UNIVERSIDAD AUTÓNOMA DE
BAJA CALIFORNIA

FACULTAD DE CIENCIAS

The 21-cm line of Hydrogen (HI) in cosmology: Simulation of the observations of the SCI-HI experiment

Hiram Kalid Herrera Alcantar

Proyecto de tesis como requisito parcial
para obtener el título profesional de
Físico

Asesora:
Dra. Alma X. González-Morales
División de Ciencias e Ingenierías
Universidad de Guanajuato

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE CIENCIAS

THE 21-CM LINE OF HYDROGEN (HI) IN COSMOLOGY: SIMULATION OF THE OBSERVATIONS OF THE SCI-HI EXPERIMENT

TESIS PROFESIONAL

QUE PRESENTA

HIRAM KALID HERRERA ALCANTAR

APROBADO POR:

DRA. ALMA X. GONZÁLEZ-MORALES
DIRECTORA DE TESIS

DRA. PRISCILLA E. IGLESIAS VÁZQUEZ
CO-DIRECTORA DE TESIS

DR. HÉCTOR ACEVES CAMPOS
SINODAL

This thesis is dedicated to my parents,
Victoria Alcantar and Javier Herrera;
to my brother Victor Herrera;
and to all my friends
for their love, comprehension,
and endless support,
without them this success would not
have been possible.

Abstract of the thesis presented by **Hiram Kalid Herrera Alcantar** as partial fulfillment of the requirements for the **Bachelor of Science in Physics**. Ensenada, Baja California, Mexico. December, 2018.

The 21-cm line of Hydrogen (HI) in cosmology: Simulation of the observations of the SCI-HI experiment

The study of the 21-cm emission line of the neutral Hydrogen (HI) is a promising technique for our comprehension of the evolution of the Universe. In this thesis a brief review of the physics behind this process is given. The projects and experiments working on measuring this signal along with their current results are described. Within these experiments there is the SCI-HI experiment; SCI-HI is a experiment consisting in a single broadband antenna used to measure the 21-cm brightness temperature, that made preliminary observations in June 2013 at Isla Guadalupe, Mexico.

A methodology to analyze the data obtained by SCI-HI is developed, this methodology consist on a simulation of the observations done by the antenna by convolving its beam pattern with the Global Sky Model (GSM) in order to obtain a theoretical predicted measured temperature, this predicted value is used to calibrate the data via a χ^2 minimization method. The results from applying this methodology to a mock Gaussian antenna pattern and the real data from SCI-HI's experiment are given.

Keywords: **Cosmology - Reionization - Hydrogen - 21 cm signal - Simulation - data analysis - data calibration**.

Abstract approved by:

Dra. Alma X. González Morales
Thesis Director

Resumen de la tesis de **Hiram Kalid Herrera Alcantar** presentada como requisito parcial para la obtención de la **Licenciatura en Física**. Ensenada, Baja California, México. Diciembre de 2018.

La línea de 21-cm del Hidrógeno Neutro (HI) en la cosmología: Simulación de las observaciones del experimento SCI-HI

El estudio de la línea de emisión de 21 cm del Hidrógeno Neutro (HI) es una técnica prometedora para la comprensión de la evolución del Universo. En esta tesis se presenta una breve revisión de la física detrás de este proceso. Se describen los proyectos y experimentos usados para medir esta señal así como sus resultados actuales. Dentro de estos experimentos se encuentra el experimento SCI-HI; SCI-HI es un experimento que consiste en una antena de banda ancha única que se utiliza para medir la temperatura de brillo de 21-cm que realizó observaciones preliminares en Junio de 2013 en la Isla de Guadalupe, México.

Se desarrolla una metodología para analizar los datos obtenidos por SCI-HI, esta metodología consiste en la simulación de las observaciones realizadas por la antena, esto se hace convolucionando el patrón de la antena con el Modelo Global del Cielo (GSM, por sus siglas en inglés) para obtener la temperaura teórica. Este valor teórico se utiliza para calibrar los datos con el método de minimización de χ^2. Se muestran los resultados de aplicar esta metodología a un patrón Guassiano simulado y a los datos reales obtenidos por SCI-HI.

Palabras clave: **Cosmología - Reionización - Hidrógeno - Señal de 21 cm - Simulación - Análisis de datos - Simulación de datos**.

Resumen aprobado por:

Dra.Alma X. González Morales
Directora de Tesis

Contents

List of Figures

List of Tables

1 Introduction

Our understanding of the Universe, its structure formation and evolution, has increased significantly over the last 60 years. The exploration of the electromagnetic spectrum outside the visible waveband, such as radio, infrared, ultraviolet, X-ray bands, etc. has made this possible.

One of the most important observables in cosmology is the Cosmic Microwave Background (CMB), predicted in 1948 (Alpher and Herman, 1949; Gamow, 1948) and observed, circumstantially, in 1965 by Arno Penzias and Robert Wilson (Penzias and Wilson, 1965) when they were testing a radio receiver intended to use for radio-astronomy located in Bell Telephone Laboratories in Murray Hill, New Jersey. In these series of tests, Penzias and Wilson detected a uniform noise source, first assumed to come from the device, after many checks on the antenna and electronics they concluded that the noise came uniformly from the sky. This discovery was later concluded to be the CMB and both Penzias and Wilson won a Nobel Prize in Physics in 1978.

Over the past years, many missions have measured and studied the CMB. The first of them the Cosmic Background Explorer (COBE) (Smoot, 1999), measured the CMB over the sky and revealed the black body spectrum of the CMB with a temperature of 2.73 K and its uniformity almost all over the sky with small anisotropies. Later, in 2003, the Wilkinson Microwave Anisotropy Probe (WMAP) (Bennett et al., 2013) studied the fluctuations found by COBE more precisely and revealed interesting results constraining the cosmological model. The most recent studies of the CMB were made by the Planck Space Telescope (Ade et al., 2016; Aghanim et al., 2018), this mission confirmed the results by WMAP and constrained more precisely the dark energy and dark matter components of the current cosmological model.

Thanks to the studies of the CMB now we know that it reveals an early epoch in the Universe's history called "recombination" during which Universe cooled enough for the protons and electrons to combine forming mostly neutral Hydrogen (HI).

The next piece of information we have about the Universe's evolution is from an epoch called reionization, where neutral atoms collapse due to gravitational attraction forming the first structures such as stars.

There is an epoch of the Universe between the two described above, of which we know very little, the "dark ages". This epoch represents one of the frontiers of the modern cosmology, since there is no structure that emits any type of radiation and therefore it is not possible to detect this epoch using the conventional techniques in astronomy.

A promising technique for studying the dark ages and reionization is through the 21-cm transition of Hydrogen (Furlanetto et al., 2006). This is the transition between the two hyperfine states of the hydrogen ground state. The hyperfine states have an energy difference of $\Delta E = 5.9 \times 10^{-6}$ eV, corresponding to a wavelength of $\lambda = 21.1$ cm and $\nu = 1420.4$ MHz frequency.

The study of this emission/absorption line has drawn attention since its prediction in 1945 (Van de Hulst, 1945) and its later observation in 1951 (Ewen and Purcell, 1951). Hydrogen is the most abundant element in the Universe and it has existed since the CMB epoch so the emission/absorption of the signal is expected to occur through all history of the Universe after the formation of the first hydrogen atoms. The main limitation to use this technique lies in the detection of this signal above the foreground and noise. The original 21 cm wavelength radiation emitted in the dark ages is redshifted due to the expansion of the Universe, thereby, the possible measurement of the signal requires instruments that work on wavelengths of the order of meters (radio-frequencies).

Given the noise an instrument can be exposed to, such as digital transmission emitters, like television, FM/AM radio and mobile phones, the instrument's proper noise, and even extraterrestrial sources like the Galaxy's radiation, this signal is really difficult to detect.

There are plenty of experiments and projects around the world trying to measure this signal, such as EDGES (Experiment to Detect the Global EoR Signature) (Bowman et al., 2008), LEDA (Large-Aperture Experiment to detect Dark Ages) (Greenhill and Bernardi, 2012), DARE (Dark Ages Radio Explorer) (Burns et al., 2012), BIGHORNS (Broadband Instrument for Global Hydrogen Reionisation Signal) (Sokolowski et al., 2015) and the Mexican experiment SCI-HI (Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro) (Voytek et al., 2014).

The main goal of this thesis is to develop a methodology to calibrate the measurements done by the SCI-HI experiment, in order to characterize the 21 cm signal. The development of this procedure is expected to serve in future observations by the SCI-HI experiments at the end of the current year.

This thesis is structured as follows:

- Fundamental definitions of cosmology and pertinent equations for this thesis are given in chapter 2.
- A description of the 21-cm signal physics and construction of the differential brightness temperature observable are given in chapter 3.
- In chapter 4 we present the state of the art on the observations and measurements of the signal, along with the description of the experiments trying to achieve this task. A brief description of parameterization used to model the signal will be given and further details of the SCI-HI experiment with its preliminary results.
- The methodology developed to simulate the observations as seen by a particular design (antenna beam pattern), the observation date and location is described in chapter 5.
- A description of the calibration method used for the HIbiscus antenna data taken in 2013 along with the results of applying a χ^2 minimization method are is described in chapter 6.
- Conclusions of the results obtained and perspectives for future observations to be done by a new antenna of the SCI-HI experiment called Mango Peel are given in chapter 7.

At the end of this thesis the developed codes used for the analysis are given in Appendices A, B, C and D; these codes were written in python language using the 2.7.15 version.

2 Fundamentals of Cosmology

Cosmology studies the origin and evolution of the Universe and its components, the theories in this area have evolved significantly in the past century. So we may give a historical review of cosmology and the ideas within it. Before describing the 21-cm transition line of Hydrogen we must first introduce some definitions used in cosmology like the cosmological principle, the expansion of the Universe, the dynamics involved in this process and the mathematics behind basic cosmology needed to understand the importance of the 21-cm line for cosmology[1].

2.1 The cosmological principle

The way how we describe the Universe and our location in it has changed through the history of humanity. Starting from the ancient belief of the geocentric Universe, where the Earth would lie in the center of the Universe, and that the Sun, Moon, planets and the stars would orbit the Earth following some strange orbits. This idea was later discarded by the introduction of Nicolaus Copernicus Heliocentric model, placing the Sun in the center of the solar system, meaning that the Earth had no

[1]Most of the content of this chapter was taken from different books like Ryden (2016), Liddle (2015) and Weinberg (2008).

such preferred place. But the question of whether the Sun was still on a special place of the Universe or not, was not answered until the studies made by William Herschel. He revealed, by studying the nearby stars, that the Sun is a typical star orbiting an assembly of stars and dust which we now know as the Milky Way galaxy. In 1952, the observations made by Walter Baade helped to prove that the Milky Way is a fairly typical galaxy, totally discarding the idea of us being in any special place in the Universe.

Nowadays the idea of there being a center, or any preferred position in the Universe is totally discarded. The idea of an isotropic and homogeneous Universe is one of the pillars of Cosmology. Isotropy states that the Universe looks the same in any direction we look, while homogeneity states that the Universe looks the same at each point, in other words there is no such special location in the Universe, this is what we call *the cosmological principle*.

It is important to remark that the cosmological principle is not applied to any scale of the Universe. For example, neither the Earth nor the Solar system are close of being isotropic or homogeneous systems. It does not even apply if we observe the galaxies in our local group. It is not until we observe the large scales of the Universe, around 100 Mpc[2] or more, where we can see that the cosmological principle holds.

2.2 Expansion of the Universe

One of the most important observations made in cosmology is that most of the galaxies and objects in the Universe appear to be moving away from us, and the

[2] A parsec (pc) is defined as the distance at which one astronomical unit subtends an angle of one arc second of parallax, this is $1\text{pc} \approx 206265\text{AU} \approx 3.086 \times 10^{16}\text{m}$.

more distant an object is the faster it appears to recede.

Usually the velocity of recession of an object in the Universe is measured via the *redshift*, which is a generalization of the Doppler effect applied to electromagnetic radiation. Most objects in the Universe, like galaxies and stars, have a set emission and absorption lines within its spectrum. The frequencies/wavelengths of these lines are well-defined, however if the object is moving towards us they will appear moved towards the blue end of the electromagnetic spectrum, this is known as blueshift. On the other hand, if the object is moving away the lines move to towards the red end of the spectrum causing the effect of redshift.

The redshift of a receding emitter is given by the formula

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}}, \tag{2.1}$$

where λ_{em} and λ_{obs} are the wavelengths of the light at the emission point and observation point (Earth), respectively, this equation can also be written as

$$1 + z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{em}}}. \tag{2.2}$$

In 1929 Edwin Hubble measured the redshift of many galaxies, his observations leaded to a linear relationship between the redshift of an observed galaxy and its distance relative to Earth, this is now known as the Hubble's Law;

$$z = \frac{H_0}{c} r, \tag{2.3}$$

where H_0 the Hubble's constant, c is the speed of the light and r is the distance of the object with respect to the Earth. The Hubble's law has another interpretation, if we take into account that Hubble observed nearby galaxies, then we can interpret the redshift of the galaxy as a shift due to the classical, non relativistic Doppler effect, this means that the redshift is related with the velocity v of the galaxy as $z = v/c$, using this relationship the Hubble's Law takes the form

$$v = H_0 r. \tag{2.4}$$

From equation (2.4) the value of the Hubble constant can be determined by dividing the velocity by the distance of the galaxies. Since the common units that these quantities are measured are Mpc and km s^{-1}, respectively, the Hubble constant is usually expressed in km s^{-1}Mpc^{-1} units.

At first the Hubble's law looks like a violation of the cosmological principle since it may seem like the Earth is in the center of the Universe and all the other galaxies are moving away from it. In fact this is what we should expect, it is true that we look that galaxies are moving away from us, but if an observer goes to another galaxy it would also observe that the galaxies are moving away from it. This is in fact an effect of the cosmological principle and the expansion of the Universe.

To clarify this effect, consider three objects in the Universe at positions $\vec{\mathbf{r_1}}$, $\vec{\mathbf{r_2}}$ and $\vec{\mathbf{r_3}}$, these define a triangle with sides of length r_{12}, r_{23} and r_{31}, with

$$r_{ij} = \left|\vec{\mathbf{r_i}} - \vec{\mathbf{r_j}}\right|, \tag{2.5}$$

if we now expand the distance between them uniformly and homogeneously, the shape of the triangles is invariant, however their relative distance changes. We may introduce the *comoving coordinates* related by the *scale factor* $a(t)$, with this definition the relative distance between the objects at a time t is given by

$$\vec{\mathbf{r_i}}(t) = a(t)\vec{\mathbf{r_i}}(t_0). \tag{2.6}$$

The scale factor measures the Universal expansion (possibly contraction) rate, it only depends on time and it represents how the separation between the objects is increasing or decreasing with time, while the initial coordinates $\vec{\mathbf{r_i}}(t_0)$ are fixed.

We can also express redshift in terms of the scale factor so that

$$1 + z = \frac{1}{a}. \tag{2.7}$$

2.3 Geometry of the Universe

In 1915 Einstein published one of the biggest theories in modern physics, the General Theory of Relativity, where gravity is described as a geometric property of space-time, and that the curvature of space-time is directly related by the mass (Einstein, 1915).

The cosmological principle demands isotropy and there are three possible curvatures that satisfy this statement, each with different implications:

Flat geometry: this geometry corresponds to zero curvature. It implies an infinite Universe, since the existence of an edge would violate the homogeneity statement of the cosmological principle. A Universe with this geometry is called *flat Universe*.

Spherical geometry: corresponds to positive curvature, a Universe with this geometry has a finite size but no boundary, due to these properties this kind of Universe is called as *closed Universe* in extent.

Hyperbolic geometry: the curvature for this geometry is negative, the properties of this geometry imply an infinite Universe in extent, this case is known as *open Universe*.

2.4 The Friedmann-Lemaître-Robertson-Walker Metric

The mathematical way of describing the space-time curvature is through the *metric*, which describes the physical distance between events in space-time and the geometry of space-time.

A metric that can hold the conditions of homogeneity and isotropy and also allows distances to expand was derived by the different contributions of Alexander Friedmann, Georges Lemaître, Howard Robertson and Arthur Walker (Ryden, 2016). This metric is now known as the Friedmann-Lemaître-Robertson-Walker (FRLW) Metric, and its described mathematically as

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right], \tag{2.8}$$

where k is the curvature of the space, s is the comoving distance, t the time and r,θ and ϕ describe the spherical coordinates.

2.5 Dynamics in cosmology

The way the metric evolves is related with the Einstein's Field equations, these are a set of 10 equations that describe the interaction of gravitation as a result of the curvature of space-time due to the presence of mass and energy.

By considering a perfect fluid and the FLRW metric in the Einstein Field equations we can obtain the following equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \tag{2.9}$$

where ρ is the mass density and G is the gravitational constant. The term $\frac{\dot{a}}{a}$ is often written as H and it is called the *Hubble parameter*. This equation is the standard form of the *Friedmann equation* and it describes the rate at which the Universe is expanding. However, it can not be solved without a notion of how the mass density evolves through time, for this we need to consider a material of mass density ρ with a pressure P.

The equation, know as the *fluid equation*, that describes how the density of the Universe evolves while time passes is

$$\dot{\rho} + 3\frac{\dot{a}}{a}\left(\rho + \frac{P}{c^2}\right) = 0. \tag{2.10}$$

We can derive a third equation by combining the two equations given above,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right), \tag{2.11}$$

this is the *acceleration equation*, it tells us how the expansion of the Universe speeds up or slows down with time.

The three equations mentioned above have three unknowns, the scale factor a, the mass density ρ and the pressure P, then three equations are needed in order to solve them, however only two of the three equations mentioned above are independent, there is need of another equation. For that we introduce the *equation of state*, which describes the relation between the pressure and the density of the material that the Universe is made of, it is assumed that the pressure is an exclusive function of the density so that $P = P(\rho)$, for the materials considered in cosmology this relation is linear

$$P = \omega \rho c^2, \tag{2.12}$$

where ω is a dimensionless number related to the type of material considered, its different values will be described later in this chapter.

Another consideration to be done, in the dynamics of the expansion, is that the observations made so far conclude that the Universe is in fact expanding in an accelerated form. However, the acceleration equation and the Friedmann equation itself don't seem to fit these observations. We introduce the *cosmological constant* Λ to fix this issue. The equations with the cosmological constant added are

$$H^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3}, \tag{2.13}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) + \frac{\Lambda c^2}{3}, \tag{2.14}$$

while the fluid equation remains unaffected.

2.6 Components of the Universe

As mentioned in the previous section to solve the fluid equation we need to use the equation of state that is related with the kind of material considered in the Universe, the kind of materials considered in cosmology are:

- **Matter:** we refer as matter to any type of non relativistic material which exerts no pressure. Matter is divided in baryonic matter and dark matter, even though these are very different they follow the same equation of state with $\omega = 0$.
- **Radiation:** particles with a relativistic movement, like photons and neutrinos, their kinetic energy makes them exert radiation pressure, which has an equation of state with $\omega = 1/3$.
- **Cosmological constant:** we can consider the cosmological constant as a fluid with an associated energy density. It follows the equation of state with $\omega = -1$.

Considering a single component and spatially flat Universe and substituting the values of ω described above in the equations given in section 2.5, we obtain the solutions given in Table 2.1.

Component	**Equation of state**	$\rho(a)$	$a(t)$	$\rho(t)$
Matter	$P = 0$	$\rho_{m,0}a^{-3}$	$(t/t_0)^{2/3}$	$\rho_{m,0}t_0^2/t^2$
Radiation	$P = \rho c^2/3$	$\rho_{r,0}a^{-4}$	$(t/t_0)^{1/2}$	$\rho_{r,0}t_0^2/t^2$
Cosmological Constant	$P = -\rho c^2$	$\rho_{\Lambda,0}$	$\exp\left[H_0(t-t_0)\right]$	$\rho_{\Lambda,0}$

Table 2.1: Solutions of the fluid and Friedmann equations for a spatially flat single component Universes. The zero suffix expresses the value at a present time. $H_0 = 8\pi G\rho/3$, $t_0 \sim H_0^{-1}$.

To solve the equations of a Universe made by a mixture of components that is not necessarily spatially flat is more complicated, the solution would be in terms of the densities of every component, $\rho = \rho_m + \rho_r + \rho_\Lambda$, for a simpler view of the behaviour of the Universe we can rewrite the Friedmann equation as

$$H^2 = H_0^2\left[\frac{\Omega_{r,0}}{a^4} + \frac{\Omega_{m,0}}{a^3} + \Omega_{\Lambda,0} - \frac{1-\Omega_0}{a^2}\right], \tag{2.15}$$

where $\Omega_{i,0} = \rho_{i,0}/\rho_c$ is the current density of every component[3] and $\Omega_0 = \Omega_{m,0} + \Omega_{r,0} + \Omega_{\Lambda,0}$.

All the values $\Omega_{i,0}$ along with H_0 in equation (2.15) have been measured by different missions and experiments trying to constrain the current cosmological model, the current values given by the study of CMB, given in (Aghanim et al., 2018), are

Parameter	**Value**
H_0	67.37 ± 0.54 km s^{-1}Mpc^{-1}
$\Omega_{m,0}$	0.3147 ± 0.0074
$\Omega_{\Lambda,0}$	0.6853 ± 0.0074
$\Omega_{r,0}$	$\sim 10^{-4}$
Ω_0	1.0000 ± 0.0074

Table 2.2: Current values of the cosmological parameters given by (Aghanim et al., 2018).

2.6.1 Time evolution of the radiation temperature

For the purpose of this thesis we may derive how the temperature of radiation varies as the Universe expands, for this we may consider the first law of thermodynamics

[3] The critical density ρ_c for a given value H is the density which would be required in order to make the geometry of the Universe flat. It equals to $\rho_c = 8\pi G/3H^2$

$$\mathrm{d}E = \mathrm{d}Q - P\,\mathrm{d}V\,, \tag{2.16}$$

where E is the internal energy, Q the heat, V is the volume and P the pressure. Considering an adiabatic expansion ($\mathrm{d}Q = 0$) of the Universe then we get

$$\mathrm{d}E = -P\,\mathrm{d}V\,, \tag{2.17}$$

considering that $E = \varepsilon V$, where ε is the energy density of radiation given by $\epsilon = \alpha T^4$, where $\alpha = 7.565 \times 10^{-16}$J m^{-3}K^{-4} (this is considering that the radiation comes from a blackbody), now since $P = \rho c^2/3 = \varepsilon/3$, we obtain

$$\alpha\big(4T^3 V\,\mathrm{d}T + T^4\,\mathrm{d}V\big) = -\frac{1}{3}\alpha T^4\,\mathrm{d}V\,, \tag{2.18}$$

after rearranging terms and simplifying we obtain

$$\frac{1}{T}\,\mathrm{d}T = -\frac{1}{3V}\,\mathrm{d}V\,, \tag{2.19}$$

now since $V \propto a^3$ we get that $\mathrm{d}V = 3a^2\,\mathrm{d}a = \frac{3V}{a}\,\mathrm{d}a$, we get that

$$\frac{1}{T}\,\mathrm{d}T = -\frac{1}{a}\,\mathrm{d}a\,, \tag{2.20}$$

the solution to this equation is

$$T = \frac{T_0}{a} \tag{2.21}$$

or in terms of redshift

$$T = T_0(1 + z). \tag{2.22}$$

2.7 Timeline of the Universe

With the current standard cosmological model we have an idea of how the Universe has evolved through time, for a better understanding this evolution is usually divided in epochs where a certain event happened (perhaps), these epochs are:

- **Big Bang:** it is theorized that the Universe began in a Big Bang, a space-time singularity with a high density and high temperature. However, it can not been observed and the current standard cosmological model can't explain this event.

- **Inflation (possibly):** this early phase of the Universe is subject to speculation, the current theories say that there is an epoch where the Universe would go through an accelerated expansion. This theory is the possible explanation of why the current curvature of the Universe seems to be flat.

- **Nucleosynthesis:** is the epoch where the Universe has cooled down enough to form the firms atomic nuclei, like hydrogen, from the previously existing nucleons, protons and neutrons. During Nucleosynthesis the only elements produced with a significant abundance were Hydrogen and Helium-4 with fraction of mass of the Universe of ~ 0.77 and ~ 0.23 respectively.

- **Recombination:** before this epoch the dominating material in the Universe was radiation, then the Universe cooled down to a point where the radiation density equals the matter density to give point to an epoch where matter domains in the evolution of the Universe. At this point energy is not enough to maintain the previously formed charged atoms ionized, letting neutral atoms to form. During recombination photons decouples from matter causing what we see now as the Cosmic Microwave Background.

- **Dark Ages:** after the emission of the CMB, the Universe was made primarily of neutral atoms and no other structures exist, so at this point there is no radiation emitted by the Universe other than the CMB and radiation occasionally released/absorbed by the neutral hydrogen (21 cm transition).
- **First stars formation:** during Dark Ages some dense regions formed in the Universe, these became denser (due to gravitational forces) and hotter until they started to burn hydrogen, forming the first stars.
- **First galaxies formation:** the dense regions of the Universe became massive due to gravitational collapse, gathering millions of stars in this process, then forming the first galaxies.
- **Reionization:** after the formation of the first luminous structures (stars and galaxies) their intense radiation started to ionize (again) the intergalactic medium.
- **Present:** the observations and current cosmological model predict that the Universe has 13.7 billion years old. Also, we observe the CMB at a temperature of approximately $T = 2.725$K. However not much of the Universe is known, so we keep studying it through different techniques.

3 | The Hydrogen's 21-cm line

The 21-cm signal is produced by a transition in the hyperfine structure of the Hydrogen atom. The hyperfine structure arises from the interaction of an atomic nucleus and the electrons, at the magnetic moments level.

In this chapter we will describe the physics involved in the 21-cm signal, like the collisional coupling and the Wouthuysen-Field effect. Then we make use of the radiative transfer equation to construct the differential brightness temperature observable.

3.1 Atomic principles of the 21-cm line

For hydrogen we have electron-proton interaction. The ground energy state of the Hydrogen ($1S$) splits in two hyperfine states. $F = 1$ corresponds to the case when the spins of both electron and proton are parallel (triplet state), and $F = 0$ for the anti-parallel case (singlet state). The energy difference between those states is $\Delta E = 5.9 \times 10^{-6}$ eV. When the Hydrogen state decays from $F = 1$ to $F = 0$, it emits a photon with this amount of energy (see Fig. 3.1). The transition from $F = 0$ to $F = 1$ can also occur if a photon of energy ΔE is absorbed. The 21-cm emission/absorption line is produced by the aforementioned transitions.

The wavelength λ and frequency ν, associated to the energy shift necessary to produce a transition is given by Planck's equation

$$\begin{aligned} \lambda &= \frac{hc}{\Delta E} = 21.106\,\mathrm{cm}, \\ \nu &= \frac{\Delta E}{h} = 1420.41\,\mathrm{MHz}, \end{aligned} \tag{3.1}$$

where h is the Planck constant and c is the speed of light.

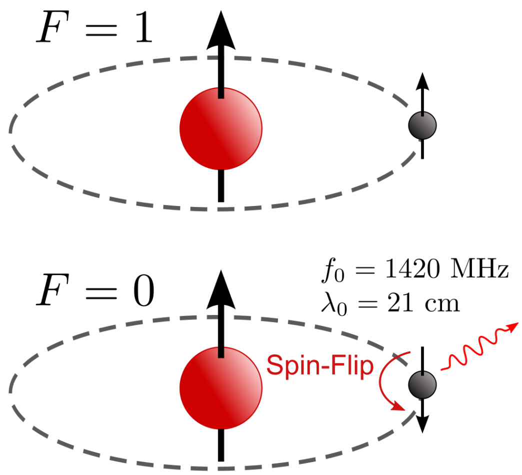


Figure 3.1: 21-cm transition between the two hyperfine states of neutral Hydrogen[1].

3.2 The 21-cm line in cosmology

As we know, from Nucleosynthesis (see section 2.7), the Hydrogen is by far the most abundant element in the Universe and it has existed since the Universe cooled enough for the electrons to merge with the nucleons (redshift $z \approx 1100$), therefore

[1]Source: https://bit.ly/2Qncktz

we can expect the 21 cm transition to happen in most of the Universe's epoch. This transition can be seen as a signal in emission or absorption.

Similar to the CMB characteristic temperature we can construct an observable for the 21-cm signal called brightness temperature, T_b. This is a function of the redshift z, since it depends on how the Universe evolves. The details about physics behind this observable is as follows.

3.3 Physics of the 21-cm line of neutral Hydrogen

The radiative transfer equation of a source emitting a 21 cm wavelength will be used to detail the 21-cm signal; suppose that a source emits a photon with a specific intensity I_ν, then this photon travels through a medium along a path described by coordinate s, the radiative transfer equation is

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = -\alpha_\nu I_\nu + j_\nu, \tag{3.2}$$

where the coefficients α_ν and j_ν describe the absorption and emission by the medium respectively.

Let us define the optical depth of the medium as the measure of the absorption occurring while light travels through the medium of a specific depth,

$$\tau_\nu(s) = \int \alpha_\nu(s')\,\mathrm{d}s'\,, \tag{3.3}$$

using Equation (3.3), we can rewrite Equation (3.2) as

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = -I_\nu + S_\nu, \tag{3.4}$$

where we have defined the ratio of the emission coefficient j_ν and the absorption coefficient α_ν as the source function, $S_\nu = j_\nu/\alpha_\nu$. By multiply both sides of equation (3.4) by e^{τ_ν} and rearranging terms, the equation obtained is

$$e^{\tau_\nu}\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} + e^{\tau_\nu}I_\nu = \frac{\mathrm{d}(e^{\tau_\nu}I_\nu)}{\mathrm{d}\tau_\nu} = e^{\tau_\nu}S_\nu, \tag{3.5}$$

the result of integrating equation (3.5) by parts is

$$I_\nu = I_{\nu,0}e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(\tau'_\nu)e^{-(\tau_\nu-\tau'_\nu)}\,\mathrm{d}\tau'_\nu\,. \tag{3.6}$$

Now let us consider that radiation is in Local Thermodynamic Equilibrium (LTE) with the medium, then the source function can be written as function of temperature (Rohlfs and Wilson, 2013) by

$$S_\nu(T) = B_\nu(T), \tag{3.7}$$

where $B_\nu(T)$ is the Planck function defined as

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/k_BT}-1}, \tag{3.8}$$

where k_B is the Boltzmann constant. Then the intensity is given by

$$I_\nu = I_{\nu,0}e^{-\tau_\nu} + \int_0^{\tau_\nu} B_\nu(T)e^{-(\tau_\nu-\tau'_\nu)}\,\mathrm{d}\tau'_\nu\,. \tag{3.9}$$

If the intergalactic medium is isothermal ($T =$ constant) then the solution to equation (3.9) is

$$I_\nu = I_{\nu,0} e^{-\tau_\nu} + B_\nu(T)(1 - e^{-\tau_\nu}). \tag{3.10}$$

Using a relation similar to equation (3.7), we can characterize the brightness temperature of an object at a frequency ν, as the temperature a black body would have the same intensity I_ν at the same frequency, this is,

$$I_\nu(T_b) = B_\nu(T_b), \tag{3.11}$$

it is correct to use assume a black body temperature distribution since the only background source we are considering is the CMB, which has been proof to have a black body spectrum.

We may use the Rayleigh-Jeans limit ($h\nu \ll kT$) because the range of frequencies where the 21-cm signal is relevant are much smaller than the peak frequency of the CMB black body radiation, then

$$e^{h\nu/k_B T} \approx 1 + \frac{h\nu}{k_B T}, \tag{3.12}$$

using this approximation, then, equation (3.8) is rewritten as

$$B_\nu(T) = 2\frac{\nu^2}{c^2} k_B T. \tag{3.13}$$

Through this relation and the considerations given above we can rewrite equation (3.10) as

$$T_b = T_\gamma e^{-\tau_\nu} + T_{\rm ex}(1 - e^{-\tau_\nu}), \tag{3.14}$$

where T_γ is the brightness temperature of the background radiation field, $T_{\rm ex}$ refers to the excitation temperature of the medium caused by absorption.

3.4 The spin temperature

For the 21-cm line the excitation temperature is known as the spin temperature T_S, and it is defined through the equation

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-E_{10}/k_B T_S}, \tag{3.15}$$

where n_1 and n_0 are the number densities of atoms in the triplet and singlet states of the hyperfine level respectively (Furlanetto et al., 2006), g_1 and g_0 are the statistical weights of these states, in this case $g_0 = 1$ and $g_1 = 3$, and E_{10} is the energy difference between the states ($E_{10} = 5.9 \times 10^{-6}$ eV).

Defining the equivalent temperature $T_* = E_{10}/k_B = 0.068$ K we obtain the following equation to describe the spin temperature given the relationship between the number densities

$$\frac{n_1}{n_0} = 3e^{-T_*/T_S}. \tag{3.16}$$

The value of T_S principally depends on three processes (Furlanetto et al., 2006):

i) The absorption or emission of 21 cm photons from the CMB radiation.

ii) Collisions between hydrogen atoms and atom-electron collisions.

iii) Scattering of Lyman-alpha ($Ly\alpha$) photons.

The photons in Lyman-alpha scattering are emitted by the de-excitation of neutral Hydrogen changing from the $2P$ state to the ground state $1S$.

The equation of the spin temperature is given by (Field, 1958)

$$T_S^{-1} = \frac{T_\gamma^{-1} + x_c T_K^{-1} + x_\alpha T_c^{-1}}{1 + x_c + x_\alpha}, \tag{3.17}$$

where x_c and x_α are the coupling coefficients for collisions and scattering of $Ly\alpha$ photons respectively, T_c is the color temperature of the $Ly\alpha$ radiation, T_γ is the temperature of the background radiation, in this case T_γ will be set by the CMB so $T\gamma = T_{\mathrm{CMB}}$, and T_K is the gas kinetic temperature.

3.5 Collisional coupling

The coefficient x_c is related to the excitation and de-excitation of the hyperfine states caused by the collision of particles. These can be: hydrogen-hydrogen (HH), hydrogen-electron (eH) and hydrogen-proton (pH), although there can be collisions between a hydrogen atom and other types of particles, we will only consider the three mentioned above.

The coefficient for every type of collision is given by

$$x_c^i = \frac{n_i \kappa_{10}^i}{A_{10}} \frac{T_*}{T_\gamma}, \tag{3.18}$$

subscript i denotes the species of the collision (HH, eH, pH), and κ_{10}^{i} is rate coefficient for spin de-excitation by collisions between hydrogen atoms and the i particle species, given in units of cm^3/s.

The total collisional coupling coefficient is the sum over all the processes,

$$\begin{aligned} x_c &= x_c^{\mathrm{HH}} + x_c^{\mathrm{eH}} + x_c^{\mathrm{pH}} \\ &= \frac{T_*}{A_{10}T_\gamma}\left(n_{\mathrm{H}}\kappa_{10}^{\mathrm{HH}} + n_{\mathrm{e}}\kappa_{10}^{\mathrm{eH}} + n_{\mathrm{p}}\kappa_{10}^{\mathrm{pH}}\right). \end{aligned} \tag{3.19}$$

The rate coefficients for de-excitation for every species changes with the kinetic temperature of the measured gas cloud, thus it is important to know their values for the epoch of interest. For example, the cosmic dark ages where coupling is dominated by these collisions; a further analysis of this behavior is given in Furlanetto et al. (2006).

3.6 The Wouthuysen-Field effect

In the previous section, the coupling due to particle collisions was analyzed, however, this process is only relevant during the dark ages of the Universe. Once that star formation begins, the scattering of $Ly\alpha$ photons becomes an important source of coupling, the mechanism from which this coupling process occurs is known as the Wouthuysen-Field effect (Wouthuysen, 1952).

To illustrate this effect, suppose a hydrogen in the hyperfine singlet state, if this hydrogen absorbs a $Ly\alpha$ photon, then it will get excited to any of the central $2P$ hyperfine states, due to the dipole selection rules allow $\Delta F = 0, 1$, and the transition $F = 0 \rightarrow 0$ is prohibited, only four of the six possible states are allowed.

After the $Ly\alpha$ photon is absorbed, the atom the ${}_0S^{\frac{1}{2}}$ state can only jump to the ${}_1P^{\frac{1}{2}}$ and ${}_1P^{\frac{3}{2}}$ states. After the absorption, relaxation may occur by emitting a $Ly\alpha$ photon while the atom decays to any of the two ground state hyperfine levels. If relaxation occurs to the triplet state ${}_1S^{\frac{1}{2}}$ then a spin-flip occurred, thus the absorption and re-emission of $Ly\alpha$ photons can produce these spin-flips relevant to the 21-cm signal.

The coupling coefficient x_α associated to this process may be written as

$$x_\alpha = \frac{P_{10}}{A_{10}} \frac{T_*}{T_\gamma}. \tag{3.20}$$

We can relate the scattering rate between the two hyperfine states with the scattering rate of the $Ly\alpha$ photons P_α by the relation $P_{10} = 4P_\alpha/27$, this is the result assuming that the radiation field is constant, so to calculate x_α we just need to know the value of P_α, however, the procedure to find this value goes beyond the purpose of this thesis, more details of this analysis can be found in Furlanetto et al. (2006), Pritchard and Loeb (2012) and Burgueño (2017).

3.7 Differential brightness temperature

According to Furlanetto et al. (2006) the optical depth of a cloud of hydrogen is

$$\tau_\nu = \int \mathrm{d}s \left[1 - e^{-T_*/T_S}\right] \sigma_0 \phi(\nu) n_0, \tag{3.21}$$

where $\sigma_0 = 3c^2 A_{10}/8\pi\nu^2$ is the effective cross-section of the cloud, $A_{10} = 2.85 \times 10^{-15}$ s^{-1} is the spontaneous decay rate of the spin-flip transition, n_0 is the hydrogen density $n_{\rm HI}$ divided by four ($n_0 = n_{\rm HI}/4$), this division is due to the fraction of hydrogen atoms in the hyperfine singlet state, and $\phi(\nu)$ is the line profile normalized so that $\int \mathrm{d}\nu\, \phi(\nu) = 1$.

Solving equation (3.21) we get that the optical depth of a hydrogen cloud while the Universe is expanding is (Furlanetto et al., 2006)

$$\tau_\nu = \frac{3}{32\pi}\frac{hc^3 A_{10}}{k_B T_S \nu_0^2}\frac{x_{\rm HI} n_{\rm HI}}{(1+z)(\mathrm{d}v_\parallel/\mathrm{d}r_\parallel)}, \tag{3.22}$$

where $x_{\rm HI}$ is the neutral fraction of hydrogen, and $\mathrm{d}v_\parallel/\mathrm{d}r_\parallel$ is the gradient of the proper velocity of the cloud along the line of sight, including the Hubble expansion and the peculiar velocity.

Finally, using equation (3.14), we can construct the differential brightness temperature δT_b by contrasting the differential temperature T_b with the CMB temperature T_γ at high redshifts

$$\delta T_b = T_b - T_\gamma = \frac{T_S - T_\gamma}{1+z}\left(1 - e^{-\tau_\nu}\right), \tag{3.23}$$

where we have used the fact that both temperatures T_b and T_γ scale proportional to redshift such that $T \propto (1+z)$, now if we expand this equation using Taylor's theorem and assume $\tau_\nu \ll 1$,

$$\delta T_b \approx \frac{T_S - T_\gamma}{1+z}\tau_\nu, \tag{3.24}$$

using equation (3.22) along with equation (3.24), we get that the differential brightness temperature is given by

$$\delta T_b\left(\nu\right) = 27 x_{\mathrm{HI}} (1+\delta) \left(\frac{1+z}{10}\right)^{1/2} \left(1 - \frac{T_\gamma}{T_S}\right) \left[\frac{H(z)/(1+z)}{\mathrm{d}v_\parallel/\mathrm{d}r_\parallel}\right] \mathrm{mK}, \tag{3.25}$$

where we have substituted the velocity $H(z)/(1+z)$ appropriate for the uniform Hubble expansion at high redshifts (Furlanetto et al., 2006), the factor $(1+\delta)$ is the fractional overdensity of baryons.

It is important to note that the detection of the differential brightness temperature depends on the value of T_S compared to T_γ, for example if $T_S \gg T_\gamma$ then δT_b saturates, while it can become arbitrarily large and negative if $T_S \ll T_\gamma$, also if $T_S = T_\gamma$ there will be no signal detected at all.

Due to the large time intervals that we are interested in this thesis we can simplify equation (3.25) to

$$\delta T_b\left(\nu\right) = 27 x_{\mathrm{HI}} \left(\frac{1+z}{10}\right)^{1/2} \left(1 - \frac{T_\gamma}{T_S}\right) \mathrm{mK}. \tag{3.26}$$

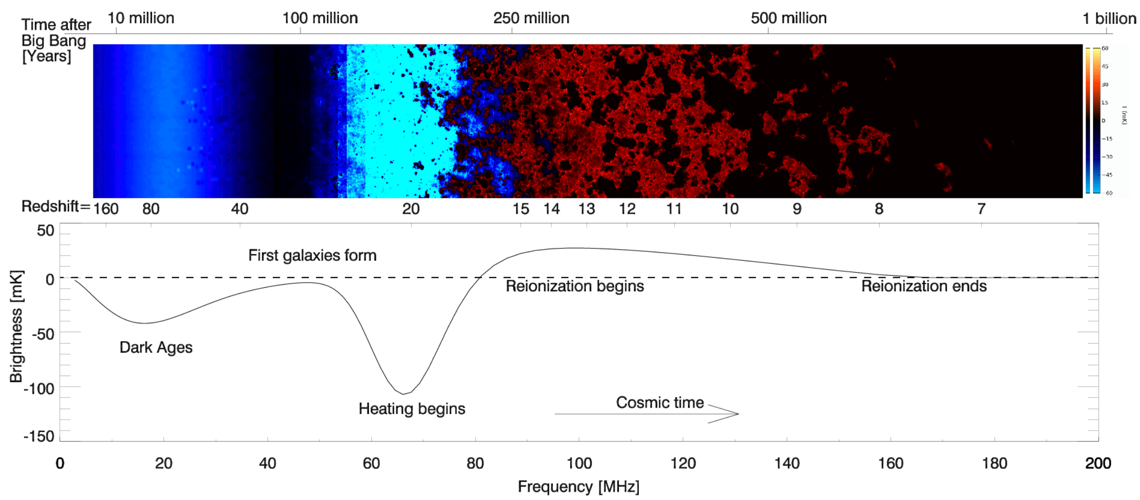


Figure 3.2: The 21-cm signal history through cosmic time. Top figure shows the time evolution of the fluctuations of the 21-cm signal from an epoch just before the first stars formed through the end of reionization epoch. Bottom figure shows the history of the expected brightness temperature of the 21-cm signal. Taken from (Pritchard and Loeb, 2012).

4 | Measuring the 21-cm signal

The effects of the expansion of the Universe make the original 21-cm signal redshift to wavelengths of the order of meters. To detect it we require instruments working in radio frequencies, and furthermore that they are located in radio-quiet zones. This is the main problem regarding the measurement of the Hydrogen transition signal.

There are plenty experiments around the world studying the 21 cm signal. There are two kinds of proposals: *i)* to measure the power spectrum of the 21 cm fluctuations and *ii)* to measure the all-sky absolute brightness temperature. All of these experiments are focused on the duration between the formation of the first protogalaxies during dark ages and the starts of the reionization epoch.

During this chapter, we describe these experiments and show their current results on the constraining of the model. Also, we detail the Mexican experiment SCI-HI and the design of its antenna HIbiscus.

4.1 Experiments for the detection of the signal

Experiments measuring the power spectrum fluctuations

- Precision Array to Probe the Epoch of Reionization (PAPER). This experiment is a low-frequency radio interferometer funded by the Nation Science Foundation with 128 antennas located in Karoo, South Africa, and 32 antennas in Green Bank, USA. The range of frequencies at which this experiment operates is around $\nu \approx 100 - 200$ MHz (Pober et al., 2013).

- Giant Meterwave Radio Telescope (GMRT) is an array of thirty steerable parabolic radio telescopes of 45 meters of diameter located near Pune, India. The antennas of this experiment are optimized to work in the frequency bands between 50 MHz and 1500 MHz (Paciga et al., 2011).

- LOw-Frequency Array (LOFAR). This experiment consists of many low-cost antennas located in thirty-six stations around the North-East of Netherlands and several international stations in Germany, Sweden, UK and France. LOFAR works with two kinds of antennas, the Low Band Antennas (LBA) operating between 10 MHz and 90 MHz, and the High Band Antenna (HBA) between 110 MHz and 250 MHz (van Haarlem et al., 2013).

- Murchison Widefield Array (MWA) is a low-frequency radio telescope array of 2048 dual-polarization dipole antennas located in the Murchison Radio-astronomy Observatory (MRO) in Western Australia. The telescope operates between 80 MHz and 300 MHz (Bernardi et al., 2013; Tingay et al., 2013).

- Hydrogen Epoch of Reionization Array (HERA) is a staged experiment located in Karoo, South Africa. HERA will consist in an array of 350 parabolic antennas installed in the South African Karoo Radio astronomy Reserve observing from 50 MHZ to 250 MHz. Its main purpose is to improve the efforts of its predecessors by bringing more sensitivity to the measurements (DeBoer et al., 2017).

- Square Kilometer Array (SKA) is an international project looking to build the world's largest radio telescope, with over a square kilometer of collecting area. SKA will eventually use thousands of dishes that will enable the exploration of the epoch of reionization, the location of SKA will be defined through the work of its precursors (LOFAR, MWA and HERA) and is expected to start its observations in 2020 (Ekers, 2012).

Experiments measuring the global 21 cm brightness temperature

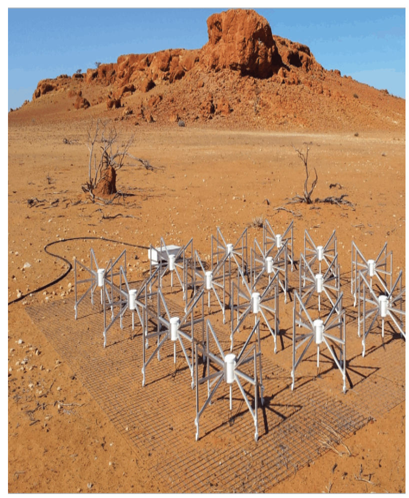
- Experiment to Detect the Global EoR Signature (EDGES) is a collaboration between Arizona State University and the MIT Haystack Observatory, it is located at the Murchison Radio-astronomy Observatory. It consists of two instruments a high-band instrument operating in 100 MHz to 200 MHz and a low-band instrument sensitive to 50 MHz up to 100 MHz (Bowman et al., 2008).

- Large-Aperture Experiment to detect Dark Ages (LEDA) is an array of antennas located in Socorro, New Mexico, it consists in 256 dipole antennas operating in the range of 10 MHz to 88 MHz (Greenhill and Bernardi, 2012).

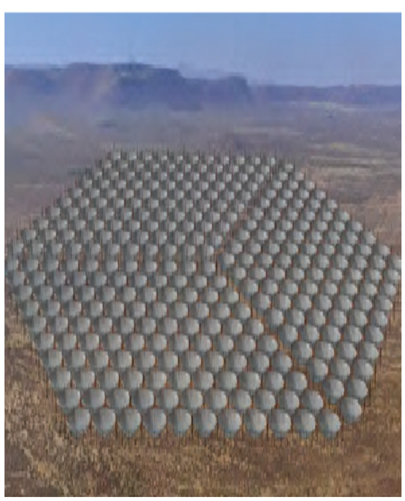
(a) PAPER

(b) GMRT

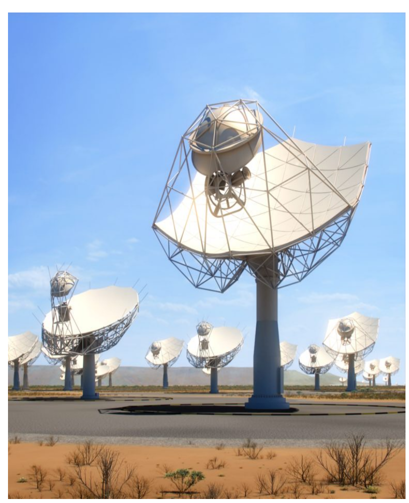
(c) LOFAR



(d) MWA



(e) HERA



(f) SKA

Figure 4.1: Mapping experiments of the 21-cm signal fluctuations power spectrum.

- Broadband Instrument for Global HydrOgen ReioNisation Signal (BIGHORNS) is an experiment located in Western Australia, it consists in a conical radiometer optimized for the frequencies between 20 MHz and 300 MHz (Sokolowski et al., 2015).
- Shaped Antenna measurement of the background RAdio Spectrum (SARAS) is a correlation spectrometer operating in the octave band 87.5-175 MHz. The antenna of SARAS is a frequency-independent Fat-dipole over an absorber ground plane (Patra et al., 2013).
- Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro (SCI-HI) is an experiment located in Isla Guadalupe, Mexico, with a modified four-square antenna optimized to work in 40-130 MHz frequencies (Voytek et al., 2014) more details are given in section 4.4.
- Dark Ages Radio Explorer (DARE). All the experiments described above are ground-based which implies that they are sensible to radio noises such as the FM radio signals, television, etc. To improve this disadvantage DARE has been proposed. DARE is a lunar orbiter mission with a wide-band bicone antenna that will observe at low-radio frequencies, 40-120 MHz. It is expected to launch in 2021 or 2022 (Burns et al., 2012).

4.2 Brightness temperature parameterizations

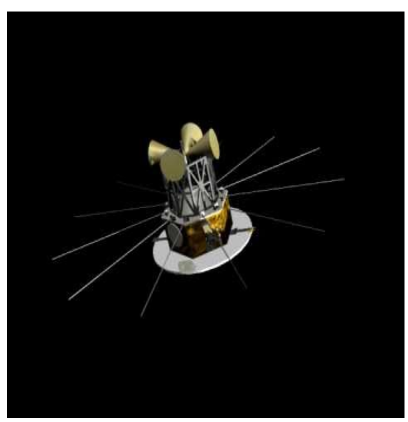
As mentioned in chapter 3 the physical model for the brightness temperature is given by equation (3.26), although comparing this model with observational data is quite difficult since solving the equations for the evolution of the intergalactic medium

(a) EDGES

(b) LEDA

(c) BIGHORNS

(d) SARAS



(e) DARE

Figure 4.2: All-sky absolute brightness measuring experiments.

(IGM) is needed, this requires knowing the possible sources of radiation that could affect signal.

Hence, to compare the measured data with a model it is convenient to approximate the signal via a simple parameterization of its value and possible errors caused by noises of other sources, like the sky temperature, the instrument proper noise, etc.

The experiments trying to measure this signal are looking towards a fixed point in the sky, most commonly the zenith, while the movement of the Earth keeps going, therefore, it is possible to model the temperature measured by the instrument with the simple equation

$$T_{\text{model}}(\nu) = T_{\text{sky}}(\nu) + \sigma(\nu), \tag{4.1}$$

where $\sigma(\nu)$ models the possible measurement errors in this case the standard deviation of the data collected is used.

Also, the sky temperature T_{sky} can be modeled by

$$T_{\text{sky}} = T_b + T_{\text{FG}}, \tag{4.2}$$

where T_b and T_{FG} are the brightness temperature and foreground (noise) temperature respectively.

Gaussian parameterization of δT_b

The simplest way to model the signal for the reionization epoch, due to its shape (see Fig. 3.2), is with a Gaussian approximation, such that

$$\delta T_b = -T_0 \exp\left(-\frac{(\nu_0 - \nu)^2}{2\sigma_0^2}\right), \tag{4.3}$$

where T_0 is a reference temperature, ν_0 is the frequency at which the signal has a minimum and σ_0 is the standard deviation of the Gaussian distribution. These parameters are quite important for the measurement of the signal since they determine the amplitude of the signal, the localization of the minimum of the signal absorption and the duration of the dark ages epoch before the beginning of the reionization epoch respectively.

tanh parameterization of δT_b

Another parameterization for the brightness temperature that has been implemented in recent works (Harker et al., 2015; Morandi and Barkana, 2012) is the tanh parameterization, this model follows the equation

$$\delta T_b = \frac{1}{2} T_0 \left(\frac{1+z}{10}\right)^{1/2} \left[1 + \tanh\left(\frac{z - z_r}{\Delta z}\right)\right], \tag{4.4}$$

where z_r is the redshift at which reionization occurs, and Δz is the duration of this process. This parameterization has the advantage that it contains information about the $Ly\alpha$ background, the temperature of the Intergalactic Medium and the ionized fraction of hydrogen (Harker et al., 2015).

Foreground models

Due to the possible measured signal has different source contribution, it is needed to extract the noise from the measurement. One of the principal sources of noise is our Galaxy, for this reason a model for the radiation contribution of the galaxy to the instrument measure is needed.

One of the models for the Galaxy temperature is the Global Sky Model (GSM), this is a model constructed with the compilation of the data sets of different large-area radio surveys (de Oliveira-Costa et al., 2008).

Another way is a simple parameterization of the temperature following a power law function, such that

$$T_{\mathrm{FG}} = T_0 \left(\frac{\nu}{\nu_0} \right)^{\alpha}, \tag{4.5}$$

where T_0 is the reference temperature of the Galaxy, ν_0 is the frequency of the Galaxy model and α is the specific coefficient to fit the GSM.

At last, we will mention another useful parameterization which is a logarithmic polynomial expansion given by

$$\log T_{\mathrm{FG}} = \sum_{n=0} a_n \log \left(\frac{\nu}{\nu_0} \right)^{n}, \tag{4.6}$$

where a_n are the expansion coefficients that will depend on the reference temperature.

With these parameterizations along with the brightness temperature model it is possible to extract the signal from a set of measurement data of an instrument.

4.3 Current results of the global brightness temperature experiments

The experiments mentioned in section 4.1 are trying to study and characterize the global 21-cm signal through their measurement, also their main goal is to constrain the model: the amplitude of the signal, the frequency of the minimum of the peak (maximum absorption) and the duration of the Dark Ages epoch.

Several reports of these measurements have occurred in recent years. In 2010, EDGES reported a constrain in the lower limit of the duration of the reionization epoch of $\Delta z > 0.06$ with 95% confidence level (Bowman and Rogers, 2010).

Later, in 2014, SCI-HI reported its firsts results for the global brightness temperature of the signal in the redshift range $14.8 < z < 22.7$ (Voytek et al., 2014). In 2016, LEDA presented its results where they constrain the signal amplitude and width to be $-890 < T_0 < 0$ mK and $\sigma_0 > 6.5$ MHz, respectively, corresponding to $\Delta z > 1.9$ at redshift $z \approx 20$ with a 95% confidence level in the range $100 > \nu > 50$ MHz (Bernardi et al., 2016). In mid 2017, EDGES reported new constrains to the model (Monsalve et al., 2017).

At last, at the beginning of 2018, EDGES reported the detection of the absorption profile centered at 78 MHz, a full width of 19 MHz and an amplitude of 500 mK with a 99% confidence level (Bowman et al., 2018).

The parameters reported by EDGES are consistent with the 21-cm signal model, however the fitting amplitude that they reported is higher than the predictions, suggesting that either the primordial gas was much colder than expected or the

CMB radiation temperature was hotter than expected (Bowman et al., 2018). Other explanations based on dark matter models have been also proposed (Berlin et al., 2018; Houston et al., 2018; Liu and Slatyer, 2018).

4.4 The SCI-HI experiment

The Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro (SCI-HI) experiment is dedicated to measure the 21 cm brightness temperature using a single broadband sub-wavelength size antenna and a sampling system for data processing and recording during observations (Jáuregui, 2016).

The antenna used by SCI-HI, HIbiscus (see Fig 4.3), was designed to work from 40-130 MHz, with a 90% antenna coupling efficiency from 55-90 MHz (Voytek et al., 2014).

Figure 4.3: Image of the HIbiscus antenna of the SCI-HI experiment in Isla Guadalupe, Mexico. Taken from Voytek et al. (2014).

SCI-HI began collecting data in June 2013 at Isla Guadalupe, Mexico (Latitude 28°58′24″N, Longitude 118°18′4″W), an island located 260 km off the Baja California peninsula in the Pacific Ocean, this location was selected after a study of the Radio Frequency Interference (RFI) noise contributions in areas located in the Mexican territory like Zona del Silencio, a zone located in the border of Coahuila, Chihuahua and Durango, and Isla Guadalupe itself (Jáuregui, 2016).

This study found that both sites are optimal for radio frequency observations since they present less RFI compared to observatories located in radio-quiet zones like Arecibo, Green Bank and Dominion Radio Astrophysical Observatory (DRAO) in Canada where the FM signal exceeds the sky signal by 10 dB over the band of 88-108 MHz (Voytek et al., 2014). Isla Guadalupe was found to be one of the best radio quiet zones in the world with about 0.1 dB residual FM noise above the Galactic foreground.

4.4.1 The HIbiscus antenna

The final design of the HIbiscus antenna was a modified four-square antenna scaled to 70 MHz, these modifications include the division of the square plates into inclined trapezoidal facets; also, additional panels were added to the side of each petal creating a strip line with fixed gap between the facets, allowing a better performance compared to a normal four-square antenna model and it has a 55° beam at 70 MHz (Jáuregui, 2016).

The signal that the antenna receives passes through a series of amplifiers and filters dedicated to remove the RFI noise contributions below 30 MHz and signals above 200 MHz. In order to measure the proper noise of the system, a switch was placed between the antenna and the first amplifier, this switch is activated in 5 minutes interval to collect the data from resistors of known temperature (50Ω, 100Ω and

Short), this data is used for calibration. The system generates a power spectra from 0-250 MHz. A basic system block diagram of the HIbiscus antenna instrumentation is shown in Figure 4.4.

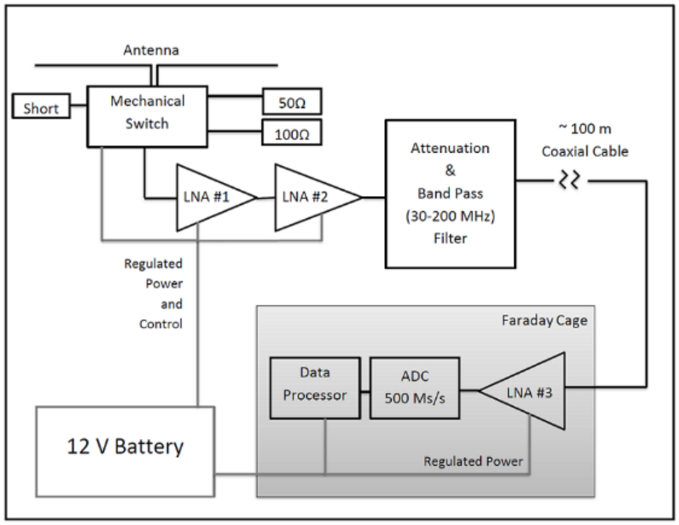


Figure 4.4: System block diagram of HIbiscus as shown in Voytek et al. (2014).

4.4.2 HIbiscus preliminary results

HIbiscus collected data in June 1-15, 2013, after this, the data was calibrated using different methods, after calibration, the results obtained by SCI-HI were presented, the diurnal variation of the temperature for the 70 MHz wavelength collected by HIbiscus can be seen in Figure 4.5, while the temperature in function of the frequency is presented in Figure 4.6.

The results reported by SCI-HI include residuals of less than 1 Kelvin, particularly near 70 MHz frequency and approximately 10 Kelvin using to different kinds of calibration respectively (Voytek et al., 2014).

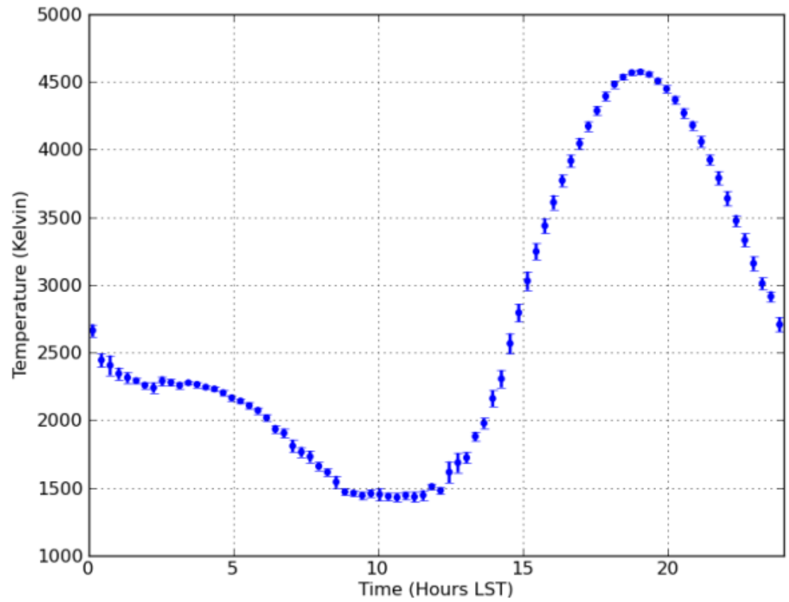


Figure 4.5: Diurnal variation of the temperature $T_{\text{meas}}(t, \nu)$ at 70 MHz frequency, the data shown is from 9 days of observation binned in 18 minutes intervals reported by Voytek et al. (2014).

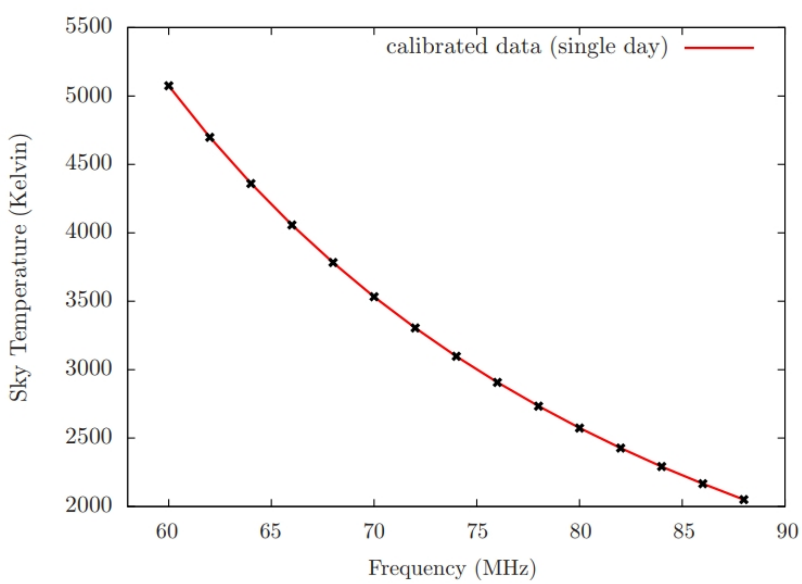


Figure 4.6: Data calibrated using the Global Sky Model, the plot shows the mean data from a single day of observation with approximately 50 minutes of integration, binned at 2 MHz intervals as seen in Voytek et al. (2014).

5 Simulations of the observed 21cm cosmological signal

The observations done by an experiment must be calibrated to interpret its results. For this are different methods that should be described in chapter 6, but before calibrating, we must have a theoretical value with which we will compare the measurements. To achieve this task, we shall simulate the observed sky by an antenna given its beam pattern and location in Earth and then obtain an expected temperature, this procedure is described next.

5.1 Antenna beam pattern and trajectory

Usually the antenna beam pattern is given in a file with its spherical coordinates. The polar angle $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, the azimuthal angle $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, notice that the θ and φ angles cover the top half of a sphere; at last the measurement of a reference temperature was given in decibel-milliwatts (dBm). First, we transform these coordinates given into Cartesian coordinates with the relations

$$
\begin{aligned}
x &= \sin\theta\cos\varphi, \\
y &= \sin\theta\sin\varphi, \\
z &= \cos\theta,
\end{aligned}
\tag{5.1}
$$

using these coordinates we get the altitude a and the azimuthal angle A coordinates of the antenna beam pattern in the horizontal coordinate system with

$$
\begin{aligned}
A &= \arctan\left(\frac{y}{x}\right), \\
a &= \frac{\pi}{2} - \arctan\left(\frac{x^2 + y^2}{z}\right).
\end{aligned}
\tag{5.2}
$$

Using the latitude ϕ of an observer in a given point of the Earth, we use the transformations to equatorial coordinates (Karttunen et al., 2016), these coordinates are given by the relations

$$
\begin{aligned}
\sin H\cos\delta &= \sin A\cos a, \\
\cos H\cos\delta &= \sin\delta\cos\phi + \cos A\cos a\sin\phi, \\
\sin\delta &= \sin a\sin\phi - \cos A\cos a\cos\phi.
\end{aligned}
\tag{5.3}
$$

where δ is the declination of an observed object and H is the hour angle. Thus,

$$
\begin{aligned}
H &= \arctan\left(\frac{sinA\cos a}{\sin\delta\cos\phi + \cos A\cos a\sin\phi}\right), \\
\delta &= \arcsin(\sin a\sin\phi - \cos A\cos a\cos\phi), \\
\alpha &= \Omega_{\mathrm{LST}} - H,
\end{aligned}
\tag{5.4}
$$

with α being the right ascension and Ω_{LST} is the local sidereal time.

With the H, α and δ coordinates, we can finally obtain the coordinates of the antenna beam pattern in the Galactic coordinate system following the equations

$$\begin{aligned} \sin(l_N - l)\cos b &= \cos\delta\sin(\alpha - \alpha_P), \\ \cos(l_N - l)\cos b &= \sin\delta\sin\delta_P - \cos\delta\sin\delta_P\cos(\alpha - \alpha_P), \\ \sin b &= \sin\delta\sin\delta_P + \cos\delta\cos\delta_P\cos(\alpha - \alpha_P), \end{aligned} \tag{5.5}$$

where the right ascension of the Galactic North Pole is $\alpha_P = 12$ h 51.4 min, its declination $\delta_P = 27°08'$ and the galactic longitude of the celestial pole $l_N = 123.0°$, there by

$$\begin{aligned} l &= l_N - \arctan\left(\frac{\cos\delta\sin(\alpha - \alpha_P)}{\sin\delta\sin\delta_P - \cos\delta\sin\delta_P\cos(\alpha - \alpha_P)}\right), \\ b &= \arcsin\big(\sin\delta\sin\delta_P + \cos\delta\cos\delta_P\cos(\alpha - \alpha_P)\big). \end{aligned} \tag{5.6}$$

The equations derived were used to simulate the antenna trajectory and simulated beam pattern; for the trajectory it is important to note that HIbiscus is pointing directly to the Zenith so the altitude of the observing point of HIbiscus is $a = 90°$, implying that, from equation (5.4), the hour angle H is zero and thereby the right ascension α is equal to the local sidereal time also the declination δ is equal to the observer's latitude ϕ, so for the antenna trajectory we have

$$\begin{aligned} l &= l_N - \arctan\left(\frac{\cos\delta\sin(\Omega_{\mathrm{LST}} - \alpha_P)}{\sin\phi\sin\delta_P - \cos\phi\sin\delta_P\cos(\Omega_{\mathrm{LST}} - \alpha_P)}\right), \\ b &= \arcsin\big(\sin\phi\sin\delta_P + \cos\phi\cos\delta_P\cos(\Omega_{\mathrm{LST}} - \alpha_P)\big), \end{aligned} \tag{5.7}$$

note that this equation entirely depends on the local sidereal time since the observer's latitude is fixed at Isla de Guadalupe $\phi = 28.9733$, in Figure 5.1 the trajectory of the antenna through the sky is shown.

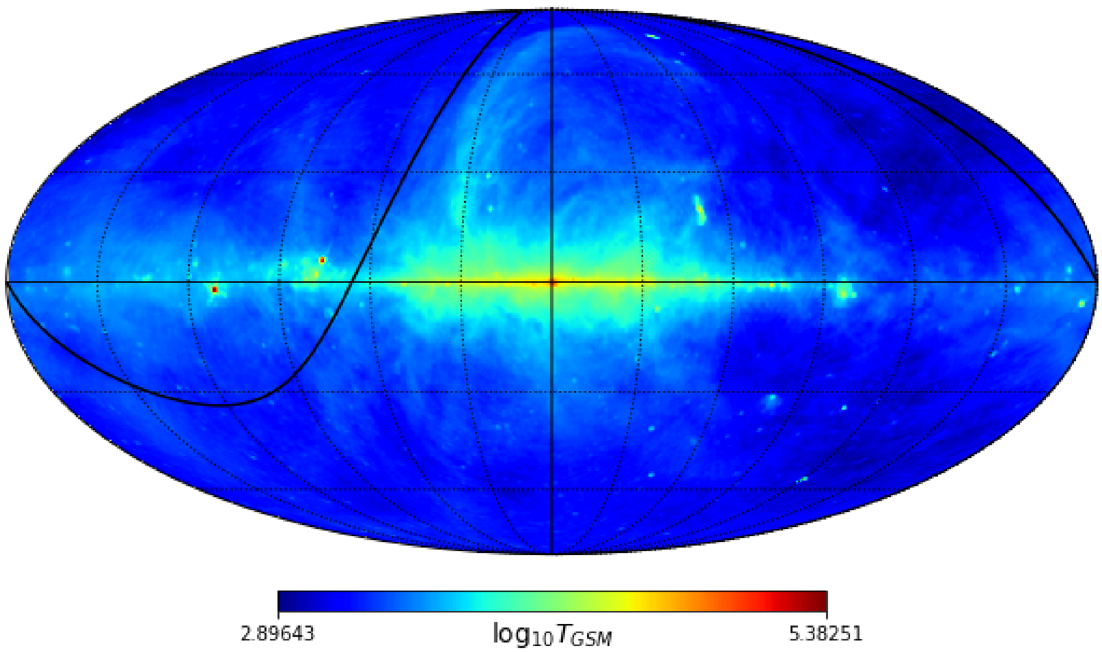


Figure 5.1: Trajectory of the HIbiscus antenna (black solid line), located at Isla de Guadalupe $\phi = 28.9733$, through the sky in a full day of observation, colors indicate the $\log_{10}$ temperature of the Galaxy in the 70 MHz wavelength using the GSM.

5.2 Simulation of the measured sky temperature

In order to obtain a simulated measured temperature $T_{\mathrm{GSM}}(t,\nu)$ for the antenna, it is needed to convolve the beam pattern temperature with the GSM temperature. For this, a conversion between the power measured by the antenna and its correspondent temperature is needed, this conversion is done by the equation (Jáuregui, 2016)

$$T = 1.36\frac{\lambda^2}{\Theta^2}S, \tag{5.8}$$

where λ is the wavelength of the measurement in centimeters, Θ is the solid angle, in arcsecs, subtended by the antenna in this case 55° and S is the flux density in miliJanskys (mJy).

The flux density is calculated by

$$S = \frac{2P}{A}, \tag{5.9}$$

where P is the measured power in Watt units, and A is the antenna area (1 m^2).

Since the data collected by HIbiscus is stored in dBm units a conversion into Watt units is needed, this is done by the following relation

$$P_{\mathrm{Watts}} = 10^{(P_{\mathrm{dBm}}-30)/10}\ \mathrm{Watts}. \tag{5.10}$$

Once we have the corresponding temperature T for the antenna beam pattern it is convolved with the GSM using the equation

$$T_{\mathrm{GSM}}(t,\nu) = \frac{\int \mathrm{GSM}(\mathrm{l},\mathrm{b},\nu)\mathcal{B}(\mathrm{l}-\mathrm{l_0(t)},\mathrm{b}-\mathrm{b_0(t)},\nu)\,\mathrm{d}\Omega}{\int \mathcal{B}(l-l_0(t),b-b_0(t),\nu)\,\mathrm{d}\Omega}, \tag{5.11}$$

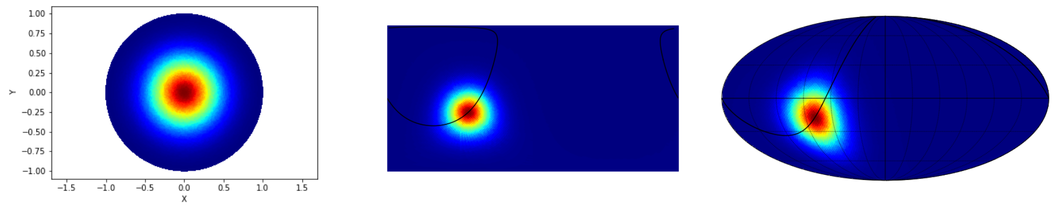
where (l,b) are the galactic coordinates, $\mathcal{B}(l,b,\nu)$ is the antenna beam pattern, $(l_0(t),b_0(t))$ is the beam center, note that these coordinates depend on time, this is due the rotation of earth, as seen on the previous section, and $\mathrm{GSM}(\mathrm{l},\mathrm{b},\nu)$ is the Global Sky Model temperature.

This method can be applied in any antenna using its beam pattern, as an example the methodology will be applied in a mock Gaussian antenna beam pattern and the HIbiscus antenna.

5.3 Applying the method on a mock Gaussian beam pattern antenna

For test purposes, a Gaussian mock pattern for an antenna located in Isla Guadalupe was generated, this is so the trajectory of the antenna beam center through the sky would be the same as the HIbiscus antenna, this pattern was used to test whether the pattern simulation was consistent with its projection in the night sky, the results are shown in Figure 5.2.

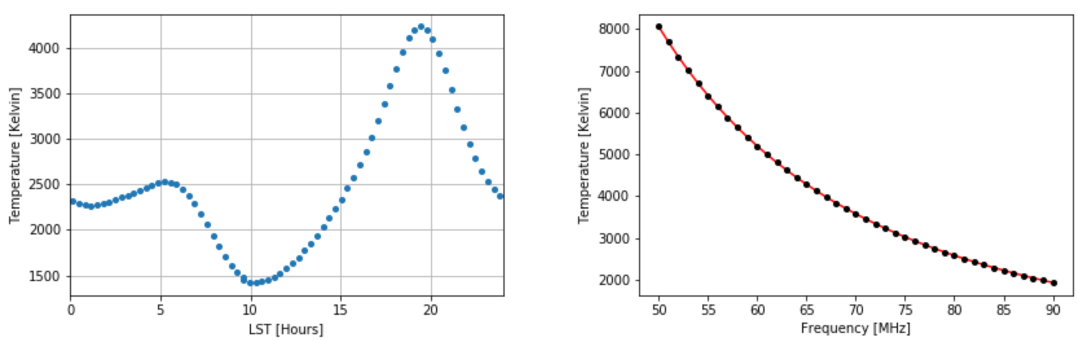
Since the pattern is Gaussian, the projected pattern into galactic coordinates is expected to be Gaussian too, this is more clearly seen in the Cartesian visualization, this can be seen in the center of Figure 5.2, proving that the pattern projection to galactic coordinates and therefore the convolution with the Global Sky Model was done correctly.



(a) (b) (c)

Figure 5.2: Mock antenna beam pattern generated to test the methodology of the pattern analysis and convolution. Figure (a) shows the pattern projection along the XY plane; figure (b) shows the projected pattern in galactic coordinates using the Cartesian visualization; figure (c) shows the projected pattern using the mollweide visualization. This pattern is for the date June 14, 2013 at 08:00:00 UTC. The black solid line in the center and right figures is the antenna trajectory.

Next, the Gaussian pattern was convolved with the GSM to get a mock theoretical temperature T_{GSM} for the antenna, once this temperature was calculated, it is possible to check the diurnal variation of the temperature during a day of observation along with the temperature in function of the frequency, this is shown in Fig. 5.3.



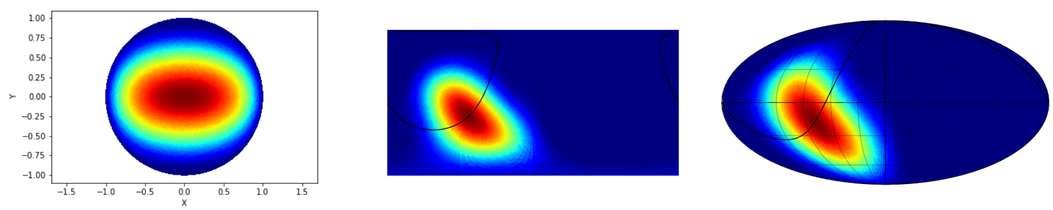
(a) (b)

Figure 5.3: Results for the mock Gaussian beam pattern antenna. Figure (a): Simulated diurnal variation for 70 MHz. Figure (b): Simulated temperature in function of the frequency for June 14, 2013 at 08:00:00 UTC.

5.4 Applying the method on the HIbiscus antenna

Once the methodology was tested to be correct, the procedure used in section 5.3 was repeated using the antenna beam pattern information of HIbiscus, this data was provided by the SCI-HI team for analysis.

The beam pattern of the HIbiscus antenna is shown in Fig. 5.4, also the beam pattern trajectory of HIbiscus for a full day in 4 hours intervals can be seen Fig. 5.5, notice how the beam center follows the predicted trajectory.

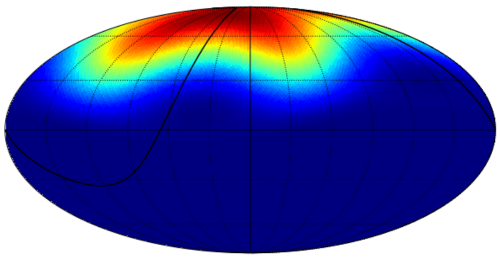


(a) (b) (c)

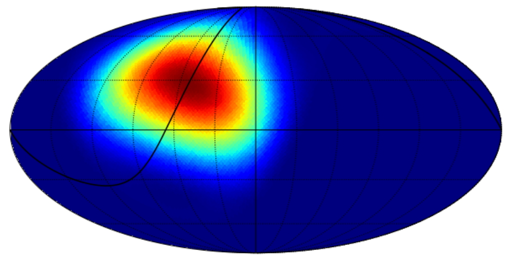
Figure 5.4: HIbiscus antenna beam pattern. As in Figure 5.2, figure (a) shows the pattern projection in the XY plane; figure (b) shows the projected pattern in galactic coordinates using the Cartesian visualization; figure (c) shows the projected pattern using the mollweide visualization. This pattern is for the date June 14, 2013 at 08:00:00 UTC. The black solid line in the center and right figures is the antenna trajectory.

After obtaining the convolved temperature $T_{\rm GSM}$, the diurnal temperature variation of the HIbiscus antenna (see Fig. 5.6) and temperature in function of the frequency (see Fig. 5.7) were simulated, in order to compare with the results reported by SCI-HI, we have added the results for the mock Gaussian beam pattern for comparison.

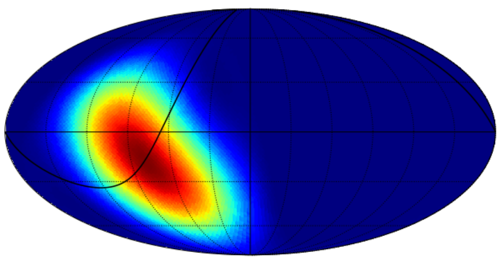
Note that the shape from both simulated diurnal variation and temperature graphs coincide with the reported by SCI-HI (Figures 4.5 and Figure 4.4.2), nevertheless the data and procedure to calculate these results are not public meaning that a direct



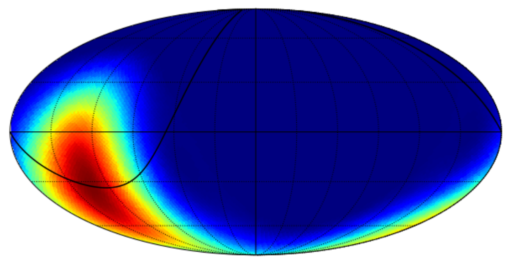
(a) UTC 04:00



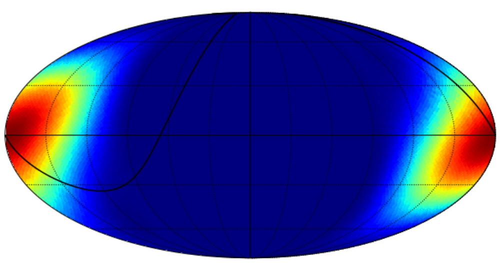
(b) UTC 08:00



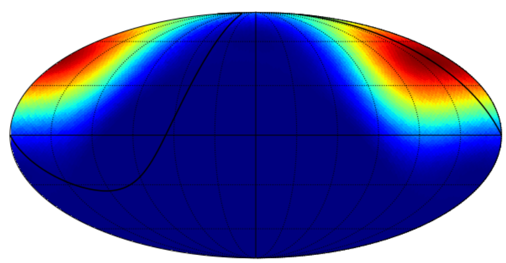
(c) UTC 12:00



(d) UTC 16:00



(e) UTC 20:00



(f) UTC 24:00

Figure 5.5: HIbiscus beam pattern, in galactic coordinates, at 70 MHz for different UTC, notice how the center of the beam follows the calculated trajectory (black solid line).

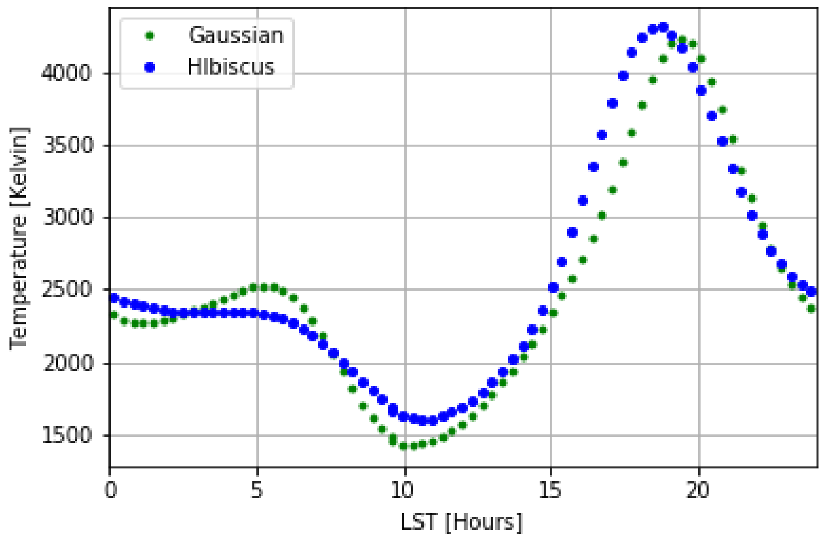


Figure 5.6: The blue dots are the simulated diurnal variation of the HIbiscus antenna for 70 MHz frequency. We have added the results for a mock Gaussian pattern (green dots) for comparison.

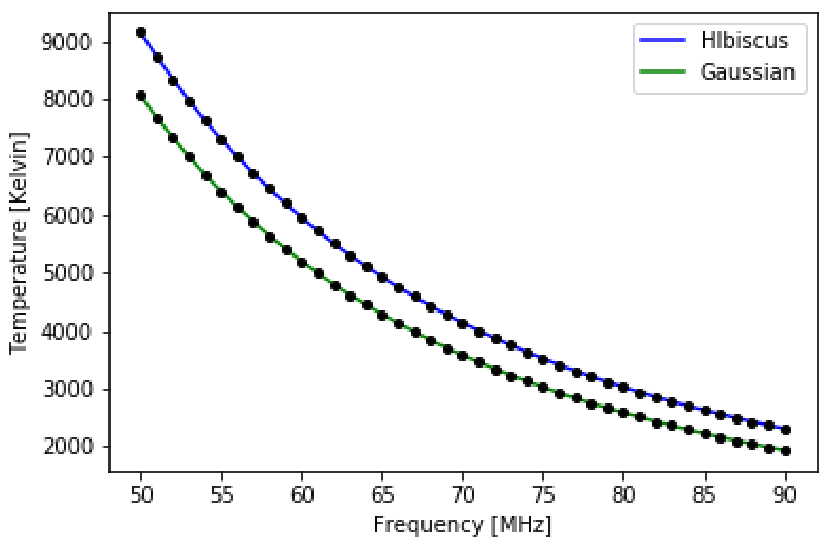


Figure 5.7: Blue line represents the simulated temperature in function of the frequency of the HIbiscus antenna for June 14, 2013 at 08:00:00 UTC. We have added the results for the mock Gaussian pattern (green line) for comparison.

comparison can not be done.

This method of simulating the measurements of an antenna given its beam pattern, can be applied in any circumstance just by having the location of the antenna and a data file given in spherical coordinates of the antenna beam pattern. Also, a possible use of this methodology is to introduce some random noise into the temperature simulation to create a mock data useful for testing the calibration methodologies of an experiment.

6 Data Calibration

As mentioned in subsection 4.4.1, HIbiscus stores its raw data in four different kinds of measured data in dBms, these data correspond to the direct antenna (P_{ant}), noise (P_{noise}), 50Ω resistor ($P_{50\Omega}$) and short (P_{short}), these data is used to calibrate the measurements obtained by HIbiscus.

This chapter details the calibration methods used for the purpose of this thesis. Also, we show the procedure of data filtering and selection and at the end we discuss the results obtained from implementing these methods.

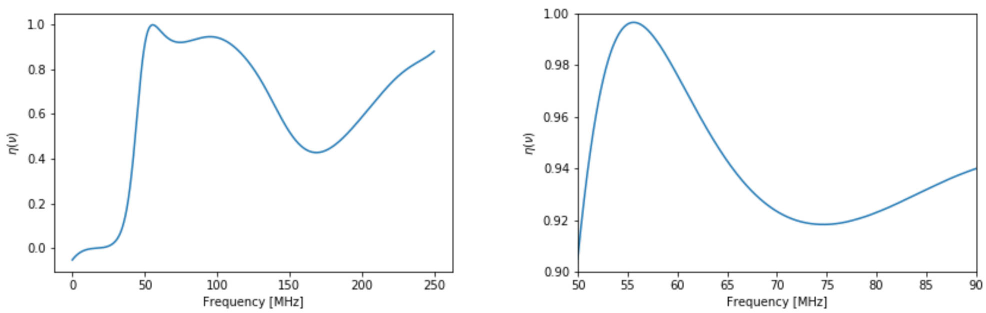
6.1 Calibration methods

First, the measured power spectrum obtained by the antenna is cleaned by removing the noise data from the antenna data $P_{\text{meas}} = P_{\text{ant}} - P_{\text{noise}}$, next the measured temperature is calculated by

$$T_{\text{meas}}(t,\nu) = K(\nu)\left[\frac{P_{\text{meas}}(t,\nu)}{\eta(\nu)} - P_{\text{short}}(\nu)\right], \tag{6.1}$$

where $K(\nu)$ is the system gain, and $\eta(\nu)$ is the transmission efficiency from the antenna measured.

The efficiency of HIbiscus was measured on site after the data collection in 2013, this efficiency for the whole power spectra generated by HIbiscus (0-250 MHz) is shown in Figure 6.1, note from Figure 6.1(b) that for frequencies of interest in the 21-cm signal (50-90 MHz) the efficiency is above 90%, as Voytek et al. (2014) mentioned.



(a) (b)

Figure 6.1: Transmission efficiency from HIbiscus. Figure (a) shows the whole efficiency of the antenna on its entire bandwidth (0-250 MHz), figure (b) shows the frequencies of interest for the 21-cm signal (50-90 MHz).

After calculating T_{meas} the system gain is calculated using either the Johnson-noise calibration or the χ^2 method, described below.

Johnson-noise Calibration (JNC)

This method uses the data from P_{short}, $P_{50\Omega}$ and the ambient temperature T_{amb} at the moment of the measurement of the data, the system gain is calculated by the equation

$$K_{\text{JNC}}(\nu) = \frac{T_{\text{amb}}}{P_{50\Omega}(\nu) - P_{\text{short}}(\nu)}. \tag{6.2}$$

χ^2 fit

This method consist in getting the gain $K_{\text{GSM}}(\nu)$ that best fits the theoretical model, this is done by minimizing the following equation

$$\chi^2(\nu) = \sum_t \frac{[T_{\text{meas}}(t,\nu) - T_{\text{GSM}}(t,\nu)]^2}{\sigma^2_{\text{meas}}(t,\nu)}, \tag{6.3}$$

where $\sigma^2_{\text{meas}}(t,\nu)$ is the standard deviation of the data for a given frequency and time.

6.2 Data filtering

The data collected from HIbiscus on June 14, 2013, and June 15, 2013, were provided for analysis. The first step to use this data is to get the measured temperature using both method described above. For the JNC method, we ignore the measurements that exceed the limits of temperature set in a maximum temperature for the JNC method of 100,000 Kelvin and a minimum temperature of 100 Kelvin, for an easier analysis this data was examined in $\log_{10}$ scale, meaning that data above $\log_{10} T_{\text{meas}} = 5$ and below $\log_{10} T_{\text{meas}} = 2$ were discarded.

The collected data before χ^2 calibration and using JNC calibration can be seen in Fig 6.2 and Fig 6.3 respectively.

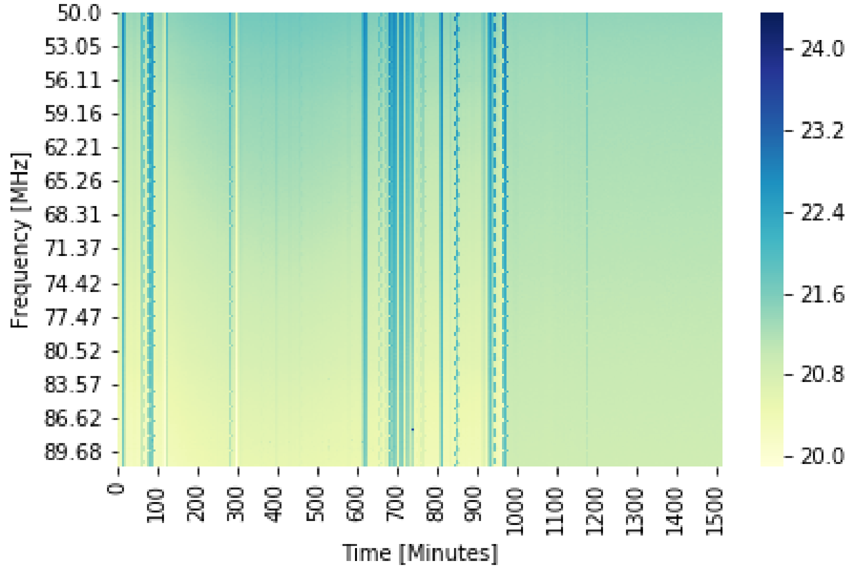


Figure 6.2: Data collected by HIbiscus in June 14, 2013 and June 15, 2013, the data was binned in 5 minutes intervals for visualization.

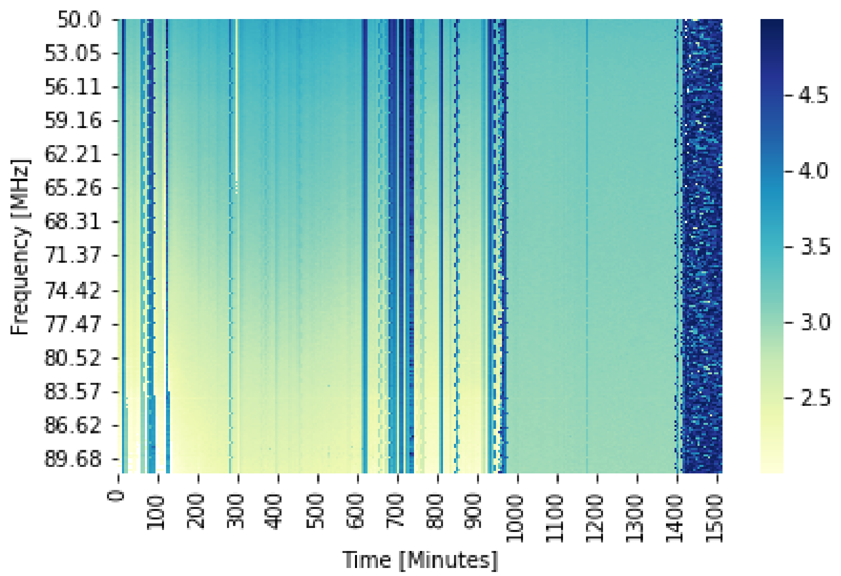


Figure 6.3: Data collected by HIbiscus in June 14, 2013 and June 15, 2013, the values were calculated using the JNC calibration, data above $\log_{10} T_{\rm meas} = 5$ and below $\log_{10} T_{\rm meas} = 2$ were discarded, also 5 minutes interval binning was done.

6.3 Analyzed data

As it can be seen in Fig 6.2 and Fig 6.3 the data collected is not uniform meaning that we can't analyze the whole data since it has some values above the limits, this is probably due to external noise to the system.

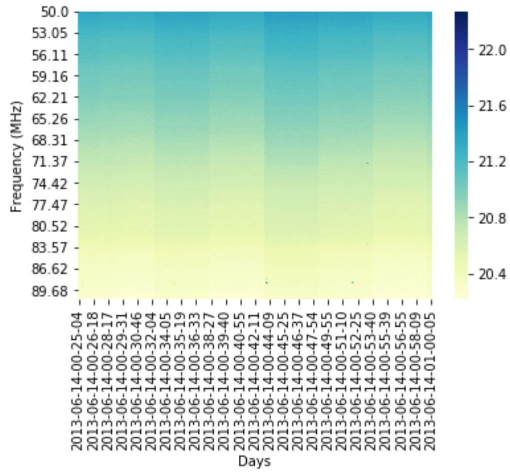
Even though not all the data is usable, we can still analyze some intervals of the observations, for this case four different time intervals were used for calibration. In this thesis we will focus on the χ^2 calibration, a further detail of the use JNC calibration method is given in Burgueño (2017).

All the time intervals chosen for analysis were from June 14, 2013, since the data from June 15, 2013, were not optimal, the time intervals chosen were 00:25:04 to 01:00:06, 04:40:07 to 05:50:01, 10:24:50 to 10:59:52 and 18:45:12 to 19:15:00. The analysis and calibration of these data intervals are presented in Figures 6.4, 6.5, 6.6 and 6.7, respectively, figures 6.4(a), 6.5(a), 6.6(a), and 6.7(a) show the data collected in the respective time interval; figures 6.4(b), 6.5(b), 6.6(b), and 6.7(b) show the data after χ^2 calibration, this calibration was done by binning the data in 5 minutes intervals and following the procedure described in Equation 6.1, figures 6.4(c), 6.5(c), 6.6(c), and 6.7(c) show the gain $K(\nu)$ used to calibrate the data; and figures 6.4(d), 6.5(d), 6.6(d), and 6.7(d) show the residue of the calibrated data with respect of the theoretical value ($T_{\text{meas}} - T_{\text{GSM}}$).

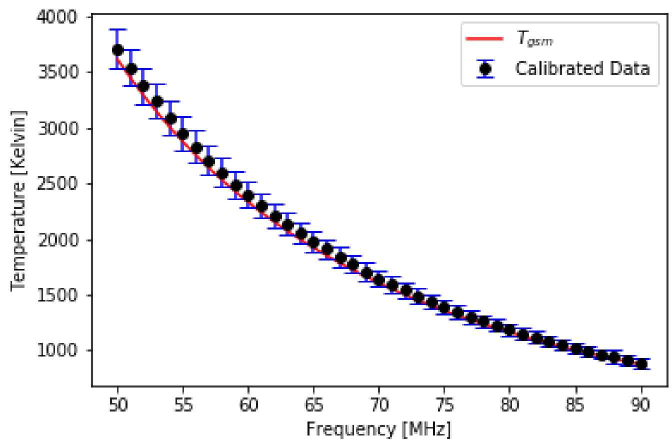
As it can be seen in figures 6.4(a), 6.5(a), 6.6(a), and 6.7(a), all the data analyzed has enough quality for the calibration, meaning that there are no noticeable noise contributions within the time intervals.

Also, from Figures 6.4(b), 6.5(b), 6.6(b) and 6.7(b) we see that the χ^2 method is a good calibration method since the results after calibration gets close to the theoretical value of the temperature T_{GSM} with a residue bellow 10^2 in most cases (see figures 6.4(d), 6.5(d), and 6.6(d)) except for Figure 6.7(b) were the maximum residue goes to $10^{2.05}$ (see Figure 6.7(d)). On figures 6.4(c), 6.5(c), 6.6(c), and 6.7(c) we see that the system gain $K(\nu)$ does not vary drastically within the frequencies of interest meaning that this calibration can be trusted.

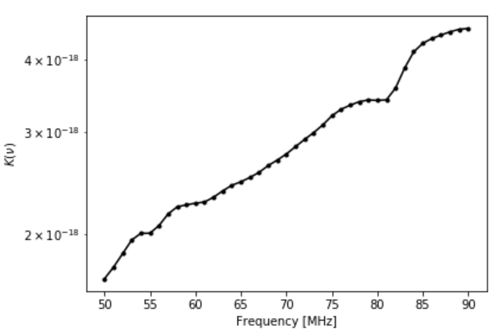
It must be noted that there must be precautions with the use of the χ^2 fit method since the procedure can attenuate the 21 cm signal. The main advantage of this method is that it does not depend on the way the data was reduced and filtered, it will always obtain a value of $K(\nu)$ to achieve the best fit to the theoretical value, unlike the JNC method that entirely depends on the data itself, this can generate greater errors than the presented above.



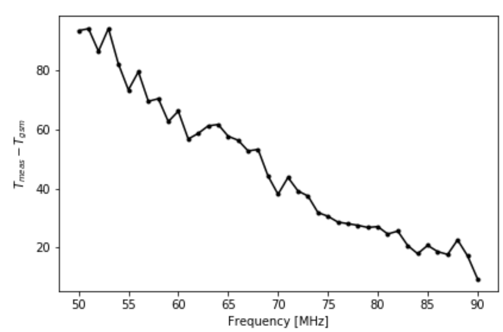
(a)



(b)

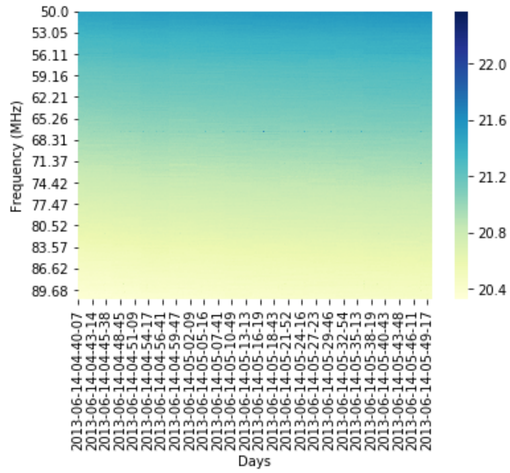


(c)

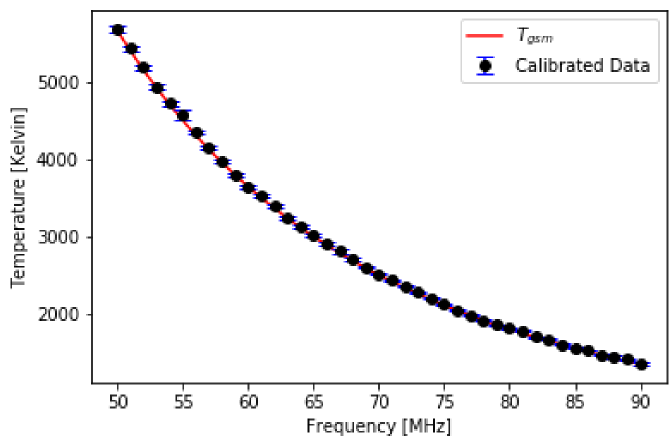


(d)

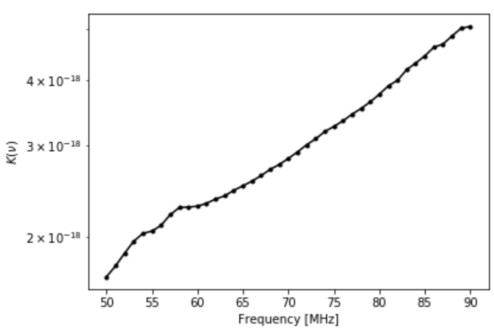
Figure 6.4: Results for June 14, 2013 during the time interval between 00:25:04 and 01:00:06. The value of $K(\nu)$ remains between the maximum value 4.523×10^{-18} and the minimum value 1.673×10^{-18}, the maximum value of the residue is 94.28 K, meaning a maximum 3.09% fractional error with respect the theoretical value of $T_{\rm GSM}$.



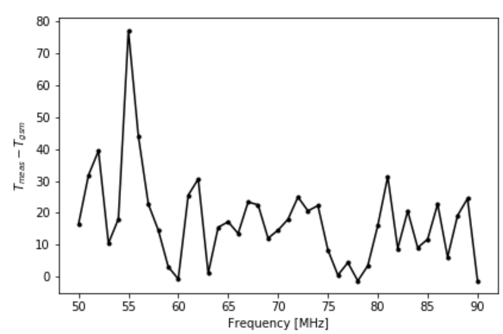
(a)



(b)

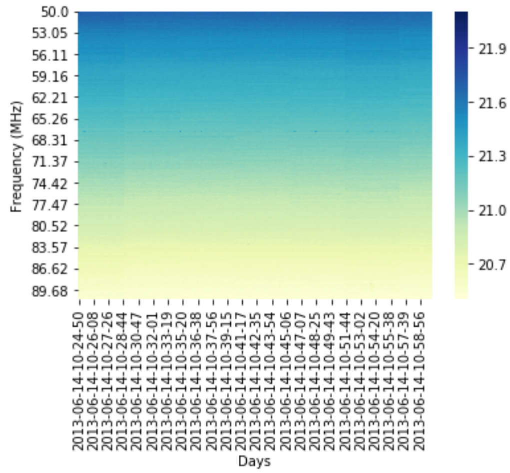


(c)

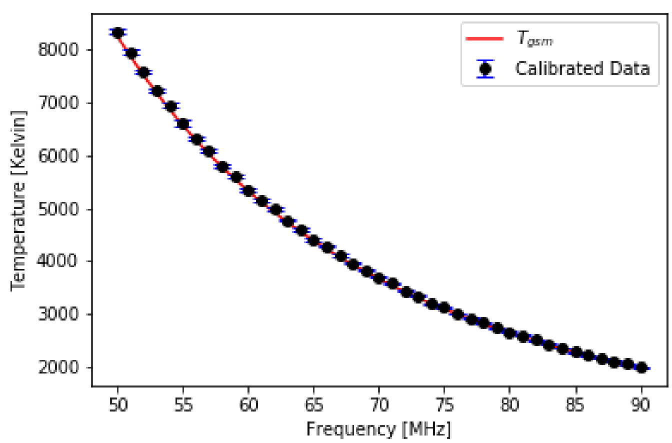


(d)

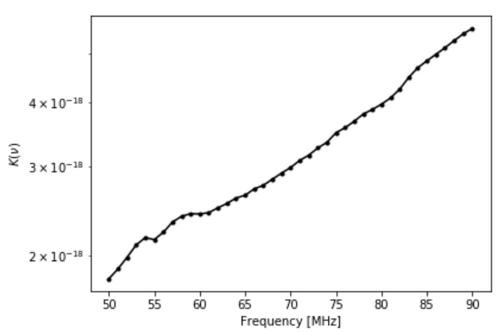
Figure 6.5: Results for June 14, 2013 during the time interval between 04:40:07 and 05:50:01. The value of $K(\nu)$ remains between the maximum value 5.067×10^{-18} and the minimum value 1.676×10^{-18}, the maximum value of the residue is 77.13 K, meaning a maximum 1.79% fractional error with respect the theoretical value of T_{GSM}.



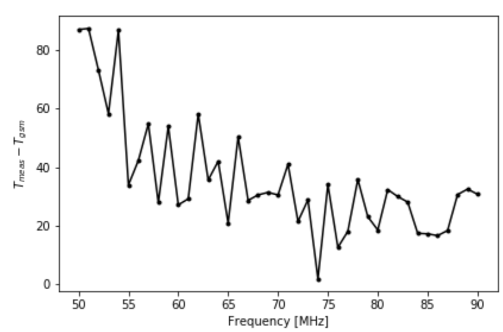
(a)



(b)

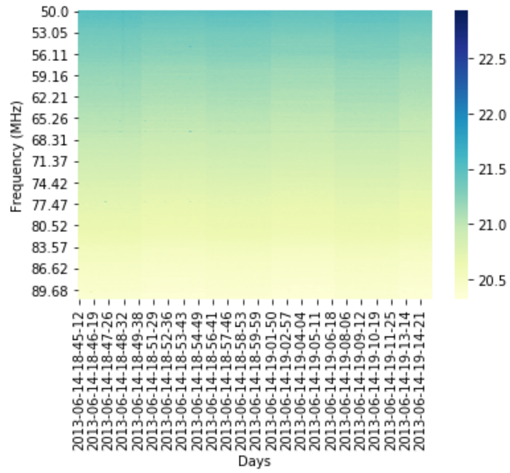


(c)

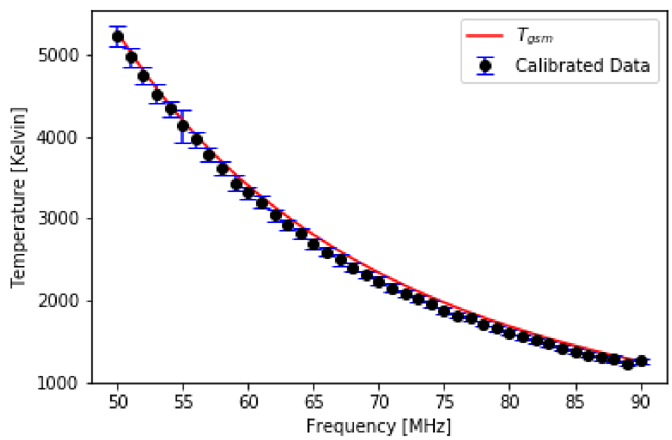


(d)

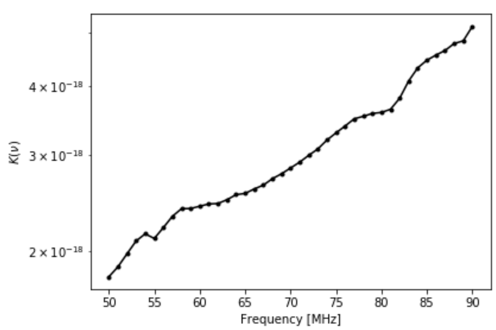
Figure 6.6: Results for June 14, 2013 during the time interval between 10:24:50 and 10:59:52. The value of $K(\nu)$ remains between the maximum value 5.610×10^{-18} and the minimum value 1.797×10^{-18}, the maximum value of the residue is 87.296 K, meaning a maximum 1.61% fractional error with respect the theoretical value of T_{GSM}.



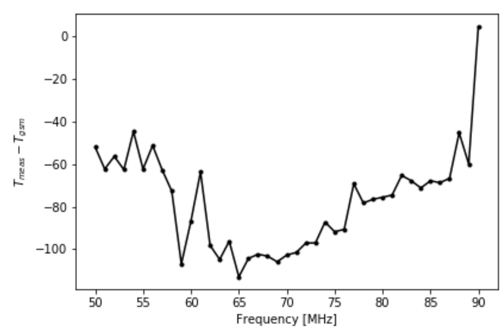
(a)



(b)



(c)



(d)

Figure 6.7: Results for June 14, 2013 during the time interval between 18:45:12 and 19:15:05. The value of $K(\nu)$ remains between the maximum value 5.134×10^{-18} and the minimum value 1.797×10^{-18}, the maximum value of the residue is 113.05 K, meaning a maximum 4.898% fractional error with respect the theoretical value of T_{GSM}.

7 | Conclusions and future work

It has been shown that the duty of detecting the 21 cm signal is not an easy job, but it is important for our understanding of the evolution of our Universe. The brightness differential temperature described in this thesis can give us a really good hint to unveil the mysteries of the cosmic structure formation and the duration of the Universe's epochs.

The study of this signal has attracted the attention of many researchers, leading to the implementation of different experiments and data analysis techniques, the most recent effort of understanding the signal was done by EDGES, reporting interesting parameters of the model of the 21 cm Hydrogen emission line, but these results need to be discussed and confirmed.

As mentioned by EDGES in their latest paper (Bowman et al., 2018), SCI-HI is one of the experiments close to achieve the performance to verify their results, so it is important that the methodology of analyzing the data measured by SCI-HI is efficient enough to possibly get a profile that proves, or maybe refutes, the profile obtained by EDGES.

As we have seen through this thesis, the methodology requires efficient data; the main problem with the data used to the calibrations for this work was the lack of quality of the whole set, the analysis had to be done within time intervals were it was believed that the data was good enough for usage. Using these data the results obtained were promising since the residues using the χ^2 calibration method were below 10^2, however for a good observation of the 21 cm signal these residues need to be reduced to 10^{-2} or better. A good improvement for future observations is to check this data real-time to make sure its quality is enough for analysis.

However, the simulation procedure, used to calculate the expected measures of an experiment, is believed to be efficient enough to get promising results once a good set of data is taken, since it only depends on the geographical location of the antenna and its pattern.

As mentioned many times, a better set of data from SCI-HI is needed, this is expected to happen during the next year, with the help of a new antenna design called "Mango Peel" (see Fig. 7.1), this antenna was developed to improve and fulfill the performance requirements that HIbiscus did not achieve. This antenna has an improved antenna pattern, uniform for the full observing range, and having fewer deviations than HIbiscus on gain (Jáuregui, 2016); also Mango Peel has a larger bandwidth than the old SCI-HI experiment.

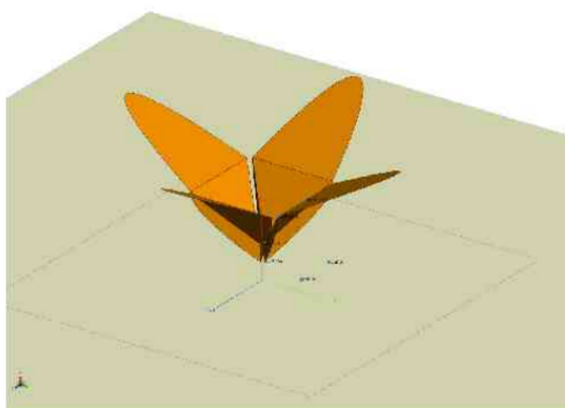


Figure 7.1: Mango Peel design. Taken from Jáuregui (2016).

Mango Peel will be ready to take measurements once it is built and installed on the observation site, possibly Isla Guadalupe or Isla Socorro, so the hope to confirm the profile obtained by EDGES is on this antenna.

Once Mango Peel starts collecting data, the methodology used for this thesis along with the Monte Carlo Markov Chain analysis done in Burgueño (2017) will be used to calibrate and compare results, although the JNC method needs to be tested and revised for a better calibration result, expecting for promising results.

Bibliography

Ade, P. A., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A., Barreiro, R., Bartlett, J., Bartolo, N., et al. (2016). Planck 2015 results-xiii. cosmological parameters. *Astronomy & Astrophysics*, 594:A13.

Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A., Barreiro, R., Bartolo, N., Basak, S., et al. (2018). Planck 2018 results. vi. cosmological parameters. *arXiv preprint arXiv:1807.06209*.

Alpher, R. A. and Herman, R. C. (1949). Remarks on the evolution of the expanding universe. *Physical Review*, 75(7):1089.

Bennett, C., Larson, D., Weiland, J., Jarosik, N., Hinshaw, G., Odegard, N., Smith, K., Hill, R., Gold, B., Halpern, M., et al. (2013). Nine-year wilkinson microwave anisotropy probe (wmap) observations: final maps and results. *The Astrophysical Journal Supplement Series*, 208(2):20.

Berlin, A., Hooper, D., Krnjaic, G., and McDermott, S. D. (2018). Severely constraining dark matter interpretations of the 21-cm anomaly. *arXiv preprint arXiv:1803.02804*.

Bernardi, G., Greenhill, L., Mitchell, D. A., Ord, S., Hazelton, B., Gaensler, B., de Oliveira-Costa, A., Morales, M., Shankar, N. U., Subrahmanyan, R., et al. (2013). A 189 mhz, 2400 deg2 polarization survey with the murchison widefield array 32-element prototype. *The Astrophysical Journal*, 771(2):105.

Bernardi, G., Zwart, J., Price, D., Greenhill, L., Mesinger, A., Dowell, J., Eftekhari, T., Ellingson, S., Kocz, J., and Schinzel, F. (2016). Bayesian constraints on the global 21-cm signal from the cosmic dawn. *Monthly Notices of the Royal Astronomical Society*, 461(3):2847–2855.

Bowman, J. D. and Rogers, A. E. (2010). A lower limit of $\delta z > 0.06$ for the duration of the reionization epoch. *Nature*, 468(7325):796.

Bowman, J. D., Rogers, A. E., and Hewitt, J. N. (2008). Toward empirical constraints on the global redshifted 21 cm brightness temperature during the epoch of reionization. *The Astrophysical Journal*, 676(1):1.

Bowman, J. D., Rogers, A. E., Monsalve, R. A., Mozdzen, T. J., and Mahesh, N. (2018). An absorption profile centred at 78 megahertz in the sky-averaged spectrum. *Nature*, 555(7694):67.

Burgueño, O. S. (2017). La línea de 21 cm del hidrógeno neutro (hi-21cm) en el contexto cosmológico: Modelación y extracción de la señal global. Master's thesis, Universidad de Guanajuato.

Burns, J. O., Lazio, J., Bale, S., Bowman, J., Bradley, R., Carilli, C., Furlanetto, S., Harker, G., Loeb, A., and Pritchard, J. (2012). Probing the first stars and black holes in the early universe with the dark ages radio explorer (dare). *Advances in Space Research*, 49(3):433–450.

de Oliveira-Costa, A., Tegmark, M., Gaensler, B., Jonas, J., Landecker, T., and Reich, P. (2008). A model of diffuse galactic radio emission from 10 mhz to 100 ghz. *Monthly Notices of the Royal Astronomical Society*, 388(1):247–260.

DeBoer, D. R., Parsons, A. R., Aguirre, J. E., Alexander, P., Ali, Z. S., Beardsley, A. P., Bernardi, G., Bowman, J. D., Bradley, R. F., Carilli, C. L., et al. (2017). Hydrogen epoch of reionization array (hera). *Publications of the Astronomical Society of the Pacific*, 129(974):045001.

Einstein, A. (1915). The field equations of gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915:844–847.

Ekers, R. (2012). The history of the square kilometre array (ska)-born global. *arXiv preprint arXiv:1212.3497*.

Ewen, H. I. and Purcell, E. M. (1951). Observation of a line in the galactic radio spectrum: Radiation from galactic hydrogen at 1,420 mc./sec. *Nature*, 168(4270):356.

Field, G. B. (1958). Excitation of the hydrogen 21-cm line. *Proceedings of the IRE*, 46(1):240–250.

Furlanetto, S. R., Oh, S. P., and Briggs, F. H. (2006). Cosmology at low frequencies: The 21 cm transition and the high-redshift universe. *Physics Reports*, 433(4-6):181–301.

Gamow, G. (1948). The evolution of the universe. *Nature*, 162(4122):680.

Greenhill, L. and Bernardi, G. (2012). Hi epoch of reionization arrays. *arXiv preprint arXiv:1201.1700*.

Harker, G. J., Mirocha, J., Burns, J. O., and Pritchard, J. R. (2015). Parametrizations of the 21-cm global signal and parameter estimation from single-dipole experiments. *Monthly Notices of the Royal Astronomical Society*, 455(4):3829–3840.

Houston, N., Li, C., Li, T., Yang, Q., and Zhang, X. (2018). A natural explanation for 21 cm absorption signals via the qcd axion. *arXiv preprint arXiv:1805.04426*.

Jáuregui, J. M. (2016). *SCI-HI: Development, testing and characterization of a 21 cm All-Sky Spectrum Experiment*. PhD thesis, Instituto Nacional de Astrofísica, Óptica y Electrónica, Tonantzintla,Puebla.

Karttunen, H., Kröger, P., Oja, H., Poutanen, M., and Donner, K. J. (2016). *Fundamental astronomy*. Springer.

Liddle, A. (2015). *An introduction to modern cosmology*. John Wiley & Sons.

Liu, H. and Slatyer, T. R. (2018). Too hot, too cold or just right? implications of a 21-cm signal for dark matter annihilation and decay. *arXiv preprint arXiv:1803.09739*.

Monsalve, R. A., Rogers, A. E., Bowman, J. D., and Mozdzen, T. J. (2017). Results from edges high-band. i. constraints on phenomenological models for the global 21 cm signal. *The Astrophysical Journal*, 847(1):64.

Morandi, A. and Barkana, R. (2012). Studying cosmic reionization with observations of the global 21-cm signal. *Monthly Notices of the Royal Astronomical Society*, 424(4):2551–2561.

Paciga, G., Chang, T.-C., Gupta, Y., Nityanada, R., Odegova, J., Pen, U.-L., Peterson, J. B., Roy, J., and Sigurdson, K. (2011). The gmrt epoch of reionization experiment: a new upper limit on the neutral hydrogen power spectrum at z$\approx$ 8.6. *Monthly Notices of the Royal Astronomical Society*, 413(2):1174–1183.

Patra, N., Subrahmanyan, R., Raghunathan, A., and Shankar, N. U. (2013). Saras: a precision system for measurement of the cosmic radio background and signatures from the epoch of reionization. *Experimental Astronomy*, 36(1-2):319–370.

Penzias, A. A. and Wilson, R. W. (1965). A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421.

Pober, J. C., Parsons, A. R., Aguirre, J. E., Ali, Z., Bradley, R. F., Carilli, C. L., DeBoer, D., Dexter, M., Gugliucci, N. E., Jacobs, D. C., et al. (2013). Opening the 21 cm epoch of reionization window: Measurements of foreground isolation with paper. *The Astrophysical Journal Letters*, 768(2):L36.

Pritchard, J. R. and Loeb, A. (2012). 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901.

Rohlfs, K. and Wilson, T. L. (2013). *Tools of radio astronomy*. Springer Science & Business Media.

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

Smoot, G. F. (1999). Summary of results from cobe. *arXiv preprint astro-ph/9902027*.

Sokolowski, M., Tremblay, S. E., Wayth, R. B., Tingay, S. J., Clarke, N., Roberts, P., Waterson, M., Ekers, R. D., Hall, P., Lewis, M., et al. (2015). Bighorns-broadband instrument for global hydrogen reionisation signal. *Publications of the Astronomical Society of Australia*, 32.

Tingay, S. J., Goeke, R., Bowman, J. D., Emrich, D., Ord, S., Mitchell, D. A., Morales, M. F., Booler, T., Crosse, B., Wayth, R., et al. (2013). The murchison widefield array: The square kilometre array precursor at low radio frequencies. *Publications of the Astronomical Society of Australia*, 30.

Van de Hulst, H. (1945). Radio waves from space. *Ned. Tijdschr. Natuurk*, 11:210–221.

van Haarlem, M. P., Wise, M., Gunst, A., Heald, G., McKean, J., Hessels, J., De Bruyn, A., Nijboer, R., Swinbank, J., Fallows, R., et al. (2013). Lofar: The low-frequency array. *Astronomy & Astrophysics*, 556:A2.

Voytek, T. C., Natarajan, A., García, J. M. J., Peterson, J. B., and López-Cruz, O. (2014). Probing the dark ages at z 20: The sci-hi 21 cm all-sky spectrum experiment. *The Astrophysical Journal Letters*, 782(1):L9.

Weinberg, S. (2008). *Cosmology*. Oxford University Press.

Wouthuysen, S. (1952). On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line. *The Astronomical Journal*, 57:31–32.

APPENDICES

Next we show the codes developed for this thesis:

- Appendix A is the code used for simulating the antenna beam pattern and trajectory through the sky, also this code generates the theoretical temperature for a given date and hour by convolving the antenna pattern with the GSM temperature.
- Appendix B has a method to filter the data collected by HIbiscus in 2013 and save it into tables in the HDF5 format for a efficient and faster analysis performance.
- Appendix C is the code used for the calibration of the data with functions that calibrate the data using the method specified either χ^2 or Johnson-noise calibration (JNC).
- Appendix D has series of transformations between the units of the data collected by SCI-HI (dBm's) into temperature units (Kelvin), these transformations are done using the HIbiscus antenna specifications.

A | Code for antenna's trajectory, beam pattern and convolution

```python
import numpy as np
import matplotlib.pyplot as plt
import healpy as hp
import pandas as pd
import os
from matplotlib import cm
from astropy.time import Time
from astropy.coordinates import SkyCoord
from astropy import units as u
from power_to_temperature import Radio_source_trans
Cmap = cm.jet
Cmap.set_under("w")

def trajectory(time,lon = -118.3011,lat = 28.9733):
    """
    Calculates trajectories for the antenna in a given time.
    Returns an array of l and b galactic coordinates values in degrees, for a given location.

    It is assumed that the antenna is looking directly to the zenit,
    so in this case the right ascension (RA) is equal to the local sidereal time of the location
    and the declination (DEC) is equal to the latitude.

    Parameters:
    time: local time of observation, can be an array or a single time. Preferred format is
            'yyyy-mm-dd hh:mm:ss'.

             WARNING: Make sure that time is in UTC.

    Optional parameters:
    lon: Default longitude is given for Isla de Guadalupe at -118.3011 degrees
    lat: Default latitude is given for Isla de Guadalupe at 28.9733 degrees
```

```
    """
    t = Time(time, location =(lon,lat))
    RA = np.array(t.sidereal_time('mean').degree)
    DEC = lat*np.ones(len(RA))
    coords = SkyCoord(ra=RA*u.degree,dec=DEC*u.degree)
    l = coords.galactic.l.degree
    b = coords.galactic.b.degree
    return l,b

def pattern(time,Freq,PATH="antenna_beam/",lon = -118.3011,lat = 28.9733):
    """
    Calculates the beam pattern for the antenna at a given time.
    Returns an array of l and b galactic coordinates values in degrees of the antenna pattern,
    for a given location and day.

    It also returns an array of Temperature lecture of the antenna for a given coordinate (l,b)

    Parameters:
    time: Time of observation, can be an array or a single time. Preferred format is
            'yyyy-mm-dd hh:mm:ss'.

             WARNING: Make sure that time is in UTC.

    Freq: Frequency of the antenna beam in MHz.

    Optional parameters:
    PATH: Folder where the pattern is stored, note that the files within this folder
            must be named with its frequency, for example 70MHz.hdf5.

          Default is antenna_beam.

    lon: Default longitude is given for Isla de Guadalupe at -118.3011 degrees
    lat: Default latitude is given for Isla de Guadalupe at 28.9733 degrees
    """
    Data = pd.read_hdf(PATH+"0%dMHz.hdf5"%Freq) #Change path if beam pattern is changed
    t = Time(time, location =(lon,lat))
    LST = t.sidereal_time('mean').degree
    theta,phi = np.radians(Data.values[:,0]),np.radians(Data.values[:,1])
    dB = Data.values[:,2]
    X,Y,Z=np.sin(theta)*np.cos(phi),np.sin(theta)*np.sin(phi),np.cos(theta)
```

```
    Rxy = np.sqrt(X**2.+Y**2.)
    colat,Az = np.arctan2(Rxy,Z),np.arctan2(Y,X)
    Alt = 0.5*np.pi-colat
    lat = np.radians(lat)
    sinDEC = np.sin(Alt)*np.sin(lat)-np.cos(Alt)*np.cos(Az)*np.cos(lat)
    DEC = np.arcsin(sinDEC)
    sinH = np.sin(Az)*np.cos(Alt)/np.cos(DEC)
    cosH = (np.cos(Az)*np.cos(Alt)*np.sin(lat)+np.sin(Alt)*np.cos(lat))
   /np.cos(DEC)
    H = np.arctan2(sinH,cosH)
    DEC = np.degrees(DEC)
    RA = LST - np.degrees(H)
    coords = SkyCoord(ra=RA*u.degree,dec=DEC*u.degree)
    l = coords.galactic.l.degree
    b = coords.galactic.b.degree
    Temp = Radio_source_trans(dB,Freq,1e6)
    return l,b,Temp

def convolve(time,Freq,PATH="antenna_beam/"):
    """
    Convolves the antenna beam pattern with the gsm map of the galaxy
   for a given frequency.
    Returns the convolved temperature of the gsm.

    Parameters:
    time: local time of observation, can be an array or a single time.
   Preferred format is
        'yyyy-mm-dd hh:mm:ss'.


         WARNING: Make sure that time is in UTC.
    Optional parameters:
    PATH: Folder where the pattern is stored, note that the files
   within this folder
           must be named with its frequency, for example 70MHz.hdf5.

          Default is antenna_beam.

    Freq: Frequency desired in MHz.

    Optional parameters:
    PATH: Folder where the pattern is stored, note that the files
   within this folder
           must be named with its frequency, for example 70MHz.hdf5.

          Default is antenna_beam.
    """
    nside = 32
    Data = pd.read_hdf("gsm_maps/gsm_%dMHz.hdf5"%Freq)
    bmap_gal = Data.values[:,0]
```

```
    bmap_gal2 = hp.ud_grade(bmap_gal,nside)
    l,b,Temp = pattern(time,Freq,PATH)
    pix = hp.ang2pix(nside,l, b, lonlat=True)
    bmap_pat = np.zeros(hp.nside2npix(nside))
    bmap_pat[pix] = Temp
    T_gsm = sum(bmap_gal2*bmap_pat)/sum(bmap_pat)
    return T_gsm

def T_gsm(time,freqs=(50,90),bins=20,days=1,PATH="antenna_beam/", OUTPUT = 'calibration'):
    """
    Provides a table of the convolved temperature of the GSM map with the
    Antenna beam pattern for a full day of observation, in a range of frequencies.
    It saves a file named Tgsm.hdf5 with the values obtained  in the calibration folder.

    Parameters:
    time: Initial date and hour of observation, preferred format is 'yyyy-mm-dd hh:mm:ss'.

            WARNING: Make sure that time is in UTC.

    Optional parameters:
    freqs: Range of frequencies, must be a tuple with initial frequency and final frequency.
           Default is 50-90

    bins: Observation interval in minutes, default is 20 minutes.
    days: Days of observation, default is 1 day.

    OUTPUT: Output folder where the data is going to be stored.

          Default is calibration.
    """
    if not os.path.exists(OUTPUT):
        os.makedirs(OUTPUT)
    Freqs = np.arange(freqs[0],freqs[1]+1)
    t0 = Time(time)
    dt = bins*u.min # Modify unit of time interval if needed
    DT = dt.to(u.hour)
    times = t0 + DT*np.arange(0,days*24/DT.value)
    data = np.zeros([len(Freqs),len(times)])
    i,j = 0,0
    for f in Freqs:
        for k in range(len(times)):
            data[i,j] = convolve(times[k],f,PATH)
            j+=1
        i+=1
```

```
        j=0
    df = pd.DataFrame(data,index = Freqs,columns = times.value)
    df.to_hdf(OUTPUT+'/Tgsm.hdf5','df')
    return df

def check_LST(time,lon = -118.3011,lat = 28.9733):
    """
    Checks the Local Sidereal time for a given time in a given location.

    Parameters:
    time: Time of observation, can be an array or a single time. Preferred format is
          'yyyy-mm-dd hh:mm:ss'.

           WARNING: Make sure that time is in UTC.

    Optional parameters:
    lon: Default longitude is given for Isla de Guadalupe at -118.3011 degrees
    lat: Default latitude is given for Isla de Guadalupe at 28.9733 degrees
    """
    t = Time(time, location =(lon,lat))
    LST = t.sidereal_time('mean')
    print 'LST time:',LST
    return LST
```

B Code for data filtering

```python
import pandas as pd
import numpy as np
import os
import glob
from power_to_temperature import *
from test import *

Freqs = np.linspace(1e-22,250,32769) # Expected range of the antenna, edit if needed.
eta_nu = eta(Freqs) # There should be a file named eta_nu.dat in the directory with the efficiency of the antenna.
mask = (Freqs>=50)&(Freqs<=91) # Edit if other bandrange is needed.
bwidth=1. # In MHz.

def Filter(PATH,Output_folder='.',outcome=0.):
    """
    Reduces the data to the selected band range (50-90 MHz) calculating the temperature of the measurements
    and stores them in HDF5 files. Data should be a table of the measurements in dBm's.

    It is expected that the files are stores in folders named with the hours of the data
    measurement.
    Example: 2013-06-13-23, 2013-06-14-00, 2013-06-14-01.

    Also the files inside these folders should end with .dat extension with different
    names for every measure type, i.e short.dat, 50ohm.dat, antenna.dat, etc.
    By default this function ignores empty .dat files.

    Parameters:
    PATH: Path of the directory with the data files.
    Output_folder: Path of the directory where the files should be stored.

    Optional parameters:
```

```
    outcome: Expected noise that should be extracted from the
   measurement, default is 0.

    Output:
    HDF5 tables of the temperature for the corresponding measure of the
     folders given.
    Also stores the gain K for the Johnson-noise calibration method.
    Also stores the ambient temperature given in the header of the
   antenna.dat files,
    if the header has no ambient temperature this script will stop.
    It is assumed that the header of the file is 19 lines long and that
     the last line of
    this header has the temperature in Celsius, if this is not the case
     modify the Tamb line.


    """
    folders = glob.glob(PATH+'/*')
    folders.sort()
    i=-1

    # Create target directories
    if not os.path.exists(Output_folder+'/short'):
        os.makedirs(Output_folder+'/short')
    if not os.path.exists(Output_folder+'/50ohm'):
        os.makedirs(Output_folder+'/50ohm')
    if not os.path.exists(Output_folder+'/antenna'):
        os.makedirs(Output_folder+'/antenna')
    if not os.path.exists(Output_folder+'/Tmeas'):
        os.makedirs(Output_folder+'/Tmeas')
    if not os.path.exists(Output_folder+'/K_jnc'):
        os.makedirs(Output_folder+'/K_jnc')

    for subdirs, dirs, files in os.walk(PATH):
        dirs[:] = [d for d in dirs if not d.startswith('.')] # Ignore
   hidden folders (ipynb checkpoints for example)
        dirs.sort()
        files.sort()
        short,antenna,_50ohm,measure,K_jnc = [],[],[],[],[]
        short_date,_50ohm_date,measure_date =[],[],[]

        # Walk through directories
        for file in files:
            path = os.path.join(subdirs,file)
            date = file.split("_")[0]
            if  os.path.getsize(path)==0: # Filtering empty data
                print 'EMPTY FILE:',path
                continue

            data =  np.loadtxt(path,unpack=True)
```

```
            if data.size == 0:
                print 'NO DATA IN FILE:',path
                continue

            elif file.endswith('short.dat'):
                T_short = Res2Temp(data,bwidth)
                short.append(T_short),short_date.append(date)
            elif file.endswith('50ohm.dat'):
                T_50ohm = Res2Temp(data,bwidth)
                _50ohm.append(T_50ohm),_50ohm_date.append(date)
            elif file.endswith('noise.dat'):
                dB_noise = data
            elif file.endswith('antenna.dat'):
                dB_antenna = data
                dB_clean = dB_antenna - dB_noise - outcome
                T_antenna = Radio_source_trans(dB_clean, Freqs, bwidth)
                T_measure = T_antenna/eta_nu - T_short # Uncalibrated measure
                Tamb = round(np.genfromtxt(path,comments='!',skip_header= 18,max_rows=1)[1]+273.15,2)
                Kjnc = Tamb/(T_50ohm-T_short) # Johnson-noise calibration coefficient
                antenna.append(T_antenna),measure.append(T_measure),K_jnc.append(Kjnc)
                measure_date.append(date)

        # HDF5 Table Generation
        if i>=0 and i<len(folders) and short and antenna and _50ohm and measure and K_jnc:
            name = os.path.normpath(folders[i])
            name = name.split("/")[1]
            short = np.transpose(short)
            antenna = np.transpose(antenna)
            _50ohm = np.transpose(_50ohm)
            measure = np.transpose(measure)
            K_jnc = np.transpose(K_jnc)

            short_table = pd.DataFrame(short[mask], index = Freqs[mask], columns = short_date)
            short_table.to_hdf(Output_folder+'/short/'+name+'.hdf5','df')
            _50ohm_table = pd.DataFrame(_50ohm[mask], index = Freqs[mask], columns = _50ohm_date)
            _50ohm_table.to_hdf(Output_folder+'/50ohm/'+name+'.hdf5','df')
            antenna_table = pd.DataFrame(antenna[mask], index = Freqs[mask], columns = measure_date)
            antenna_table.to_hdf(Output_folder+'/antenna/'+name+'.hdf5','df')
```

```
            measure_table = pd.DataFrame(measure[mask], index = Freqs[mask], columns = measure_date)
            measure_table.to_hdf(Output_folder+'/Tmeas/'+name+'.hdf5','df')
            Kjnc_table = pd.DataFrame(K_jnc[mask], index = Freqs[mask], columns = measure_date)
            Kjnc_table.to_hdf(Output_folder+'/K_jnc/'+name+'.hdf5','df')
        i+=1
```

C | Code for calibration

```python
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sb
import scipy.optimize as op

def Calibrate(PATHS, dates, calibration = 'Chisq'):
    """
    Returns the Gain K for the calibration selected, also returns
    T_gsm,T_meas and T_std, the theoretical gsm temperature, the measured
    and its standard deviation respectively.

    Parameters:
    PATHS: Path of the data to calibrate, if JNC calibration is chosen this
           variable must be a tuple in the form (DATA_PATH,KJNC_PATH).

           KJNC_PATH must be the path to the Kjnc generated by the filter function
           in the filtering process.


    dates: 4 dim Array, must contain the initial date for the measured data,
           the final date for the measured data, initial date for the
           theoretical data and final date for theoretical data.

           (Date_data0,Date_data1,Date_teo0,Date_teo1)

    Optional:
    calibration: Choose between 'Chisq' or 'JNC', it defines how the data
                 calibration will be done.


    RETURNS: 4 variables, K,T_gsm,T_meas,T_std
    """
```

```python
    if isinstance(PATHS,(tuple)):
        PATH = PATHS[0]
        PATH_JNC = PATHS[1]
    else:
        PATH = PATHS

    Temps = pd.read_hdf(PATH)
    Date_data0,Date_data1,Date_teo0,Date_teo1 = dates
    temp = Temps.loc[:,Date_data0:Date_data1]
    data = temp.values
    data_gsm = pd.read_hdf('calibration/Tgsm.hdf5')
    data_gsm = data_gsm.loc[:,Date_teo0:Date_teo1]
    T_gsm = data_gsm.values
    freqs = data_gsm.index.values
    index = Temps.index.values

    if calibration=='JNC':
        Kjnc = pd.read_hdf(PATH_JNC)
        Temps = Kjnc*Temps

    freq_bins = []
    for f in freqs:
        mask = (index>=f)&(index<f+1)
        freq_bins.append(np.mean(temp[mask]))
    Temp_binfreq = np.array(freq_bins)

    bins = int(np.shape(Temp_binfreq)[1]/25. +1)
    # Data binning in time, 25 is because every 25 data is 5 mins (Aprox 12 secs between data)
    T_meas = np.zeros(np.shape(T_gsm))
    T_std = np.zeros(np.shape(T_gsm))
    for i in range(bins):
        T_meas[:,i] = np.mean(Temp_binfreq[:,25*i:25*(i+1)], axis =1)
        T_std[:,i] = np.std(Temp_binfreq[:,25*i:25*(i+1)], axis =1)

    if calibration=='Chisq':
        def Chisq(k,Tmeas,Tgsm, error):
            Chi = (Tgsm- k*Tmeas/1.e19)**2./(error/1.e19)**2.
            return sum(Chi)
        fun = lambda *args: Chisq(*args)

        K = np.zeros(len(freqs))
        for i in range(len(freqs)):
            K[i] = op.minimize(fun,1,args=(T_meas[i,:],T_gsm[i,:],T_std[i,:])).x[0]
        K = K*1e-19
    elif calibration =='JNC':
        K_jnc = Kjnc.loc[:,Date_data0:Date_data1]

        freq_bins = []
```

```
        for f in freqs:
            mask = (index>=f)&(index<f+1)
            freq_bins.append(np.mean(K_jnc[mask]))
        Kjnc_binfreq = np.array(freq_bins)
        K_meas = np.zeros(np.shape(T_gsm))

        for i in range(bins):
            K_meas[:,i] = np.mean(Kjnc_binfreq[:,25*i:25*(i+1)], axis=1)
        K = np.mean(K_meas,axis=1)

    Tgsm = np.mean(T_gsm,axis=1)
    Tmeas = np.mean(T_meas,axis=1)
    Tstd = np.mean(T_std,axis=1)
    return K,Tgsm,Tmeas,Tstd

def Check_quality(PATH,dates,savepath=False):
    """
    Function for plotting the desired interval data for
    visual check of quality.

    Parameters:
    PATHS: Path of the data to check.
    dates: 2 dim Array, must contain the initial and final date
           of the data to check.

    Optional:
    savepath: Path to save the figure
    """
    Temps = pd.read_hdf(PATH)
    Temps.index = Temps.index.values.round(2)
    Date_0,Date_1 = dates
    temp = Temps.loc[:,Date_0:Date_1]
    sb.heatmap(np.log10(temp), cmap='YlGnBu', yticklabels=400)
    plt.xlabel('Days')
    plt.ylabel('Frequency (MHz)')
    if savepath != False:
        plt.savefig(savepath+'data_%s_%s.png'%(Date_0,Date_1), bbox_inches='tight')
    plt.show()
```

D Code for dBm's to temperature transformations

```python
import numpy as np

k = 1.38064852e-23 #Boltzmann constant
c = 299792458.0    #speed of light m/s

#dBm's to Power conversion
P = lambda source: 10.0**((source - 30.) / 10.0)

def deg2arcsec(angle):
    """
    Transformation of a angle into arcsecs.

    Parameters:
    angle: antenna beam solid angle in deg for transformation to
    arcsecs.
    """

    asec = angle * 3600.0
    return asec


def Radio_source_trans(Radio_source, freqs, Bwidth):
    """
    Transforms dBm's to Temperature for the HIbiscus antenna

    Parameters:
    Radio_source: Data of the antenna in dBm's to be converted to
    temperature
    freqs: The frequency range in MHz
    Bwidth: Bandwidth in Hz
    """

    area = 1.0            # m^2
    angle = 55.0     #degrees
    theta = deg2arcsec(angle)
    power = P(Radio_source)
```

```

    #the units of the flux density are W m^-2 MHz^-1
    flux = (2.0 * power / area) * Bwidth
    flux_Jy = flux * 1e26  # Jy
    flux_Jy = flux_Jy * 1e3 # mJy
    freq = freqs * 1e6 #Hz
    wavelength = (c / freq) * 100.  # cm
    T = 1.36 * flux_Jy  * wavelength**2 / theta**2
    return T

def Res2Temp(res_source, Bwidth):
    """
    Transfomation to temperature for a electronic source

    Parameters:
    res_source: Data of the antenna in dBm's to be converted to temperature
    Bwidth: Bandwidth in Hz
    """
    power = P(res_source)
    T = power / (k * Bwidth)
    return T
```