpc^{-1}
redshift bin Δz	0.1064
$N_{\rm boxes}$	494

Table 3.2: Assumed survey characteristics calculated from the values in table 3.1 and several equations explained above. The dish size sets Δk while the minimum baseline sets $k_{\rm min}$. Usually $k_{\rm min}$ is larger than Δk. Saying they are equal is an approximation only valid for compact arrays which is the case of HIRAX as explained in section 3.1.

3.2 Simulating HI power spectrum for the HIRAX instrument

To compute the HI power spectrum we use the models explicitly stated in chapter 2 and we compute the error bars for a measurement with the HIRAX survey. We start by deriving the power spectrum variance. Besides a normalization due to the

window function, an estimator for the power spectrum may be written as [57],

$$P_{\rm HI}^{\rm obs} = \frac{1}{N_k}\sum_{i=1}^{N_k}|\delta(k_i)|^2\Big|_{|k_i-k|\leq\Delta k}, \tag{3.8}$$

where $\delta(ki)$ is a Fourier transform of the density field in position space, Δk is the k bin size and N_k is the number of independent k-modes available per bin. This estimator is unbiased because

$$\left\langle P_{\rm HI}^{\rm obs}(k)\right\rangle = \frac{1}{N_k}\sum_{i=1}^{N_k}\left\langle|\delta(k_i)|^2\right\rangle = \left\langle|\delta(k_i)|^2\right\rangle = P_{\rm HI}(k), \tag{3.9}$$

where $P_{\rm HI}(k)$ is the underlying HI power spectrum. The variance of this estimator is given by:

$$\left\langle\left(\frac{P_{\rm HI}^{\rm obs}(k)-P_{\rm HI}(k)}{P_{\rm HI}(k)}\right)^2\right\rangle = 1-2\frac{\left\langle P_{\rm HI}^{\rm obs}(k)\right\rangle}{P_{\rm HI}(k)}+\frac{1}{N_k^2P_{\rm HI}(k)^2}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\left\langle\delta(k_i)^*\delta(k_i)\delta(k_j)^*\delta(k_j)\right\rangle . \tag{3.10}$$

Assuming that the density field is a Gaussian random variable with its variance given by:

$$\left\langle\delta_i^*\delta_j\right\rangle = P_{\rm HI}(k)\delta_{ij} . \tag{3.11}$$

We use Wick's theorem for evaluating the last double summation over i an j in equation 3.10. We got,

$$\begin{aligned}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\left\langle\delta_i^*\delta_i\delta_j^*\delta_j\right\rangle &= \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\left[\left\langle\delta_i^*\delta_i\right\rangle\left\langle\delta_j^*\delta_j\right\rangle+\left\langle\delta_i^*\delta_j\right\rangle\left\langle\delta_j^*\delta_i\right\rangle+\left\langle\delta_i^*\delta_j^*\right\rangle\left\langle\delta_i\delta_j\right\rangle\right]\\ &= N_k^2\left|P_{\rm HI}(k)\right|^2+N_k\left|P_{\rm HI}(k)\right|^2 .\end{aligned} \tag{3.12}$$

Recall equation 3.12 and substitute the value of this summation to equation 3.10 to get the expression of the variance in equation 3.13,

$$\left\langle \left(P_{\rm HI}^{\rm obs}(k) - P_{\rm HI}(k) \right)^2 \right\rangle = \frac{P_{\rm HI}(k)^2}{N_k} = \sigma(k)^2, \tag{3.13}$$

which yields the formula to compute the error bar of the HI power spectrum in equation 3.14,

$$\sigma\,(k) = \frac{P_{\rm HI}(k)}{\sqrt{N_k}}\,. \tag{3.14}$$

Actually the formula being calculated in equation 3.14 is only for one box (one patch) of the sky. So the variance for the whole survey (see 3.2) of $N_{\rm boxes}$ is as follows:

$$\left\langle \left(P_{\rm HI}^{\rm obs}(k) - P_{\rm HI}(k) \right)^2 \right\rangle = \frac{P_{\rm HI}(k)^2}{N_k N_{\rm boxes}} = \sigma\,(k)^2, \tag{3.15}$$

which gives the actual error for the noise free power spectrum,

$$\sigma\,(k) = \frac{P_{\rm HI}(k)}{\sqrt{N_k}\sqrt{N_{\rm boxes}}}\,. \tag{3.16}$$

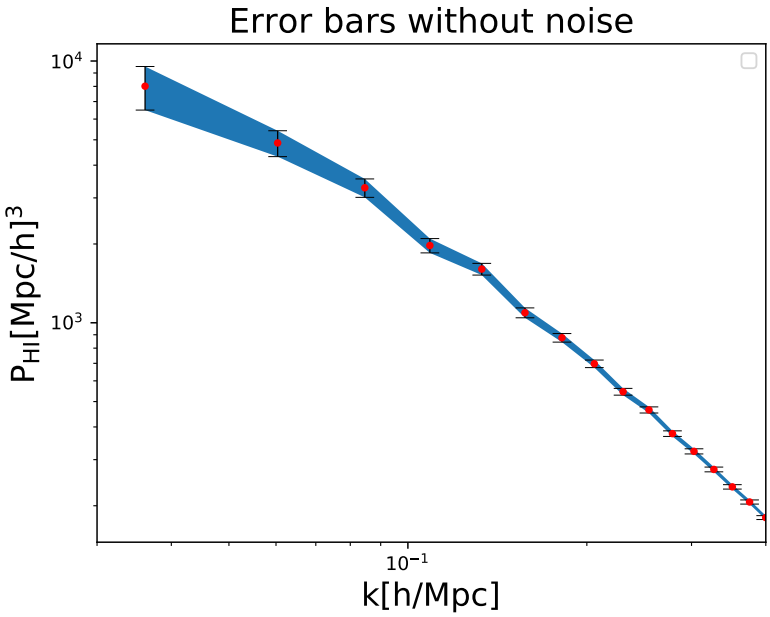


Figure 3.1 HI power spectrum with error bars and no instrumental noise. For a HIRAX type survey at $z \sim 1$.

Including the noise from the instrument the formula in equation 3.16 becomes as follows:

$$\sigma(k) = \frac{(P_{\mathrm{HI}} + P_{\mathrm{Noise}})}{\sqrt{N_k}\sqrt{N_{\mathrm{boxes}}}}. \tag{3.17}$$

In order to calculate the noise power spectrum for HIRAX, we will need to know the uv distribution of the interferometer. This is a function of the dish distribution and the amount of time spent observing a certain patch of the sky. In order to proceed, we used the results from the code developed in Witzemann et al. **2018**. From it, we obtain fractional errors, $\frac{\sigma_p}{P}$, on the HI power spectrum for HIRAX, where P is the HI power spectrum at a given z and σ_p is the error on the power spectrum including noise and cosmic variance (and already factoring in the field of view and survey area, e.g. using equation 3.16). As shown in figure 3.2, $\frac{\sigma_p}{P}$ is already given at different redshifts. We need to interpolate these values to our target redshift $z = 1$. Moreover, the error here was calculated assuming a different redshift and k bin, which changes

the cosmic variance. So we need to rescale as follows,

$$\sigma_P = \frac{\frac{\sigma_P}{P} P_{\rm HI}}{\sqrt{\frac{dz^{'}}{dz}}\sqrt{\frac{dk^{'}}{dk}}}, \tag{3.18}$$

such that $dz = 0.1064$ (see table 3.2) and $dz^{'} = 0.2$.

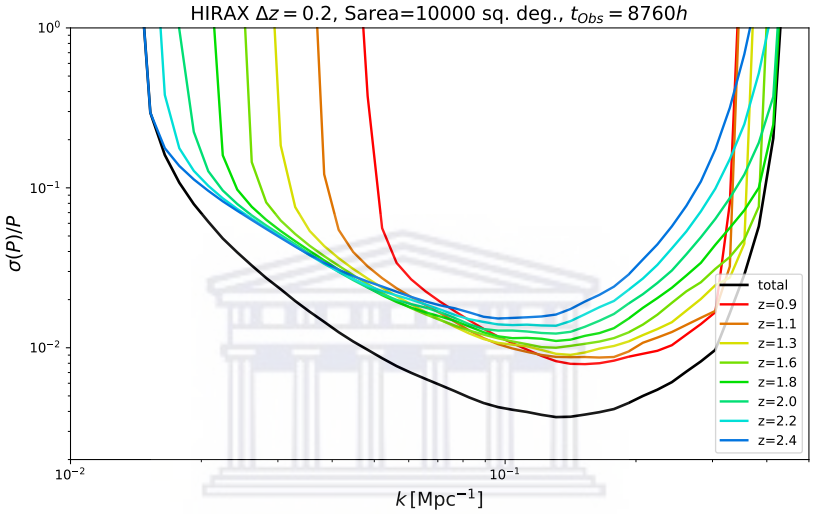


Figure 3.2 Error to power spectrum ratio computed from HIRAX at various redshifts. We interpolate $z = 0.9$ and $z = 1.1$ to get our $\frac{\sigma_p}{P}$ at the desired redshift $z = 1$.

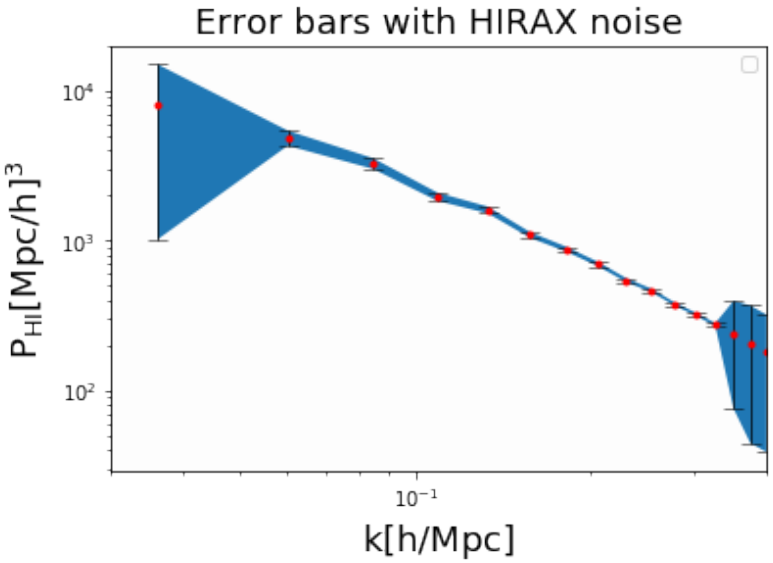


Figure 3.3 HI power spectrum with error bars, including instrumental noise from HIRAX. The error is high at large scales (small k) due to cosmic variance and large at small scales (large k) due to the lack of long baselines in the interferometer.

The resulting HI power spectrum with error bars is shown in figure 3.1 and figure 3.3 without and with instrumental noise, respectively. Without instrumental noise, we recover the increase of error bars on large scales due to cosmic variance. Instrumental noise can increase the error bars on both large and small scales due to the baseline distribution, although for compact arrays like HIRAX we except such noise power spectrum to be more or less constant in k with a sharp cutoff both on large and small k. Moreover the ratio of signal to noise with decrease for large k since the HI signal goes down as the noise stays constant or increases. This can be clearly seen in figure 3.3. More importantly, the error is quite small over the BAO region.

3.3 Fisher matrix formalism and biases evaluation

Fisher matrix measures the information provided by an experimental set-up for estimation of the parameters (Zimmer 2016). It can be used to calculate the confidence

intervals for each parameter or a multidimensional confidence area for all parameters. So we can say that it is used to predict the error in parameters for a given model, and without any data required, under the assumption that the likelihood is a multivariate Gaussian (Fisher **1935**). In fact, the Fisher matrix formalism translates errors on observed quantities measured directly in the survey into constraints on parameters of interest in the underlying model. To put it more directly, it is the elegant way of doing propagation of errors in the case of multiple, correlated, measurements and many parameters (Tegmark et al. **1997**). We write the Fisher matrix F_{ij} as follows,

$$F_{ij} = \sum_k \frac{1}{\sigma(k)^2} \frac{\partial X(k)}{\partial \theta_i} \frac{\partial X(k)}{\partial \theta_j} \tag{3.19}$$

where $\sigma(k)$ is the error on the observable X (e.g. power spectrum measurement) and θ_i are parameters (e.g. cosmological parameters). To get the equation 3.19, let θ be a vector of parameters and θ^* be the fiducial model. We can expand the likelihood about θ^*:

$$\ln\mathcal{L}(\theta^* + \delta\theta) = \ln\mathcal{L}(\theta^*) + \sum_{\mathrm{i}} \left.\frac{\partial \ln\mathcal{L}(\theta)}{\partial \theta_{\mathrm{i}}}\right|_{\theta=\theta^*} \delta\theta_{\mathrm{i}} + \frac{1}{2} \sum_{\mathrm{i,j}} \left.\frac{\partial^2 \ln\mathcal{L}(\theta)}{\partial \theta_{\mathrm{i}} \partial \theta_{\mathrm{j}}}\right|_{\theta=\theta^*} \delta\theta_{\mathrm{i}} \delta\theta_{\mathrm{j}} + \ldots \tag{3.20}$$

The first term is a constant which depends only on the fiducial model. Since, after many data realisations, we expect the fiducial model to be the point of maximum likelihood, then by definition the second term will vanish. The third term is the curvature matrix or Hessian of the likelihood and defines the Fisher matrix(Fisher **1935**):

$$F_{ij} = \left\langle -\frac{\partial^2 \ln\mathcal{L}(\theta)}{\partial \theta_i \partial \theta_j} \right\rangle. \tag{3.21}$$

The angle brackets represent the expectation value. To write the Fisher matrix in terms of the theoretical predictions for the observables, X, first note that the

likelihood is given by:

$$\mathcal{L} = \frac{1}{\sqrt{(2\pi)^N |C|}} \exp\left(-\frac{1}{2}\Delta^{\mathrm{T}} \mathrm{C}^{-1} \Delta\right). \tag{3.22}$$

Where $\Delta = X - d$ (the difference between the theoretical prediction and the measured quantity), C is the data covariance matrix and N is the number of data points. Substituting this into equation 3.21, we find:

$$F_{ij} = \frac{\partial X^T}{\partial \theta_i} C^{-1} \frac{\partial X}{\partial \theta_j} + \frac{1}{2}\left(C^{-1}\frac{\partial C}{\partial \theta_i} C^{-1} \frac{\partial C}{\partial \theta_j}\right). \tag{3.23}$$

If the data covariance matrix, which also includes effects like cosmic variance, is independent of the parameters, as is often the case (Bassett et al. **2011**), and if the data are uncorrelated, the Fisher matrix becomes:

$$F_{ij} = \frac{\partial X^T}{\partial \theta_i} C^{-1} \frac{\partial X}{\partial \theta_j} = \sum_n \frac{1}{\sigma_n^2} \frac{\partial X_n}{\partial \theta_i} \frac{\partial X_n}{\partial \theta_j}. \tag{3.24}$$

Where n is the index over the data. In our case, the observable X is the power spectrum of neutral hydrogen P_{HI}. The inverse of the Fisher matrix estimates the parameter covariance matrix. In the case for an unbiased estimator (i.e. the expected value of θ corresponds to θ^*) and where one does not marginalise over any parameters, then the expected error satisfies the Cramer-Rao bound: $\Delta\theta_i \geq 1/\sqrt{F_{ii}}$.

In the more realistic case which includes marginalisation, this inequality becomes: $\Delta\theta_i \geq \sqrt{F_{ii}^{-1}}$. It can be shown that the error obtained when marginalising is always greater than or equal to the error obtained without marginalisation. It should also be noted that in the case where the likelihood is exactly Gaussian, the Cramer-Rao bound becomes an equality.

The Fisher information matrices can be calculated for different experimental set-ups allowing the selection of the most informative design. This procedure is

called experimental design. The goal is to obtain a parameter estimate that is as precise as possible, which means that its variance is as small as possible.

Throughout this thesis, we will be concentrating on the full non-linear power spectrum. However, the question arises of what will happen to our measurements of the linear power spectrum instead is assumed in parameter fitting. In the era of high precision cosmology, this could have a high impact. The Fisher information matrix approach also allows to calculate how the best fit parameters will be biased by using the wrong fitting model.

Following Shimon et al. **2013** and Linder **2006**, the bias $\delta\theta_i$ on each cosmological parameter θ_i depends on the inverse of the Fisher matrix. Parameter biases $\delta\theta_i$ induced from offsets ΔO_k in the observable quantities are calculated using the Fisher formalism, where maximizing the likelihood leads to (to linear order):

$$\delta\theta = A\Delta O = (U^T C^{-1} U)^{-1} U^T C^{-1} \Delta O\,, \tag{3.25}$$

where O is the vector of expected observations, C is the covariance matrix of observational errors, and $U = \dfrac{\partial O}{\partial \theta}$. Put more simply, when the covariance matrix is diagonal,

$$\delta\theta_i = (F_{ij}^{-1}) \sum_k \frac{\partial O_k}{\partial \theta_j} \frac{1}{\sigma_k^2} \Delta O_k, \tag{3.26}$$

and O_k is the kth observable (e.g. power spectrum), ΔO_k is the observational quantity offset, and F is the Fisher matrix from all observables. In our case and to linear order, we write equation 3.26 as follows:

$$\delta\theta_i = \sum_{j,k} F_{ij}^{-1} \frac{\Delta P_{\rm HI}(k)}{\sigma(k)^2} \frac{\partial P_{\rm HI}^{LIN}(k)}{\partial \theta_j}\,, \tag{3.27}$$

where

$$\Delta P_{\rm HI} = P_{\rm HI}^{\rm NL} - P_{\rm HI}^{\rm LIN}\,, \tag{3.28}$$

such that $P_{\rm HI}^{\rm NL}$ is the total power spectrum that includes non-linear terms and $P_{\rm HI}^{\rm LIN}$ is the linear power spectrum. Even if $\Delta P_{\rm HI}$ is small it can still result in a much more significant bias in parameter inference if the sensitivity of $P_{\rm HI}$ to small variations in the cosmological model is large.

Chapter 4

Results and discussion

We are now going to show the results based on the theory of the HI power spectrum developed in chapter 2 and the methodology of error bars and biases forecasts from chapter 3. At the end, we will show the effects on the BAO and in particular discuss the changes in the peak location. Our forecasts in this chapter will be based on the HIRAX experiment.

4.1 Biases and errors in cosmological parameters

With the results shown in table 4.1 and table 4.2, we bring answers to the following questions: how well can we measure the cosmological parameters when considering the non-linear model? How much is the measurement of cosmological parameters biased when using the linear model? We follow the Fisher matrix formalism developed in the previous chapter. Note that we do not include priors from other experiments such as Planck.

In table 4.1, we show the results of errors and biases in the cosmological parameters without including the instrumental noise from HIRAX in the estimation. In the last column, we have the ratio of the bias to the error that allows us to know whether

biases are significant compared to the errors. We found that values of the ratio are at the same level as the errors and even higher (from 0.348 to 7.351). Ratios of 1 and above signal that the bias is relevant and may affect our conclusions on what is the best cosmological model given the data. Most parameters suffer from a strong bias.

Forecast without instrumental noise

	Cosmo parameters	Error (Non-linear case)	Bias	$\text{Bias/Error}_{\text{NL}}$
Ω_m	0.2670	0.0071 (~2.67%)	0.0024 (~0.93%)	0.348
h	0.67115	0.0089 (~1.34%)	0.0078 (~1.17%)	0.873
Ω_b	0.0489	0.0018 (~3.80%)	0.0026 (~5.33%)	1.402
n_s	0.9608	0.0028 (~0.30%)	0.0063 (~0.66%)	2.200
σ_8	0.826	0.0030 (~0.37%)	0.0224 (~2.72%)	7.351
ω_0	-1.0	0.0064 (~0.64%)	0.0126 (~1.26%)	1.968

Table 4.1: This table gives the forecast results of bias and error in cosmological parameters without instrumental noise. In the first column, we have the cosmological parameters whose errors and biases are respectively computed in the second column and the third column. In the last column, we evaluated the bias to error ratio in order to see how significant is the bias compared to the error.

Forecast Results with HIRAX instrumental noise

	Cosmo parameters	Error (Non-linear case)	Bias	$\text{Bias/Error}_{\text{NL}}$
Ω_m	0.2670	0.0081 (~3.07%)	0.0017 (~0.67%)	0.218
h	0.67115	0.0132 (~1.98%)	0.0053 (~0.79%)	0.398
Ω_b	0.0489	0.0020 (~4.24%)	0.0015 (~3.21%)	0.757
n_s	0.9608	0.0054 (~0.56%)	0.0038 (~0.40%)	0.714
σ_8	0.826	0.0133 (~1.62%)	0.0151 (~1.83%)	1.129
ω_0	-1.0	0.0096 (~0.96%)	0.0102 (~1.02%)	1.062

Table 4.2: This table gives the forecast results of bias and error in cosmological parameters with HIRAX instrumental noise from HIRAX telescope at $z = 1$ using equation 3.26. In the first column we have the cosmological parameters whose errors and biases are respectively computed in the second column and the third column. In the last column we evaluated the bias to error ration in order to see how significant is the bias compared to the error.

In table 4.2, we include the noise from HIRAX when forecasting biases and errors.

The Biases decrease as expected from the equation 3.27 because the error, $\sigma(k)$, increases. For an experiment with errors like HIRAX, it seems that the bias imposed on the cosmological parameters due to ignoring the non-linear effects is not negligible. The situation will get worse if for instance we include further priors. We also show the fractional variation of the bias to error ratio, which we dubbed "p", in table 4.3.

Fractional variation of the bias to error ratio p

	$\frac{\Delta p}{p} = \frac{\lvert p_1 - p_0 \rvert}{p_0}$	$\frac{\Delta p}{p}$ %
Ω_m	0.373	37
h	0.524	52
Ω_b	0.460	46
n_s	0.675	67
σ_8	0.846	84
ω_0	0.460	46

Table 4.3: Variation of bias to error ratio. p_1, given in the third column of 4.2, is the ratio of the bias to the error of the forecast with noise from HIRAX. p_0, given in the third column of 4.1, is the ratio of the bias to error of the forecast noise free. Each value in the third column tells us by how many percent the bias to error ratio changes once noise is added.

4.2 Baryonic acoustic oscillations peak

We now turn our attention to the Baryon Acoustic Oscillations. Several experiments are planned to measure this effect with great accuracy in order to probe dark energy. This includes the HIRAX telescope which is the focus of this thesis. One of the great advantages of the BAO is that we can get constrains on the angular distance and Hubble parameter just by using the position of the peaks and troughs. This is very robust to uncertainties in calibration thus allowing for a much more clean measurement of dark energy. It is therefore crucial that we test the effect of neglecting

the non-linearities we have been considering on the position of the BAO wiggles. We use the method given in Bull et al. **2015** to extract the BAO signals. The idea is to split the power spectrum into the smooth part and oscillatory part. We construct a preliminary oscillatory function by dividing the sampled $P(k)$ (matter or HI) by the splined function (not its logarithm), then fit another cubic spline to the result and find the zeros of its second derivative with respect to k. These are the points at which the first derivatives of the oscillatory function are maximal/minimal, and in some sense define "mid-points" of the function and its overall trend. We construct a cubic spline through these too, and then divide the preliminary oscillatory function by it to detrend. This leaves $f_{\rm bao}(k)$ as the final result. We start by showing the expected error on the BAO wiggles in figure 4.1. The high sensitivity of HIRAX shows that any deviations on the peak positions can bias the results.

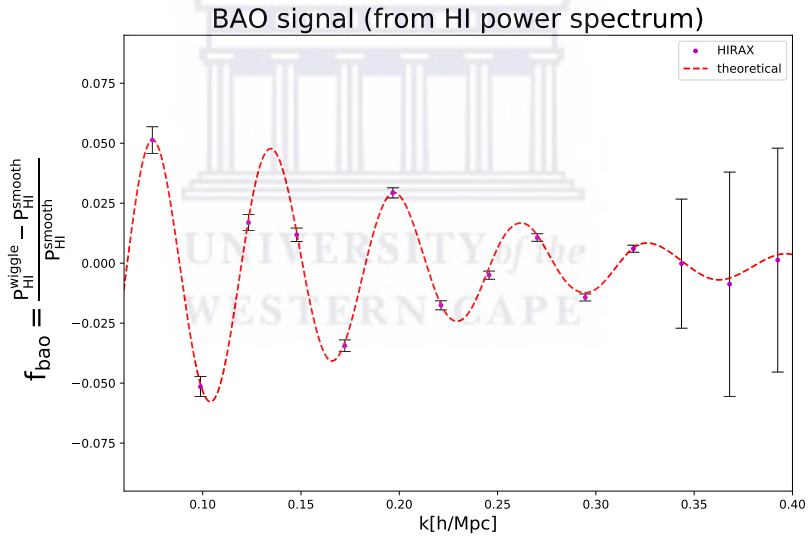


Figure 4.1 Baryonic acoustic oscillations feature with error bars from HIRAX at $z = 1$. This figure shows that HIRAX can detect well the BAO for the values of the mode k that range from $\approx$ 0.050 [h/Mpc] to 0.325 [h/Mpc], beyond that the errors bars are high. The dashed red line is the theoretical fit of the BAO features which shows a good overlap with the data point with error bars noise included from HIRAX instrument (which include binning).

In figure 4.2 we compare the BAO from the linear and non-linear matter power spectrum using CAMB. Non-linearity solely affects the amplitude of the signal. On the other hand, in figure 4.3 we show the BAO signal from the HI power spectrum from both the linear and non-linear models described previously. We don't expect differences between the linear HI BAO and dark matter BAO. For the non-linear HI, we still see a reduced amplitude of the acoustic oscillations as before. However, we also notice a phase shifting. We plot the evolution of that shift in figure 4.4. We see that we have a good overlap of peaks in the first peaks and then the separation grows at small scales.

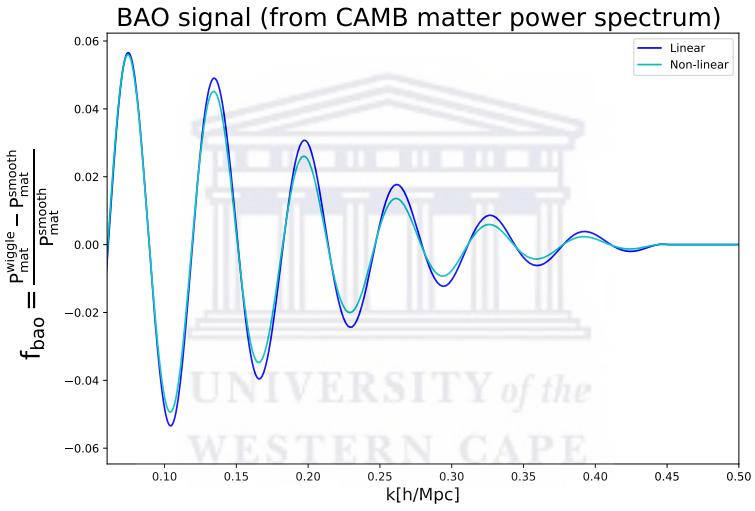


Figure 4.2 Baryonic acoustic oscillations features, $f_{\rm bao} = \frac{P_{\rm mat}^{\rm wiggle} - P_{\rm mat}^{\rm smooth}}{P_{\rm mat}^{\rm smooth}}$, from matter power spectrum from CAMB. $P_{\rm mat}^{\rm smooth}$ is the smooth component of $P_{\rm mat}^{\rm wiggle}$. There is a reduction of the amplitude but peaks remain at the same location.

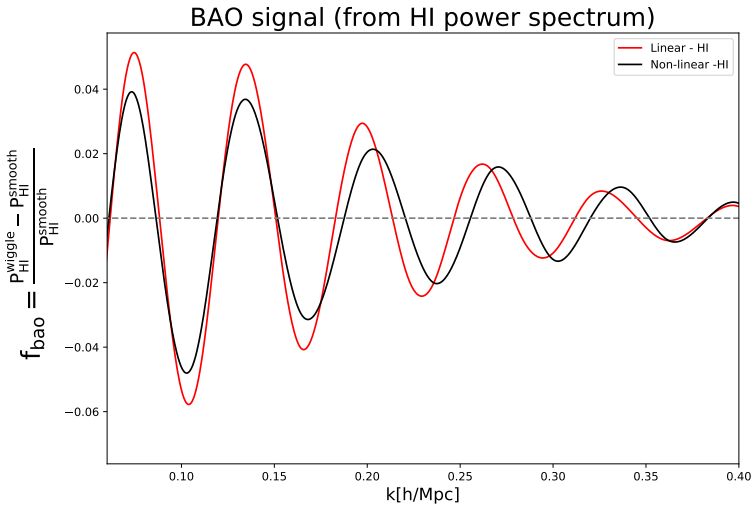


Figure 4.3 This plot shows the baryonic acoustic oscillations features, $f_{\rm bao} = \frac{P_{\rm HI}^{\rm wiggle} - P_{\rm HI}^{\rm smooth}}{P_{\rm HI}^{\rm smooth}}$, from HI power spectrum. $P_{HI}^{\rm smooth}$ is the smooth component of $P_{\rm HI}^{\rm wiggle}$. There is a reduction of the amplitude and change in peak location especially at small scales compared to Fig 4.2

At large scales we can see, in figure 4.3, that the amplitudes are changing but peaks (the first and the second peaks) remains at the same location. We expected that as $P_{13}^{\rm HI}$ is a dominant negative component at large scales while $P_{22}^{\rm HI}$ is a not significant at these scales. However, at small scales, peaks of the BAO signal are shifted although the amplitude almost remain unchanged. These change in peak location is due to the contribution of $P_{13}^{\rm HI}$ on the wiggles. The third order component $P_{13}^{\rm HI}$, which also carries wiggles, tends to balance with the second order component, $P_{22}^{\rm HI}$, as we shown in figure 2.2. From figure 4.5 we can see that the amount of the amplitude suppression at small-scales tend to be smaller than at large scales. The shift of the peaks starts at $k = 0.187$ h/Mpc.

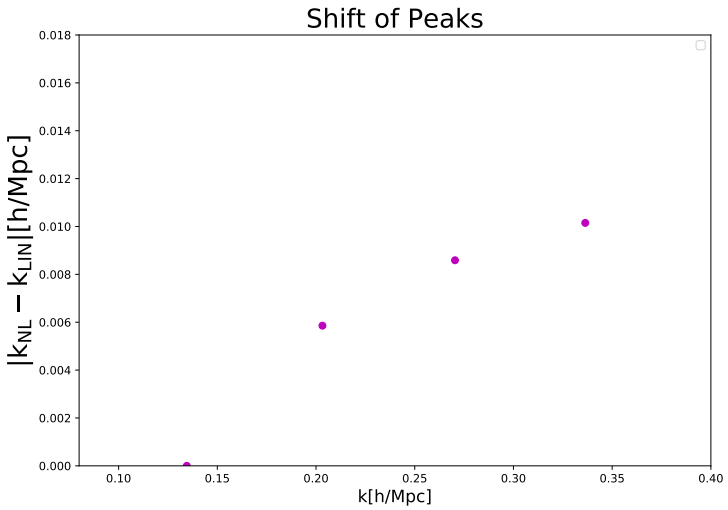


Figure 4.4 This plot show the shift of peaks in magnitude given by the difference of location of peaks of the BAO features from the linear model and the non-linear model. The BAO signals are highly shifted at small scales.

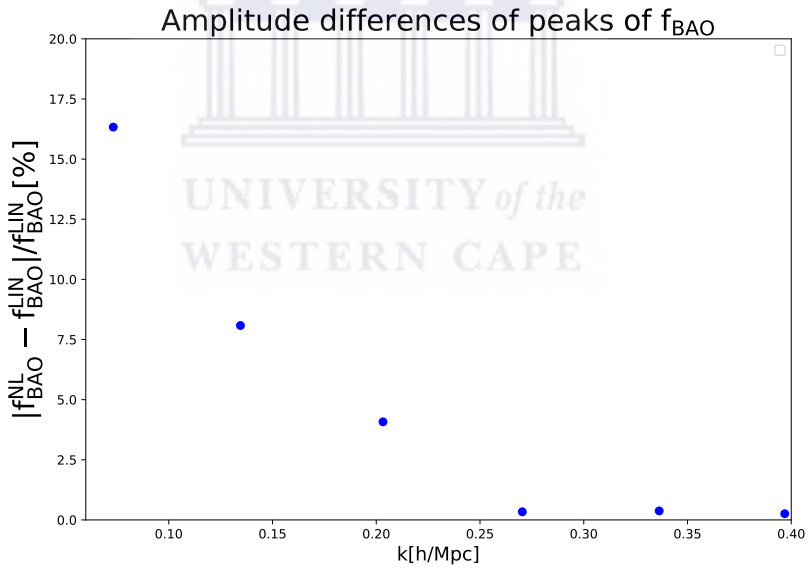


Figure 4.5 This shows reduction in amplitude of BAO features of the non-linear model compared to the linear model. We can see the variation in percentage given by the ratio $\left|f^{NL}_{\text{bao}} - f^{LIN}_{\text{bao}}\right| / f^{LIN}_{\text{bao}}$. The BAO signals get damped at large scales.

The value in percentage of the reduction of the amplitude of the BAO features, in

Figure 4.5 range from 16.33% for the first peak to 0.33% from those at small scales. For the change of the location of the shift, the value of the shift is from 0 h/Mpc to 0.01 h/Mpc which is a significant value at small scales.

In summary, the first peak and secondary peaks change in amplitude. Further secondary peaks change in location. For this reason, we get biased in constraining and forecasting cosmological parameters such as the angular diameter distance and Hubble, if we use the linear model.

Chapter 5

Conclusions

In this thesis, I have tested the non-linear model of the power spectrum of the neutral hydrogen and benchmark it against the linear model. The non-linear HI power spectrum was calculated with a full one loop correction. The code was implemented in Python and produces the HI non-linear power spectrum using as input the CAMB linear matter power spectrum and the different HI bias parameters from the Halo model.

We aimed to forecast the biases in cosmological parameters and to study the BAO. The errors are forecasted using the Fisher matrix formalism. We used the biases formula for a given model according to Shimon et al. **2013** and Linder **2006** which also depends on the Fisher information matrix F_{ij}. To implement this study, we designed a survey of 10,000 squared degrees with Hydrogen Intensity and Real-time eXperiment (HIRAX). It is a 400-800MHz radio interferometer which will map most of the southern sky over the course of four years (Nishimichi et al. **2007**). HIRAX will observe unresolved sources of neutral hydrogen via their redshifted $21-$ cm emission line. Indeed, it is convenient for our study at $z = 1$ as its redshifts range from 0.8 to 2.5.

We find that the biases are mostly at the same level as the errors, where the ratio of the bias to error ranges from 0.218 to 1.129 (going up to a factor of 7 without

noise). That means that if we do not consider the non-linear components of the HI power spectrum in our forecast we can get biased in our constraints for cosmological parameters, although the effect is small given the noise levels. Using the true model (non-linear model), the accuracy in measurement of cosmological parameters, with HIRAX, is 0.67% for Ω_m, 0.79% for h, 3.21% for Ω_b, 0.40% for n_s, 1.83% for σ_8, 1.02% for ω_0.

In order to consider in more detail the effect on parameters that rely on baryon acoustic oscillations, we also looked at the extraction of the BAO features f_{bao}, which is the oscillatory part of P_{HI}^{wiggle}, from the non-linear HI power spectrum and compared it to the f_{bao} signal from the HI linear power spectrum in order to investigate the discrepancy. Here, the P_{HI}^{smooth} is obtained following the method in Bull et al. **2015** (see chap 4). Then we choose two reference values of k that bound the region in which the oscillations are significant in total HI power spectrum and find the points at which the first derivatives of the oscillatory function are maximal or minimal. Then we construct a cubic spline and detrend it to get f_{bao} in Fig 4.3.

Our results about the BAO features show that there is a change in the location of the peaks (see Fig 4.4). The amount of the shift has a maximum value of 10^{-2} h/Mpc, although, the first peak is not shifted at all. We also have noticed a change in the amplitude of f_{bao} (see Fig 4.5. The reduction of the amplitude is high at the first and second peaks which goes up to 16.33% and becomes insignificant at small scales. Briefly, we can affirm that the non-linearities have a significant effect both at large and small scales. We then should adopt the non-linear model to gain more precision and accuracy.

Lastly, as a prospect of future work, I believe it is possible to carry out this same work on the redshift space distortions. In fact, redshift-space distortions has an effect on the shift of the BAO characteristic scales in a weakly nonlinear regime

using one-loop correction from perturbation theory (Nishimichi et al. **2007**). Further studies can also be performed at different redshifts and more parameters can also be considered such as ω_0 and ω_a. It will be important to confirm that the non-linear model of the power spectrum is crucial to anyone who works on constraining dark energy.

References

Book Sources

(14) Dodelson, S., *Modern cosmology*; Academic Press: San Diego, CA, 2003.

(22) Bradt, H., *Astrophysics Processes: The Physics of Astronomical Phenomena*; Cambridge University Press: 2008.

(39) Peebles, P. J. E., *The large-scale structure of the universe*; Princeton University Press: 1980.

Other Sources

(1) Ellis, G. F. R. *Quarterly Journal of the Royal Astronomical Society* **1975**, *16*, 245–264.

(2) Er, X.; Li, G.; Mao, S.; Cao, L. *Monthly Notices of the Royal Astronomical Society* **2013**, *430*, 1423–1432.

(3) Carlstrom, J. E.; Holder, G. P.; Reese, E. D. *Annual Review of Astronomy and Astrophysics* **2002**, *40*, 643–680.

(4) Nock, K.; Percival, W. J.; Ross, A. J. *Monthly Notices of the Royal Astronomical Society* **2010**, *407*, 520–532.

(5) Tegmark, M. *Physical Review Letters* **1997**, *79*, 3806–3809.

(6) Schwarz, D. J. In *Fundamental Interactions: A Memorial Volume for Wolfgang Kummer. Edited by GRUMILLER DANIEL ET AL. Published by World Scientific Publishing Co. Pte. Ltd., 2010. ISBN #9789814277839, pp. 267-276*, Grumiller, D., et al., Eds.; World Scientific Publishing Co: 2010, pp 267–276.

(7) Park, C.-G.; Hyun, H.; Noh, H.; Hwang, J.-c. *Mon. Not. Roy. Astron. Soc.* **2017**, *469*, 1924–1931.

(8) Scrimgeour, M. I. et al. *Monthly Notices of the Royal Astronomical Society* **2012**, *425*, 116–134.

(9) Das, A.; Banerjee, A.; Chakraborty, S.; Pan, S. *Pramana* **2018**, *90* 19, 19.

(10) Planck Collaboration et al. *arXiv e-prints* **2018**, arXiv:1807.06209.

(11) Klypin, A. A.; Trujillo-Gomez, S.; Primack, J. *apj* **2011**, *740* 102, 102.

(12) López-Corredoira, M. *Foundations of Physics* **2017**, *47*, 711–768.

(13) Baumann, D. et al. In *American Institute of Physics Conference Series*, ed. by Dodelson, S.; Baumann, D.; Cooray, A.; Dunkley, J.; Fraisse, A.; Jackson, M. G.; Kogut, A.; Krauss, L.; Zaldarriaga, M.; Smith, K., 2009; Vol. 1141, pp 10–120.

(15) Eisenstein, D. J. et al. **2005**, *633*, 560–574.

(16) Bassett, B. A.; Hlozek, R. **2009**.

(17) Zamudio-Fernandez, J.; Okan, A.; Villaescusa-Navarro, F.; Bilaloglu, S.; Derin Cengiz, A.; He, S.; Perreault Levasseur, L.; Ho, S. *arXiv e-prints* **2019**, arXiv:1904.12846, arXiv:1904.12846.

(18) Bagla, J. S.; Khandai, N.; Datta, K. K. *Mon. Not. Roy. Astron. Soc.* **2010**, *407*, 567.

(19) Wyithe, J. S. B.; Loeb, A. *mnras* **2009**, *397*, 1926–1934.

(20) Wyithe, J. S. B.; Warszawski, L.; Geil, P. M.; Oh, S. P. *mnras* **2009**, *395*, 311–318.

(21) Zwaan, M. A.; Meyer, M. J.; Staveley-Smith, L.; Webster, R. L. *mnras* **2005**, *359*, L30–L34.

(23) Morales, M. F.; Wyithe, J. S. B. **2010**, *48*, 127–171.

(24) Pourtsidou, A. **2016**, *461*, 1457–1464.

(25) Wolz, L.; Blake, C.; Wyithe, J. S. B. **2017**, *470*, 3220–3226.

(26) Hall, A.; Bonvin, C. **2017**, *95* 043530, 043530.

(27) Carucci, I. P. In *Journal of Physics Conference Series*, 2018; Vol. 956, p 012003.

(28) Padmanabhan, H.; Choudhury, T. R.; Refregier, A. **2015**, *447*, 3745–3755.

(29) Pénin, A.; Umeh, O.; Santos, M. G. **2018**, *473*, 4297–4305.

(30) Santos, M.; Bull, P.; Alonso, D.; Camera, S.; Ferreira, P.; Bernardi, G.; Maartens, R.; Viel, M.; Villaescusa-Navarro, F.; Abdalla, F. B.; Jarvis, M.; Metcalf, R. B.; Pourtsidou, A.; Wolz, L. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 2015, p 19.

(31) Santos, M. G. et al. In *Proceedings, MeerKAT Science: On the Pathway to the SKA (MeerKAT2016): Stellenbosch, South Africa, May 25-27, 2016*, 2017.

(32) Battye, R. et al. *arXiv e-prints* **2016**, arXiv:1610.06826, arXiv:1610.06826.

(33) Villaescusa-Navarro, F.; Bull, P.; Viel, M. *Astrophys. J.* **2015**, *814*, 146.

(34) Cooray, A.; Sheth, R. K. *Phys. Rept.* **2002**, *372*, 1–129.

(35) Battye, R. A.; Browne, I. W. A.; Dickinson, C.; Heron, G.; Maffei, B.; Pourtsidou, A. *Monthly Notices of the Royal Astronomical Society* **2013**, *434*, 1239–1256.

(36) Kaiser, N. *Mon. Not. Roy. Astron. Soc.* **1987**, *227*, 1–21.

(37) Eisenstein, D. J.; Hu, W. *apj* **1998**, *496*, 605–614.

(38) Linder, E. V. *prd* **2005**, *72* 043529, 043529.

(40) Bernardeau, F.; Colombi, S.; Gaztañaga, E.; Scoccimarro, R. *physrep* **2002**, *367*, 1–248.

(41) Umeh, O.; Maartens, R.; Santos, M. *jcap* **2016**, *3* 061, 061.

(42) Umeh, O. *JCAP* **2017**, *06*, 005.

(43) Cooray, A.; Sheth, R. *physrep* **2002**, *372*, 1–129.

(44) Sheth, R. K.; Tormen, G. *mnras* **1999**, *308*, 119–126.

(45) Seehars, S.; Paranjape, A.; Witzemann, A.; Refregier, A.; Amara, A.; Akeret, J. *jcap* **2016**, *2016* 001, 001.

(46) Bagla, J. S.; Khandai, N.; Datta, K. K. **2010**, *407*, 567–580.

(47) Sarkar, D.; Bharadwaj, S.; Anathpindika, S. **2016**, *460*, 4310–4319.

(48) Villaescusa-Navarro, F.; Viel, M.; Datta, K. K.; Choudhury, T. R. **2014**, *2014* 050, 050.

(49) Villaescusa-Navarro, F.; Planelles, S.; Borgani, S.; Viel, M.; Rasia, E.; Murante, G.; Dolag, K.; Steinborn, L. K.; Biffi, V.; Beck, A. M.; Ragone-Figueroa, C. **2016**, *456*, 3553–3570.

(50) Barnes, L. A.; Haehnelt, M. G. **2010**, *403*, 870–885.

(51) Barnes, L. A.; Haehnelt, M. G. **2014**, *440*, 2313–2321.

(52) Padmanabhan, H.; Choudhury, T. R.; Refregier, A. **2016**, *458*, 781–788.

(53) Bullock, J. S.; Kolatt, T. S.; Sigad, Y.; Somerville, R. S.; Kravtsov, A. V.; Klypin, A. A.; Primack, J. R.; Dekel, A. *Monthly Notices of the Royal Astronomical Society* **2001**, *321*, 559–575.

(54) Padmanabhan, H.; Kulkarni, G. **2017**, *470*, 340–349.

(55) Padmanabhan, H.; Refregier, A.; Amara, A. *Mon. Not. Roy. Astron. Soc.* **2017**, *469*, 2323–2334.

(56) Masui, K. W.; Switzer, E. R.; Banavar, N.; Bandura, K.; Blake, C.; Calin, L. .-.-M.; Chang, T. .-.-C.; Chen, X.; Li, Y. .-.-C.; Liao, Y. .-.-W.; Natarajan, A.; Pen, U. .-.-L.; Peterson, J. B.; Shaw, J. R.; Voytek, T. C. **2013**, *763* L20, L20.

(57) Grasshorn Gebhardt, H. S.; Jeong, D.; Awan, H.; Bridge, J. S.; Ciardullo, R.; Farrow, D.; Gebhardt, K.; Hill, G. J.; Komatsu, E.; Molina, M.; Paulino-Afonso, A.; Saito, S.; Schneider, D. P.; Zeimann, G. *apj* **2019**, *876* 32, 32.

(58) Witzemann, A.; Bull, P.; Clarkson, C.; Santos, M. G.; Spinelli, M.; Weltman, A. **2018**, *477*, L122–L127.

(59) Zimmer, C. Experimental Design for Stochastic Models of Nonlinear Signaling Pathways Using an Interval-Wise Linear Noise Approximation and State Estimation., https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0159902&type=printable, [Online; accessed 2019], 2016.

(60) Fisher, R. A. *Journal of the Royal Statistical Society* **1935**, *98*, 39–82.

(61) Tegmark, M.; Taylor, A. N.; Heavens, A. F. *apj* **1997**, *480*, 22–35.

(62) Bassett, B. A.; Fantaye, Y.; Hlozek, R.; Kotze, J. *International Journal of Modern Physics D* **2011**, *20*, 2559–2598.

(63) Shimon, M.; Itzhaki, N.; Rephaeli, Y. *jcap* **2013**, *2013* 009, 009.

(64) Linder, E. V. *Astroparticle Physics* **2006**, *26*, 102–110.

(65) Bull, P.; Ferreira, P. G.; Patel, P.; Santos, M. G. *Astrophys. J.* **2015**, *803*, 21.

(66) Nishimichi, T.; Ohmuro, H.; Nakamichi, M.; Taruya, A.; Yahata, K.; Shirata, A.; Saito, S.; Nomura, H.; Yamamoto, K.; Suto, Y. **2007**, *59*, 1049.

(67) Jimenez, R.; Maartens, R.; Khalifeh, A. R.; Caldwell, R. R.; Heavens, A. F.; Verde, L. *JCAP* **2019**, *1905*, 048.

(68) Nesseris, S.; Trashorras, M. **2019**, *99* 063539, 063539.

(69) Komatsu, E. et al. **2009**, *180*, 330–376.

(70) Planck Collaboration et al. **2016**, *594* A13, A13.

(71) Baryshev, Y. V.; Sylos Labini, F.; Montuori, M.; Pietronero, L. *Vistas in Astronomy* **1994**, *38*, 419–500.

(72) Komatsu, E. et al. **2011**, *192* 18, 18.

(73) Riess, A. G. et al. **1998**, *116*, 1009–1038.

(74) Perlmutter, S. et al. **1999**, *517*, 565–586.

(75) Schwarz, D. J. In *Fundamental Interactions: A Memorial Volume for Wolfgang Kummer*, Grumiller, D., et al., Eds.; World Scientific Publishing Co: 2010, pp 267–276.

(76) Bharadwaj, S. Nonlinear distortions: The two-point correlation function.

(77) Pénin, A.; Umeh, O.; Santos, M. G. *Mon. Not. Roy. Astron. Soc.* **2018**, *473*, 4297–4305.

(78) M.zaldarriaga 21 centimeter fluctuations from cosmic gas at high redshifts.

(79) Magira, H. Cosmological Redshift-Space Distortion on Clustering of High-Redshift Objects: Correction for Nonlinear Effects in the Power Spectrum and Tests with N-Body Simulations., http://iopscience.iop.org/article/10.1086/308170/fulltext/.

(80) Percival, W. . Baryon Acoustic Oscillations., https://pdfs.semanticscholar.org/8cd5/c5d9.pdf.

(81) Padmanabhan Hydrogen in the post-reionization universe., https://arxiv.org/pdf/1712.01296.pdf.

(82) Okumura, T. Distribution function approach to redshift space distortions. Part III: halos and galaxies.

(83) B.Bassett, R. H. Baryon accoustic oscillations. Dark Energy, Ed. P. Ruiz-Lapuente.

(84) Bowman, J. HI and Cosmology: What We Need To Know.

(85) Raccanelli, A. Measuring redshift-space distortions with future SKA surveys.

(86) D.Wittman Fisher Matrix for Beginners., http://wittman.physics.ucdavis.edu/Fisher-matrix-guide.pdf.

(87) http://www.elementsdatabase.com/Hydrogen-H-1-element/.

(88) https://www.craf.eu/iau-list-of-important-spectral-lines/, (accessed: 01.09.2018).

(89) Marisa Advanced Statistical Methods for Astrophysical Probes of Cosmology textbook.

(90) M.Santos 21cm cosmology , NASSP MSc Lecture notes 2017.

(91) B.Bassett NASSP MSc Lecture notes 2017.

(92) Chang, T.-C.; Pen, U.-L.; Bandura, K.; Peterson, J. B. *arXiv e-prints* **2010**, arXiv:1007.3709, arXiv:1007.3709.

Appendix A

Bias formula derivation from Maximum Likelihood Estimation (MLE)

Bias formula in equation A.1 is derived from MLE principle.

$$\delta\theta_i = \sum_{j,k} F_{ij}^{-1} \frac{\Delta P_{HI}(k)}{\sigma(k)^2} \frac{\partial P_{HI}^{LIN}(k)}{\partial \theta_j} \tag{A.1}$$

Assume that we have a function of distribution that follows the the Gaussian law

$$f\left(x_k;\theta\right) = \frac{1}{(2\pi|c|)^{N/2}} exp\left[-\frac{1}{2}\left(x_k - H\theta\right)^T C^{-1}\left(x_k - H\theta\right)\right] \tag{A.2}$$

where C is the covariance matrix and

$$x_k = \sum_{l=1}^{p} h_l\theta_l + w_k = H\theta + w_k \tag{A.3}$$

so the likelihood function is as follows,

$$L\left(x_1, x_2, ..., x_k;\theta\right) = \frac{1}{(2\pi|c|)^{N/2}} exp\left[-\frac{1}{2}\sum_{k=1}^{k}\left(x_k - H\theta\right)^T C^{-1}\left(x_k - H\theta\right)\right] \tag{A.4}$$

Trying to find the $\theta = \theta_{MLE}$ that maximise $L\left(x_1, x_2, ..., x_k;\theta\right)$ given $x_1, ..., x_k$ is the same as maximising $ln\left[L\left(x_1, x_2, ..., x_k;\theta\right)\right]$

$$\theta_{MLE} = argmax_{[\theta]}\, ln\left[L\left(x_1, x_2, ..., x_k;\theta\right)\right] \tag{A.5}$$

Which is equivalent of

$$\theta_{MLE} = min_{[\theta]} \sum_{k=1}^{k} (x_k - H\theta)^T C^{-1} (x_k - H\theta) \tag{A.6}$$

$$\begin{aligned} \theta_{MLE} = min_{[\theta]}\ & \theta^T H^T C^{-1} H\theta \ - \ \theta^T H^T C^{-1} \sum_{k=1}^{k} x_k \\ & - \left(\sum_{k=1}^{k} x_k \right)^T C^{-1} H\theta \ + \ \sum_{k=1}^{k} x_k^T C^{-1} x_k \end{aligned} \tag{A.7}$$

$$\begin{aligned} \theta_{MLE} = min_{[\theta]}\ & \theta^T \left[H^T C^{-1} H \right] \theta \ - \ \theta^T H^T C^{-1} \bar{x} \\ & - \bar{x}^T C^{-1} H\theta \ + K \end{aligned} \tag{A.8}$$

$$\bar{x} \ = \ \frac{1}{k} \sum_{k=1}^{k} x_k$$

We can complete the square and we have

$$\begin{aligned} \theta_{MLE} = min_{[\theta]}\ & \left(\theta \ - \ \left[H^T C^{-1} H \right] H^T C^{-1} \bar{x} \right)^T H^T C^{-1} H \left(\theta \ - \ \left[H^T C^{-1} H \right] H^T C^{-1} \bar{x} \right) \\ & + K \ - \ \bar{x}^T C^{-1} H \left(H^T C^{-1} H \right)^{-1} H^T C^{-1} \bar{x} \end{aligned} \tag{A.9}$$

This yields

$$\theta_{MLE} \ = \ \left[H^T C^{-1} H \right]^{-1} H^T C^{-1} \bar{x} \tag{A.10}$$

Equation A.10 is the general form of equation A.1. We just need to substitute terms such that, $H = \frac{\partial P_{HI}}{\partial \theta}$, $x = \Delta P_{HI}$ and C is related to the covariance matrix of observational errors .

Appendix B

Bardeen equation

From Bardeen equation B.1 We can respectively get the matter and radiation densities contrast in equations B.6, B.7 in the Large scales structure.

Bardeen equation is written as follow,

$$\Phi'' + 3\mathcal{H}(1 + c_s^2)\Phi' + [2\mathcal{H}' + (\mathcal{H}^2 - K)(1 + 3c_s^2)]\Phi - c_s^2\Delta\Phi = 0 \quad \text{(B.1)}$$

and the comoving density contrast is

$$(\Delta + 3K)\Psi = \frac{\kappa}{2}a^2\rho\delta^c \quad \text{(B.2)}$$

Here $\Phi = \Psi$

Denote $y = \frac{a}{a_{eq}}$ then one can show that using the Bardeen equation(B.1) with two fluids we come up with these three governing equations given as follows:

$$\Phi'' + (7 - \frac{1}{1+y} + \frac{8}{4+3y})\frac{\Phi'}{2y} + \frac{\Phi}{(y(1+y)(4+3y))} = \frac{2}{(4+3y)y^2}(\delta^c - \frac{yS}{1+y}) \quad \text{(B.3)}$$

$$\delta^c = -\frac{4}{3}(\frac{k}{k_{eq}})^2\frac{y^2}{1+y}\Phi \quad \text{(B.4)}$$

$$S'' + \frac{3y+2}{2y(1+y)}S' = \frac{2}{4+3y}(\frac{k}{k_{eq}})^2(\delta^c - \frac{yS}{1+y}) \quad \text{(B.5)}$$

such that $S = \delta_m - \frac{3}{4}\delta_r$. Then for each fluid we can write:

$$\delta_m = \frac{3(1+y)\delta^c/4 + S}{1 + 3y/4} \tag{B.6}$$

$$\delta_r = \frac{(1+y)\delta^c - S}{1 + 3y/4} \tag{B.7}$$

As initial conditions we have : $\Phi = \Phi_i, \Phi' = 0, S = 0, S' = 0$ These initial conditions.

Let ζ be the curvature perturbation.

$$\zeta = \Phi + \frac{2}{3(1+\omega)\mathcal{H}}(\Phi' + \mathcal{H}\Phi) \tag{B.8}$$

For radiation $\omega = \frac{1}{3}$ then $\zeta = \frac{3}{2}\Phi_{rad}$ and for matter $\omega = 0$ then $\Phi_{rad} = \frac{3}{5}\zeta$

Thus,

$$\Phi_{mat} = \frac{9}{10}\Phi_{rad} \tag{B.9}$$