21 cm Hydrogen Line and Its Power Spectrum

Levent Toksoz
Supervised by: Dragan Huterer

April 18, 2017

Contents

Abstract

In this project, a 21 cm hydrogen line was investigated to forecast the precision of cosmological parameters defined by a chosen model in the presence of statistical and systematic errors. The 21 cm line was a signal used to determine the distribution of matter in the universe and was observed using radio surveys. Using this signal, one can plot its power spectra and analyze how parameters from a given model affect this spectra. This allows the sensitivity of the power spectrum to be determined with respect to changes in each parameter. By combining this information with the noise and the cosmic variance of the data and utilizing statistical techniques, such as the Fisher matrix, the cosmological constraints imposed on the parameter space was obtained and the precision of the results was predicted.

1 Acknowledgements

I would like to thank my research advisor, Professor Huterer, for providing me the oppurtunity to work on this project. He guided and supported me throughout this journey. I am grateful for all of his invaluable contributions to my undergraduate education.

2 Introduction

As the early universe expanded, it began to cool down, leading to universe a phase called the epoch of recombination. During this phase, ionized plasma of H, created through nucleosynthesis, began a phase transition and quickly became neutral. After the recombination era, photons decoupled from matter in the universe, traveling freely and forming cosmic microwave background (CMB) radiation. As a result, the universe become transparent. These neutral hydrogen and helium atoms became the seeds of the first cosmological structures, such as stars and quasars. Later, high-energy photons created by these stars and quasars began to reionize the hydrogen in the intergalactic medium, eventually leading to the formation of large-scale structures. [1]

By observing the neutral hydrogen atom and analyzing its power spectrum, the statistical distribution of matter over time can be inferred and the evolution and formation of the very first galaxies and stars can be analyzed. However, neutral atoms do not emit light in a traditional sense; thus, quantum mechanical transitions between hyperfine energy levels in the hydrogen must be used. This hyperfine splitting is caused by the interaction of the magnetic moments of the proton and electron, which together makes up the hydrogen atom. As result of this splitting, the neutral hydrogen atom emits microwave radio radiation with a wavelength of 21 cm and a frequency of 1,420 MHz. The lifetime of neutral hydrogen before spontaneously de-exciting to a lower energy state is approximately 10 million years; however, its high abundance in any given region makes it possible to analyze neutral hydrogen and obtain its power spectrum.

Using the 21 cm measurements, cosmologists seek to detect radiation from the far edges of the universe to map distribution of the mater within the universe using this signal. There are currently ongoing projects to detect these signals throughout the world, but the significant systematic and statistical noise caused by cell phones and other human-related emissions can be as large as five orders of the magnitude of the cosmological signal, making it difficult to detect the 21 cm signals accurately.

3 Aim and Motivation

In this project, a 21 cm hydrogen line and power spectrum was investigated to infer the precision of cosmological parameters in the presence of statistical and systematic errors. The model parameters were chosen from the proposed model, which is described in part five of this thesis. A python and code for anisotropies in the microwave background (CAMB) were used to simulate data for a 21 cm hydrogen line and to obtain the line's power spectrum. The CAMB can be used to obtain power spectra from various correlation functions, such as temperature cross temperature (TxT). However, in this project, the power spectra of gravitational fluctuations were represented in the CAMB as a gravitational cross gravitational (wxw) correlation function. Gravitational fluctuations are explained in the fourth chapter of this thesis. Noise and cosmic variance in the data were analyzed and used to compute the Fisher matrix for the following three variables: $\Omega_b h^2$,$\Omega_c h^2$, and H. The specifics of these computations are discussed in their respective chapters. Using the Fisher matrix and approximating the data as Gaussian, the cosmological constraints imposed on the three-variable parameter space were obtained and the precision of the results was predicted.

4 Theoretical Background

On a scale of roughly 100 Mpc or more, the universe is both homogeneous and isotropic. On smaller scales (around 50 Mpc and less) the universe is not homogeneous, and it contains density fluctuations varying from fluctuations in subatomic levels to those in large superclusters and voids. These fluctuations occur due to the slight inhomogeneity of the early universe. Modern cosmology can trace these fluctuations to the time of the last scattering, shortly after the photons decoupled from matter in the universe, by following their stamp on the CMB. These fluctuations were amplified during the matter-dominated phase of the universe, during which overdense regions of the universe began to collapse and form bound objects, such as clusters. At the same time, these regions were drawing matter from surrounding underdense regions, and as this process continued, denser regions became denser and underdense regions became less dense, eventually leading to the formation of large-scale cosmological structures that were larger than individual galaxies.

Putting this process into a quantitative context and using the notation and formulas given in Ryden‘s Introduction to Cosmology[1], let $\epsilon(\vec{r}$,t) represent the enrgy density of the universe as a function of position and time. For any given time t, the spatially averaged energy density is

$$\overline{\epsilon}(t) = \frac{1}{V}\int_V \epsilon(\vec{r},t)d^3r \tag{1}$$

In this equation V represents a volume larger than the size of the largest structure in the universe. Now, define dimensionless density fluctuations as

$$\delta(\vec{r},t) \equiv \frac{\epsilon(\vec{r},t) - \overline{\epsilon}(t)}{\overline{\epsilon}(t)} \tag{2}$$

In this equation, δ takes negative values in underdense regions and positive values in overdense regions. The minimum value of δ is -1, which correponds to $\epsilon = 0$. In this notation, small fluctuations correpond to $|\delta|\ll 1$. For simplicity, static is homogenous and a matter-only universe has a uniform mass density of $\overline{\rho}$. To simplify computations even further, the evolution of density fluctuations within a sphere of radius R is considered, and by adding small amounts of mass to this sphere, the density within the sphere is $\overline{\rho}(1+\delta)$. Assuming that excess density (δ) is uniform within the sphere, the gravitational acceleration due this excess mass is computed as

$$\frac{\ddot{R}}{R} = -\frac{4\pi G\overline{\rho}\delta(t)}{3} \tag{3}$$

Notice that for $\delta>0$, $\ddot{R}<0$, which means excess mass represented by δ causes the sphere to collapse inwards.

Now, using the conservation of mass (M) is

$$M = \frac{4\pi}{3}\overline{\rho}(1+\delta(t))R(t)^3 \tag{4}$$

Because M is constant, the equation is rewritten as

$$R(t) = R_0(1+\delta(t))^{-\frac{1}{3}} \tag{5}$$

where

$$R_0 = \left(\frac{3M}{4\pi\overline{\rho}}\right)^{\frac{1}{3}} \tag{6}$$

For $\delta \ll 1$, it can be approximated as

$$R(t) \approx R_0\left(1 - \frac{\delta(t)}{3}\right) \tag{7}$$

Taking the second derivative and performing further approximations, yields

$$\frac{\ddot{R}}{R} \approx \frac{-\ddot{\delta}}{3} \tag{8}$$

Combining this with Equation 3, the following equation is obtained,

$$\ddot{\delta} = 4\pi G\overline{\rho}\delta \tag{9}$$

The general solution is

$$\delta(t) = A_1 e^{\frac{t}{t_d}} + A_2 e^{\frac{-t}{t_d}} \tag{10}$$

where A_1 and A_2 come from the initial conditions of the sphere, and for simplicity, can be set to $\delta(0)/2$ and

$$t_d = \frac{1}{(4\pi G\overline{\rho})^{\frac{1}{2}}} \tag{11}$$

If Equation 10 is analyzed after several t_d's, only the exponential growth term remains significant, which indicates that gravity causes small density fluctuations to grow exponentially within the sphere.

Finally, the density fluctuations can be used to compute the power spectrum. For a given comoving volume (V) that is not affected by the expansion of space, the density fluctuation field, $\delta(\vec{r})$, is given as [1]

$$\delta(\vec{r}) = \frac{V}{(2\pi)^3} \int \delta_{\vec{k}} e^{-i\vec{k}.\vec{r}} d^3k \tag{12}$$

where $\vec{k}$ represents the wave number;$\delta_{\vec{k}}$ represents the individual Fourier components and is computed by

$$\delta_{\vec{k}} = \frac{1}{V} \int \delta(\vec{r}) e^{i\vec{k}.\vec{r}} d^3r \tag{13}$$

By performing a Fourier transformation, the statistical properties of the field are obtained rather than the exact locations of density maxima and minima. The mean square amplitude of the Fourier components gives the power spectrum.

$$P(k) = <\mid \delta_{\vec{k}} \mid^2> \tag{14}$$

where the average is taken over all possible orientations of $\vec{k}$. In this project, the CAMB was used to obtain the power spectrum, which uses a formula that is similar to the one in Equation 14 to compute the 21 cm powers pectrum.

5 Choosing a cosmological model and obtaining a power spectrum using CAMB

Due to the complicated nature of the 21 cm power spectrum, it is not possible to obtain simple analytical expressions for this spectrum; thus, CAMB sources were used to simulate it. The CAMB is a code package written in Fortran and used to simulate cosmological events. CAMB sources are code packages based on the CAMB; however, these are not as well tested and accurate as the CAMB. For the purposes of this project, CAMB sources were sufficiently accurate to obtain reasonable values for the power spectrum and cosmological constraints.

According to the CAMB sources website, the CAMB uses finite-width window functions, integrated with parts and Limber approximations to obtain smooth window functions for all L values. More specifically, the CAMB uses direct integration of source functions against spherical Bessel functions for low L values. The results obtained at these values are spherical.[2] More detailed information can be found on the CAMB sources website.

The CAMB can produce auto- and cross-correlation spectra of various correlation functions, such as TxT and CMB lensing potential (ΦxΦ) cross-correlation spectra. It can also produce angular power spectra of these functions. However, in this project, the focus were wxw cross-correlation spectra. The obtained power spectra are unlensed and dimensionless.

To obtain the power spectrum, a set of variables was chosen and assigned as input for the CAMB sources. Some of the input variables are as follows,

Table 1: CAMB parameters

Parameters	Value	Physical Interpretation
$\Omega_b h^2$	0.0226	Physical Baryon Density
$\Omega_c h^2$	0.112	Physical Dark Matter Density
$\Omega_\nu h^2$	0.00064	Physical Neutrino Density
Ω_k	0	Curvature
H	70	Hubble
w	-1	Dark Energy
T_{cmb}	2.7255	Temperature of CMB
$He_{fraction}$	0.24	Helium Fraction

Here, $\Omega_b h^2$,$\Omega_c h^2$ and H were used as parameter spaces to analyze how the power spectrum changed with respect to these parameters and how this change can be used to infer the cosmological constraints that are imposed on these paramaters. To achieve this, the Fisher matrix of these parameters was analyzed to compute the noise of the data and cosmic variances.

6 Modelling the Noise and Cosmic Variance

The statistical and systematic errors were due to two factors: instrumental noise and cosmic variance. Cosmic variance represents errors due to a finite number of samples of the sky, which is common at large scales, and corresponds to low multipole(ℓ) values. Noise represents errors due to finite resolution, which is common at small scales, and corresponds to high ℓ values. This dependence is illustrated in Figures 2 and 3.

Cosmic variance can be computed by

$$\sigma_{cv} = \sqrt{\frac{2}{2\ell + 1}} C_l \tag{15}$$

Shiraishi, Muñoz, Kamionkowski, and Raccanelli's model for noice was used to effectively simulate noise. [3] The noise formula is given as

$$\ell^2 C_l^N = \frac{(2\pi)^3 T_{sys}^2(\nu)}{\Delta\nu t_0 f_{cover}^2} \left(\frac{\ell}{l_{cover}(\nu)} \right)^2 \tag{16}$$

where

$$T_{sys}(\nu) = 180 \left(\frac{\nu}{180MHz} \right)^{-2.6} K \tag{17}$$

and

$$l_{cover}(\nu) = 2\pi D_{base}/\lambda \tag{18}$$

In these equations, ν represents the frequency , $\Delta\nu$ represents the bandwith of the chosen survey, $l_{cover}(\nu)$ represents the maximum multipole observable,D_{base} represents the largest baseline of the interferometer, T_{sys} represents the synchrotron temperature of the observed sky, and T_{sys} is the amplitude of the noise. A square kilometre array (SKA) was used for the reference survey, and noise was calculated based on the parameters of this survey. Parameter values of the SKA were taken from Shiraishi, Muñoz, Kamionkowski, and Raccanelli‘s [3] and Cooray‘s papers[4]. The results are shown in Figure 1-3.

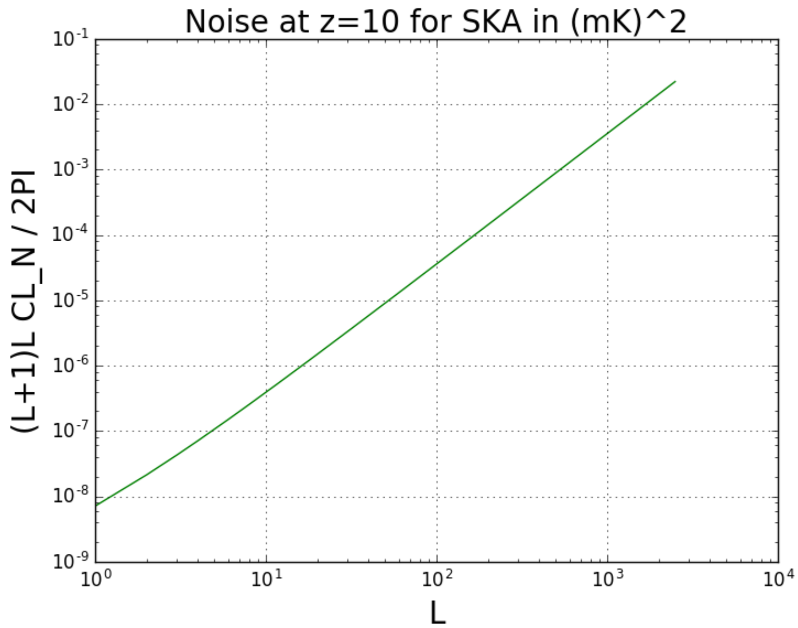

Figure 1: Evolution of noise at z=10 for SKA. Equation 16 is used to obtain this graph.

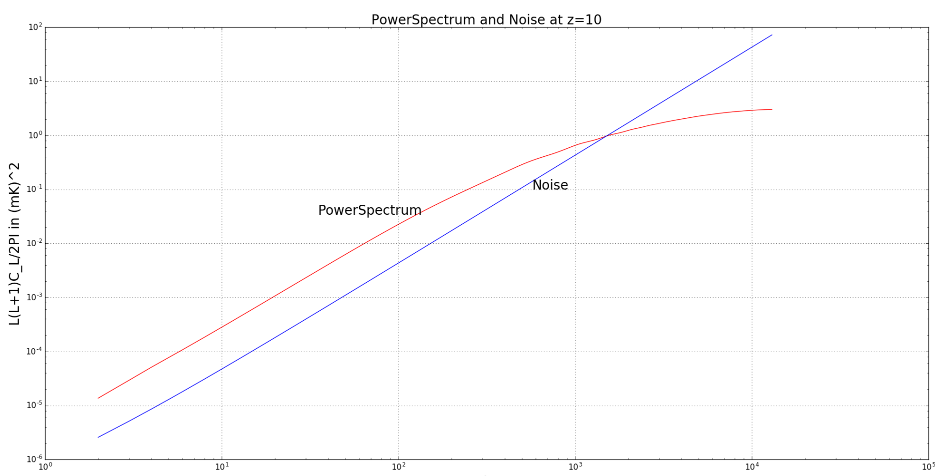

Figure 2: Power spectrum of signal and noise at z=10. At high ℓ noise dominates the graph.

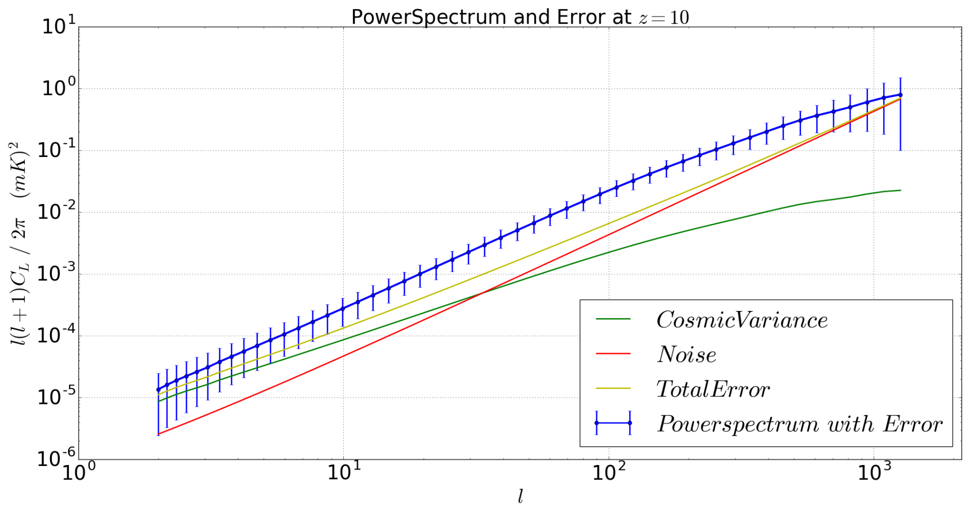

Figure 3: Power spectrum and Total Error at z=10. Equation 15 is used to obtain the CV curve. At low ℓ values CV dominates the graph.

In Figures 2 and 3, noise dominates at small scales (i.e., high ℓ values), and cosmic variance dominates at large scales (i.e., at low ℓ values), as expected. The total error in Figure 3 was computed by adding cosmic variance and noise values.

To analyze how each parameter from the chosen paramater space affects the power spectrum, each parameter was varied slightly to obtain the following power spectrum.

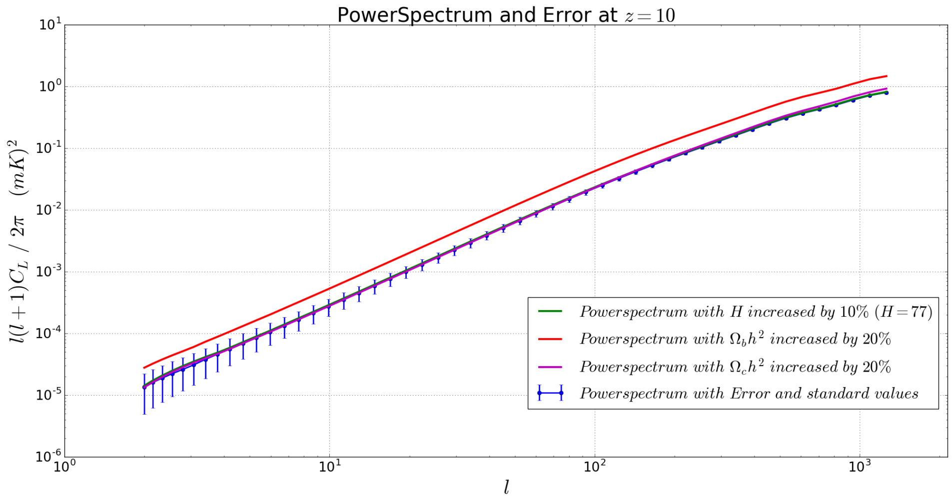

Figure 4: Power spectrum for varying parameters. The value of each parameter is changed slightly while keeping others constant.

Figure 4 illustrates how sensitive the power spectrum is based on the changes in each parameter. For example, varying H and $\Omega_c h^2$ does not change the power spectrum significantly. However, changing $\Omega_b h^2$ does affect the power spectrum significantly.

7 Fisher Matrix and Cosmological Constraints due to our Cosmological model

Fisher matrices use Gaussian uncertanities to forecast the precision of experimental data and offer an approximation of the actual distribution [5]. In this thesis, the Fisher matrix was used to derive cosmological constraints that are imposed on parameter space to infer the precision of the data. Using the power spectrum and the total error, the Fisher matrix was calculated as follows:

$$\delta_R = \sqrt{\frac{2}{(2l+1)f_{sky}}}(C_l + N_l) \tag{19}$$

where δ_R represents the total error, C_l term represents the contribution from the cosmic variance, and N_l term represents the contribution from the noise. $f_s ky$ term is added to accommodate a limited coverage of the sky. Then the Fisher matrix is given as

$$F_{ij} = \sum_l \frac{dC_l}{dp_i}\frac{dC_l}{dp_j}\frac{1}{\delta_R^2} \tag{20}$$

Where p_i represents the chosen parameter i. The covariance matrix is the inverse of the Fisher matrix. Thus, by inverting the Fisher matrix and taking the square root of the diagonal elements in the resultant matrix, the standard deviation of each parameter was computed. By doing this, the cosmological constraints that are imposed on the parameters were identified to predict the precision of the data. More quantatively,

$$[C] = [F]^{-1} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{bmatrix} \tag{21}$$

where C represents the covariance matrix and F represents the Fisher matrix. In this case, x $=\Omega_b h^2$, y$=\Omega_c h^2$, and z=H. Thus, following this procedure, the following sigma values are obtained: $\sigma_{\Omega_b h^2}$= 0.00033, $\sigma_{\Omega_c h^2}$=0.00514, and σ_H=16.8. Most of the values are approximately $\pm 2\sigma$ away from the expected value, and the uncertainity in the parameters can be approximated as $\pm 2\sigma$ to obtain

$$\begin{aligned} \delta_{\Omega_b h^2} &= 0.0006, \\ \delta_{\Omega_c h^2} &= 0.010, \\ \delta_H &= 33. \end{aligned}$$

The standard deviation (sigma values) obtained in this chapter was consistent with the sensitivity of the power spectrum with respect to each parameter. The more sensitive (i.e., the smaller the standard of deviation) a parameter was, the more significantly the parameter changed the power spectrum; this allowed the precision of the spectrum to be predicted more effectively because slightly changing a parameter caused relatively large changes in the power spectrum. Thus, only small changes are required for each parameter to obtain most of the acceptable values. Figure 4 shows how varying H and $\Omega_c h^2$ does not change the power spectrum significantly; thus, relatively large sigma values are seen, and precision cannot be predicted accurately. However, changing $\Omega_b h^2$ affects the power spectrum significantly and allows relatively small sigma values to be obtained and the precision of $\Omega_b h^2$ to be predicted accurately. Thus, the spectrum used was more effective for computing the precision of $\Omega_b h^2$ than that of H and $\Omega_c h^2$.

8 Conclusion

To sum up, in this project a 21 cm hydrogen line and power spectrum was analyzed to predict the precision of cosmological parameters in the presence statistical and systematic errors. The model parameters were chosen from a three variable model space defined as $\Omega_b h^2$, $\Omega_c h^2$, H. Analyzing how parameters from a three variable model affect this spectra, sensitivity of the power spectrum was determined with respect to changes in each parameter. According to the three variable model, the most sensitive parameter is $\Omega_b h^2$. By combining this information with the noise and the cosmic variance of the data and utilizing statistical techniques, such as the Fisher matrix, the cosmological constraints imposed on three variable model space was obtained. From these constraints its concluded that the proposed model is more effective for computing the precision of $\Omega_b h^2$ than that of H and $\Omega_c h^2$.

References

[1] Barbara Ryden. *Introduction to Cosmology.* Addison-Wesley, San Francisco, 2003.

[2] Camb sources. [Online; accessed 02-April-2017].

[3] Maresuke Shiraishi, Julian B. Muñoz, Marc Kamionkowski, and Alvise Raccanelli. Violation of statistical isotropy and homogeneity in the 21-cm power spectrum. *Phys. Rev.*, D93(10):103506, 2016.

[4] Asantha Cooray. Large-scale non-Gaussianities in the 21 cm background anisotropies from the era of reionization. *Mon. Not. Roy. Astron. Soc.*, 363:1049, 2005.

[5] D. Coe. Fisher Matrices and Confidence Ellipses: A Quick-Start Guide and Software. *ArXiv e-prints*, June 2009.