

# SKA DATA CHALLENGE-3

Submitted by: **Santanu Das**

Roll no: **PH22C040**

Guided by: **Dr. Samir Choudhuri**



**Indian Institute of Technology Madras**  
**Department of Physics**

November 29, 2024

# Declaration

I, **Santanu Das**, hereby declare that the project work titled “**SKA DATA CHALLENGE-3**” is a result of my efforts and has been carried out under the guidance of **Dr. Samir Choudhuri** in the Department of Physics, **Indian Institute of Technology Madras**.

I confirm that:

1. The data, findings, and interpretations presented in the project are original and have not been copied from any other sources.
2. Any references or sources used in the project have been appropriately cited and acknowledged.
3. I have followed all ethical guidelines and standards required for the completion of this project.
4. The project has not been submitted, either in part or full, for any other degree or qualification at this or any other institution.

I understand that any violation of the above statements will nullify my work and may result in further academic or legal consequences.

## SUBMITTED BY

Santanu Das  
Roll: PH22C040

## SUPERVISED BY

Dr. Samir Choudhuri  
Assistant Professor, IIT Madras  
Chennai-600036  
Date: 20/05/2024

# Acknowledgements

I want to express my genuine gratitude to my project supervisor Dr. Samir Choudhuri for his constant guidance and support throughout this project work. His enthusiasm and interest in cosmology and astronomy to constantly engage me in novel research work. I truly value my time working with him. I am also thankful to Shouvik Sarkar and Shashank Ramesh for supporting me through their knowledge and guidance. I learnt a lot from them, which would help me with my future academic endeavours. It was a great experience for me to work with them.

# Abstract

The 21-cm wavelength, emitted by neutral hydrogen atoms, serves as a powerful tool in our quest to understand the universe's cosmic evolution. With advancements in radio astronomy technology and sophisticated data analysis techniques, the study of the 21-cm signal promises to unlock key insights into fundamental cosmological questions and draw an idea about the process of cosmic dawn and epoch of reionization. In this project, we will apply TGE(Tapered Gridded Estimator) to the data that we got from the SKA1-Low(Square Kilometre Array) telescope. We have generated an Angular and cylindrical power spectrum and tried to compare it with the ito validator spectrum for the validation of our method. We also compare the estimated Angular power spectrum with the existing cosmological 21-cm angular power spectrum features.

# Contents

<b>1</b>	<b>Introduction:</b>	<b>5</b>
<b>2</b>	<b>Theory</b>	<b>7</b>
2.1	How 21 cm signal is produced: . . . . .	7
2.2	Concept of redshift and how 21-cm signal stretched out: . . . . .	8
2.3	Concept of visibility and sky image: . . . . .	8
2.4	Sky image: . . . . .	9
2.5	introduced primary beam pattern in the visibility expression: . . . . .	10
2.6	TGE(tapered gridded estimator): . . . . .	10
2.7	Angular power spectrum: . . . . .	12
2.8	Cylindrical power spectrum: . . . . .	12
<b>3</b>	<b>Data specification:</b>	<b>13</b>
3.1	SKA Data Specification . . . . .	13
3.2	Telescope specification: . . . . .	13
3.2.1	Sensitivity of SKA1-LOW: . . . . .	15
3.2.2	Primary beam pattern for some specific channels: . . . . .	15
<b>4</b>	<b>Data Processing Steps and Results</b>	<b>17</b>
4.1	Data Reduction: . . . . .	17
4.2	Processing the Combined UVfits file: . . . . .	18
4.2.1	Angular power spectrum $cl(\Delta\nu)$ : . . . . .	21
4.2.2	Estimating cylindrical power spectrum of the test data for the validation of the TGE estimator: . . . . .	24
4.2.3	Percentage deviation of the generated Power Spectrum from the model power spectrum: . . . . .	25
4.2.4	Cylindrical power spectrum for true data: . . . . .	26
4.3	Conclusion: . . . . .	26

# Chapter 1

## Introduction:

The study of 21-cm astronomy presents a remarkable opportunity to unlock the secrets of the universe's evolution. This field focuses on observing the faint radio emissions at a wavelength of 21 centimeters, emitted by neutral hydrogen atoms. These emissions provide a unique window into the early universe, offering insights into its formation, structure, and evolution. One of the most significant contributions of 21-cm astronomy lies in its ability to probe the cosmic dawn, a period shortly after the Big Bang when the universe transitioned from a hot, ionized state to a cooler, neutral one. During this epoch, neutral hydrogen atoms began to form as the universe expanded and cooled. By detecting the 21-cm radiation emitted by these atoms, we can study the distribution and properties of neutral hydrogen gas in the early universe. This allows us to trace the formation of the first stars and galaxies. Furthermore, 21-cm observations provide crucial insights into the process of the cosmic reionization phase in the universe's history. This epoch marked the ionization of the intergalactic medium by the first luminous sources. By mapping the evolution of the 21-cm signal across cosmic time, astronomers can trace the progress of reionization and constrain the properties of the sources responsible for it. This includes the timing of the first star formation and the growth of early galaxies, as well as their impact on the surrounding intergalactic gas. Moreover, 21-cm astronomy offers a unique perspective on the "dark ages" of the universe. During this epoch, neutral hydrogen dominated the cosmos, leaving behind a faint imprint in the form of 21-cm radiation. By studying the characteristics of this signal, we can investigate the properties of primordial gas clouds, the nature of dark matter, and the mechanisms governing the transition from the dark ages to the cosmic dawn.

The detection of the 21 cm signal is facilitated by several major telescopes operating in the low-frequency radio range. The Low-Frequency Array (LOFAR), situated in Europe, and the Murchison Widefield Array (MWA), located in Western Australia, are two prominent instruments dedicated to probing the cosmos at frequencies between approximately 30 to 300 MHz. These telescopes have played pivotal roles in studying the cosmic dawn and the epoch of reionization, where the 21 cm signal carries crucial information about the formation and evolution of the universe's first structures. Additionally, the Precision Array for Probing the Epoch of Reionization (PAPER) and the Giant Metrewave Radio Telescope (GMRT) are instrumental in detecting the 21 cm signal, focusing

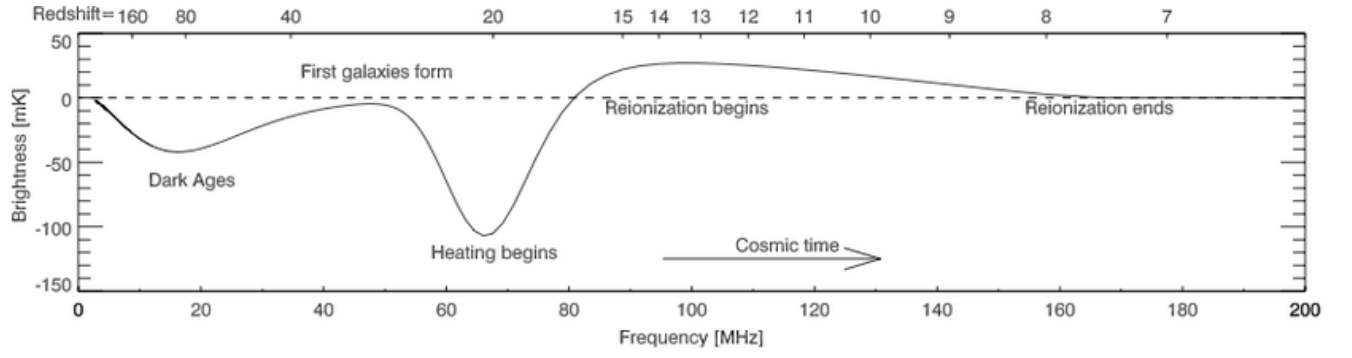


Figure 1.1: Evolution of universe through global 21-cm signal. The above figure is taken from [Pritchard and Loeb](#) here, we see that the 21cm signal goes through various phases(dark ages( $200 > z > 25$ ), cosmic dawn ( $25 > z > 16$ ), Heating( $16 > z > 10$ ), Epoch of reionization( $10 > z > 6$ ) and post reionization phase) as the universe expands( $z$  stands for redshift).

particularly on the reionization era. Operating in the frequency range of approximately 100 to 200 MHz, PAPER employs an interferometric array of antennas to capture the faint signals from neutral hydrogen emissions. Meanwhile, the GMRT, located in India, provides high sensitivity and resolution for studying the 21 cm signal and other astrophysical phenomena. Furthermore, the Australian Square Kilometre Array Pathfinder (ASKAP) and the Square Kilometre Array (SKA) promise to revolutionize 21 cm astronomy with their advanced capabilities. ASKAP operates in the frequency range of 700 MHz to 1.8 GHz, while SKA will cover frequencies from 70 MHz to 10 GHz. These telescopes will offer unprecedented sensitivity, wide field-of-view surveys, and high-resolution imaging, enabling comprehensive studies of the 21 cm signal across different cosmic epochs.

SKA1-Low boasts a high level of sensitivity. This enhanced sensitivity allows it to detect faint signals from neutral hydrogen emissions, even amidst cosmic noise and interference. It covers a broad frequency range, including the critical frequency range associated with the redshifted 21 cm signal. Its ability to observe multiple frequency bands simultaneously enables comprehensive studies of the 21 cm signal across different cosmic epochs. With its wide field of view, SKA1-Low can survey large regions of the sky efficiently. This capability is crucial for mapping the distribution of neutral hydrogen over cosmological volumes and for conducting large-scale surveys to study cosmic structure formation. It offers high angular resolution, allowing astronomers to resolve fine-scale structures in the 21 cm signal.

In this project, firstly we are trying to validate TGE by applying it to the test data, which is given by the SKA data challenge for the validation(SKAL-low telescope). Secondly, we also have planned to make 6 bands of frequency from the actual data, and then we will apply TGE there to estimate the cylindrical power spectrum. We also have a plan to construct the angular power spectrum as the cylindrical power spectrum is nothing but the Fourier transform of it.

# Chapter 2

## Theory

### 2.1 How 21 cm signal is produced:

The 21 cm signal theory is a fundamental concept in astrophysics and astronomy, providing crucial insights into the properties of neutral hydrogen in the universe. This signal arises from the hyperfine transition of neutral hydrogen atoms, emitting radiation at a wavelength of approximately 21 centimetres. It is a powerful tool in astrophysics, offering a unique window into the universe's early history and evolution.

Within the hydrogen atom, the electron and proton possess intrinsic angular momentum or "spin." When the electron and proton spins are aligned in the same direction (parallel configuration, Total angular momentum of the atom  $(\vec{S} + \vec{I}) \vec{F}=1$ ), the atom is in a higher energy state known as the *triplet* state. Conversely, when the spins are anti-parallel, the atom is in a lower energy state called the *singlet* state ( $\vec{F}=0$ ).

The transition between these two states emits radiation at a wavelength of approximately 21 centimetres, corresponding to a frequency of approximately 1420.4 MHz. This transition is known as the *spin-flip* transition, where the electron and proton spin change from being parallel to antiparallel or vice versa, it would cause emission or absorption of the 21-cm radiation.

#### Redshift and Observation

Due to the universe's expansion, the emitted radiation from distant sources undergoes cosmological redshift. As a result, the rest-frame wavelength of 21 cm radiation is stretched as the universe expands, causing it to be observed at longer wavelengths or lower frequencies. By analysing these signals we can learn about the evolution of the neutral hydrogen density in the universe, which gives us various important cosmological insights.



## 2.2 Concept of redshift and how 21-cm signal stretched out:

In cosmology, redshift ( $z$ ) describes the phenomenon where light from distant objects in the universe is shifted to longer wavelengths as the universe expands. This shift also affects the frequency of the light. Mathematically, the redshift of an object is defined as :

$$z = \frac{f_{\text{emitted}} - f_{\text{observed}}}{f_{\text{observed}}}$$

where  $f_{\text{observed}}$  is the frequency of the light observed by the observer, and  $f_{\text{emitted}}$  is the frequency of the light emitted by the source. A redshift of  $z = 0$  corresponds to no change in frequency, indicating that the object is relatively nearby and not moving away from us. A redshift of  $z > 0$  indicates greater distances or early stages of the universe.

If the emitted frequency is  $f_{\text{emitted}} = 1420$  MHz and the redshift is  $z$ , the observed frequency ( $f_{\text{observed}}$ ) can be calculated using the formula for redshift:

$$f_{\text{observed}} = \frac{f_{\text{emitted}}}{1 + z} = \frac{1420}{1 + z} \text{ MHz}$$

So, we will observe 21-cm at lower frequencies.

## 2.3 Concept of visibility and sky image:

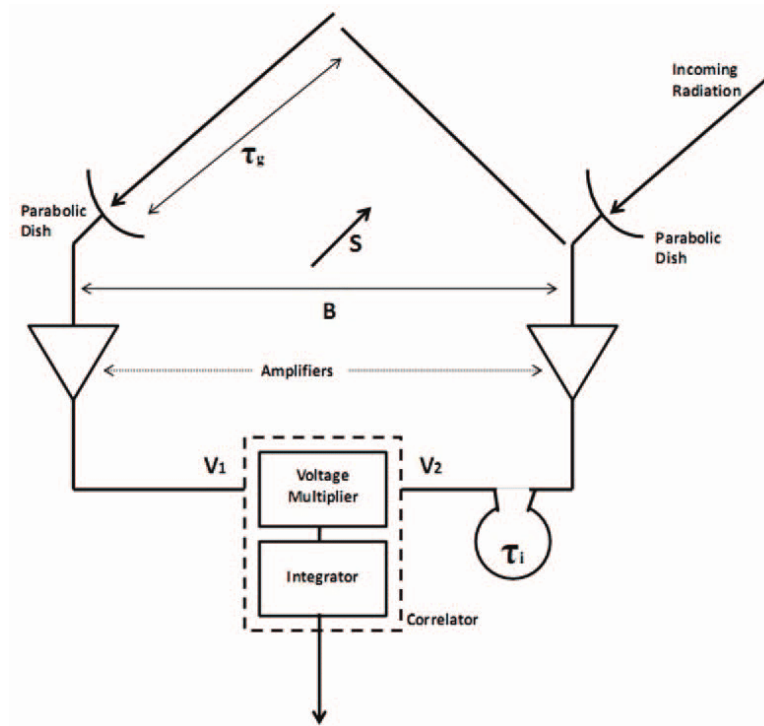


Figure 2.1: two-element interferometer

As shown in Figure 2.1 the signal which is coming from any astrophysical source is collected by the two different parabolic antennas, and then it is converted into a voltage signal. Thereafter, the

two voltage signals are passing through the correlator. The correlator has two parts voltage multiplier used for multiplying the signals and an integrator used to integrate the product of the signals, these two together cross-correlate the signal. There is a time delay or the path difference of the signal. These cross-correlations of two different signals give the value of visibility in the uv-plane.

## 2.4 Sky image:

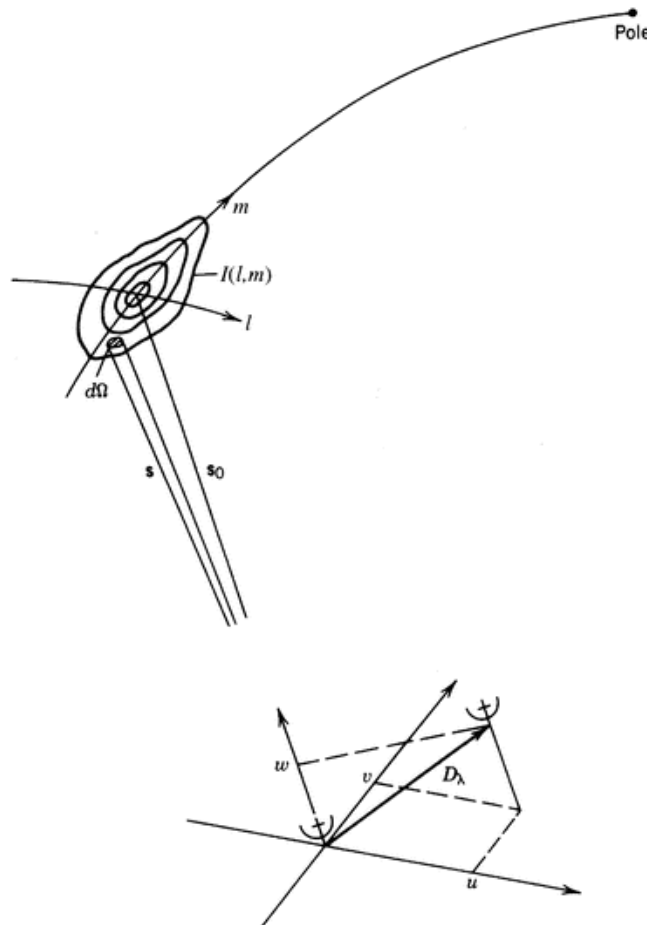


Figure 2.2: relation sky image plane and the  $uv$ -plane(the picture taken from "Interferometry and the synthesis in radio astronomy by A.Thompson and M.Moran"[2])

Sky image is the Fourier transform of the visibility and vice-versa.

$$V(u, v) = \iint I(l, m) e^{-i2\pi[l u + m v]} dl dm$$

Where:  $u$  and  $v$  are the coordinates in the  $uv$ -plane(baseline coordinate) and  $l$  and  $m$  are the angular coordinates in the sky image plane.  $I(l, m)$  is the sky brightness distribution.

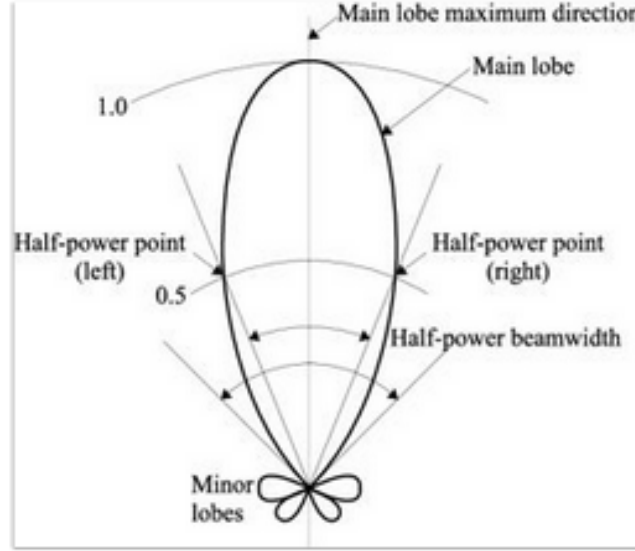


Figure 2.3: station beam pattern

## 2.5 introduced primary beam pattern in the visibility expression:

In the figure-2.3, we can see one normalized station beam pattern, it consists of one primary beam (main lobe) and a few secondary beams (minor lobes). Incorporating the primary beam pattern into the expression accounts for the sensitivity pattern of the antenna to radiation from different directions in the sky. This ensures that only the radiation from the desired direction contributes to the measured visibility. The Primary beam is used to trace out a particular portion of the sky.

### Visibility with Primary Beam Pattern

The visibility function  $V(u, v)$  with the primary beam pattern incorporated is given by:

$$V(u, v) = \iint I(l, m) \cdot B(l, m) \cdot e^{-i2\pi[l u + m v]} dl dm$$

Where:  $V(u, v)$  is the visibility function,  $I(l, m)$  is the sky brightness distribution,  $B(l, m)$  is the primary beam pattern (response function of the antenna to incoming radiation),  $u$  and  $v$  are baseline co-ordinates and  $l$  and  $m$  are direction cosines corresponding to the angular coordinates in the sky image plane.

## 2.6 TGE(tapered gridded estimator):

Point sources near these nulls and sidelobes can create challenges for estimating the angular power spectrum of diffuse background radiation emission. It would also complicate the foreground separation from the cosmological 21-cm signal. Fourier transform of all individual visibility points is computationally very expensive. These two challenges, TGE can easily incorporate.

## Tapering and grinding of visibility:

The Gaussian window function is defined as:

$$W(\theta) = e^{-\frac{\theta^2}{\theta_w^2}}$$

where  $\theta_w = f\theta_0$  with  $f \leq 1$  to ensure cutoff before the first null.

$$V_c(U) = \tilde{w}(U) \otimes V(U)$$

The tapering is achieved by convolving the measured visibilities  $V(U)$  with the Fourier transform of the window function  $\tilde{w}(U)$ .

$$\tilde{w}(U) = \mathcal{F}\{W(\theta)\} = \pi\theta_w^2 e^{-\pi^2 U^2 \theta_w^2}$$

The convolved visibilities  $V_c(U)$  are then the Fourier transform of the product  $W(\theta)A(\theta)\delta I(\theta)$ , allowing for better control of the sky response through the window function as mentioned in [Choudhuri et al.](#)

The measured visibilities  $V_i$  decompose into sky signal  $S(U_i)$  and system noise  $N_i$ .

$$V_i = S(U_i) + N_i$$

Tapering is achieved by convolving visibilities with the Fourier transform of the window function and  $U_g$  are the different grid points in the uv-plane.

$$V_{cg} = \sum_i \tilde{w}(U_g - U_i) V_i$$

The estimator for the angular power spectrum is defined:

$$\hat{E}_g = M_g^{-1} \left( |V_{cg}|^2 - \sum_i |\tilde{w}(U_g - U_i)|^2 |V_i|^2 \right)$$

$M_g$ , which is the normalization constant for the TGE. The values of  $M_g$  depend on the baseline distribution and the form of the tapering function  $W(\theta)$  as mentioned in [Choudhuri et al.](#)

## 2.7 Angular power spectrum:

The angular power spectrum  $C_\ell$  can be realized by the ensemble average of the two-point correlation function of brightness temperature fluctuations  $\tilde{T}$ .

$$\langle \tilde{T}(U) \tilde{T}^*(U') \rangle = \delta_D^2(U - U') C_{2\pi U}$$

Where,  $\tilde{T}(U)$  and  $\tilde{T}^*(U')$  are the brightness temperature fluctuations at two different baselines.  $\delta_D^2(U - U')$  is a 2D Dirac delta function ensuring that the ensemble average is non-zero only when  $U$  is equal to  $U'$ .  $C_{2\pi U}$  is the angular power spectrum, where  $2\pi U = \ell$  is the angular multipole.

Averaging over multiple realizations of the estimator gives the angular power spectrum.

$$\langle \hat{E}_g \rangle = C_{\ell_g}$$

Here,  $\ell_g = 2\pi U_g$  is the angular multipole, 'g' refers to grid points, given in [Choudhuri et al.](#)

## 2.8 Cylindrical power spectrum:

The Cylindrical power spectrum is the Fourier transform of  $C_\ell(\Delta\nu)$ , under flat sky approximation [Bharadwaj et al.](#), the expression is shown below,

$$P(k_\perp, k_\parallel) = r^2 r' \int_{-\infty}^{\infty} d(\Delta\nu) e^{-ik_\parallel r' \Delta\nu} C_\ell(\Delta\nu)$$

Where,  $P(k_\perp, k_\parallel)$  is the 2D cylindrical power spectrum,  $k_\perp$  and  $k_\parallel$  are the perpendicular and parallel wavenumbers, respectively and which correspond to perpendicular to the line of sight and parallel to the line of sight, the comoving distance  $r$ , its frequency derivative  $r'$ ,  $\Delta\nu$  is the frequency separation, and  $C_\ell(\Delta\nu)$  is the power spectrum of brightness temperature fluctuations,

# Chapter 3

## Data specification:

### 3.1 SKA Data Specification

In this section, we provide an overview of the data used in our study. The test data includes 150 uvfits test files provided to validate the accuracy of the power spectrum estimator. The main dataset consists of 900 uvfits files with a total size of 7.5TB. Additionally, there is a Station Beam Image Fits File which provides information about the station beam, including a field of view of  $5^\circ \times 5^\circ$  in the sky at Right Ascension (RA), Declination (Dec) = 0h,  $-30^\circ$ . The observational parameters are as follows: The frequency coverage ranges from 106MHz to 196MHz, corresponding to redshift  $z = 12.39$  to  $z = 6.24$ . The channel width is 100kHz with 900 channels. The observation time is 4 hours with an integration time of 10 seconds.

### 3.2 Telescope specification:

#### Antenna and station design:

The basic telescope model makes use of the SKA-Low configuration of 512 stations. The station layout is the so-called “Vogel” layout, a one-arm spiral configuration with a uniform areal density of antennas and a maximally diverse azimuthal sampling. A single station layout has been adopted to reduce computational expense, rather than a full set of 512 diverse station layouts. Each station contains 256 dual linearly polarized antennas, so a total of 131,072 antennas will be configured with a dense central core of antennas, surrounded by three spiral arms. The longest distance (or baseline) between antennas will be 65km. Long baselines mean a higher resolution.

Figure 3.1 shows the station pattern for 200, 350 and 512 of the antenna station and as we increase the station numbers, we can see the formation of three spiral arms.

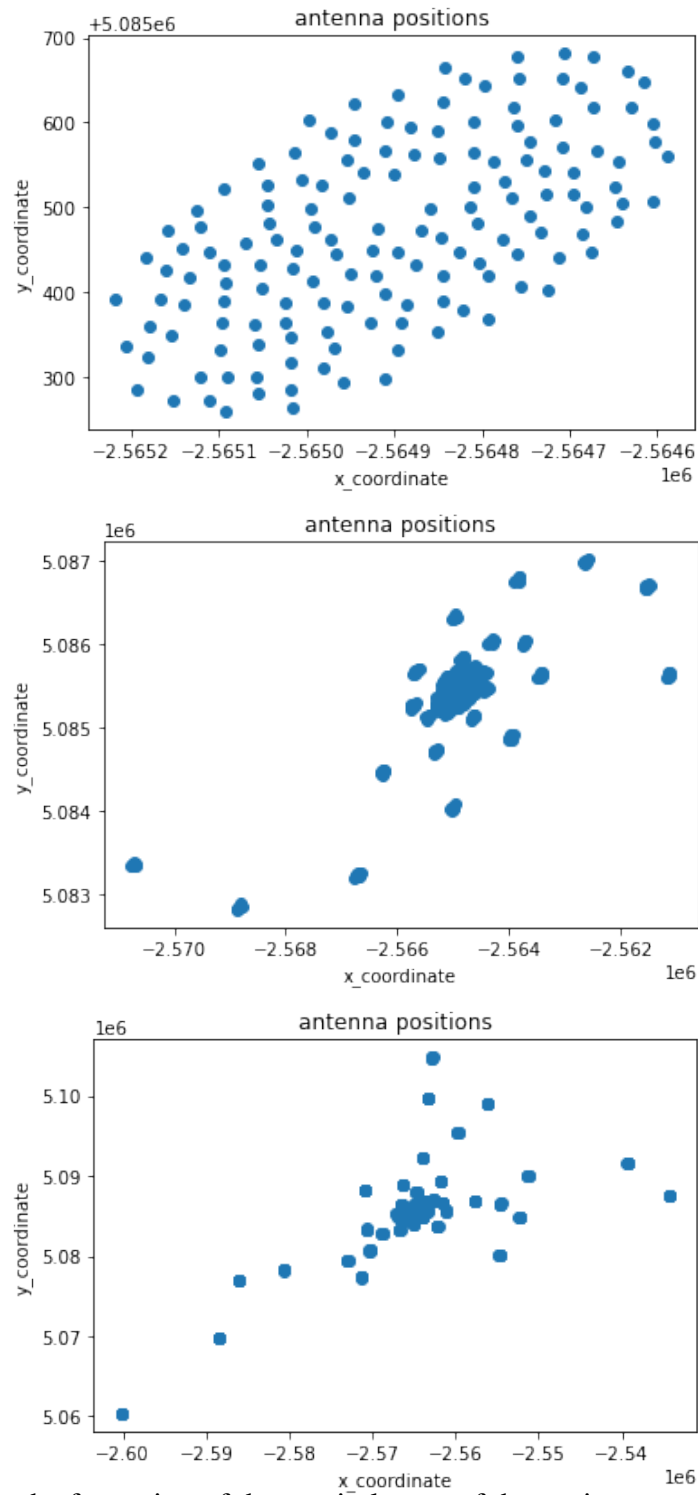


Figure 3.1: the formation of three spiral arms of the station-antenna pattern.

To conduct extensive sky surveys effectively, a telescope needs a broad field of view to observe large sky regions efficiently. SKA-Low employs a method known as digital beamforming instead of physically relocating the antennas.

This technique digitally adjusts the telescope's primary beam "pointing" direction by combining signals from multiple antennas within the array. The telescope achieves an extensive field of view by generating numerous beams simultaneously, each pointing to various sky locations. This approach significantly enhances survey speeds compared to traditional telescopes

### 3.2.1 Sensitivity of SKA1-LOW:

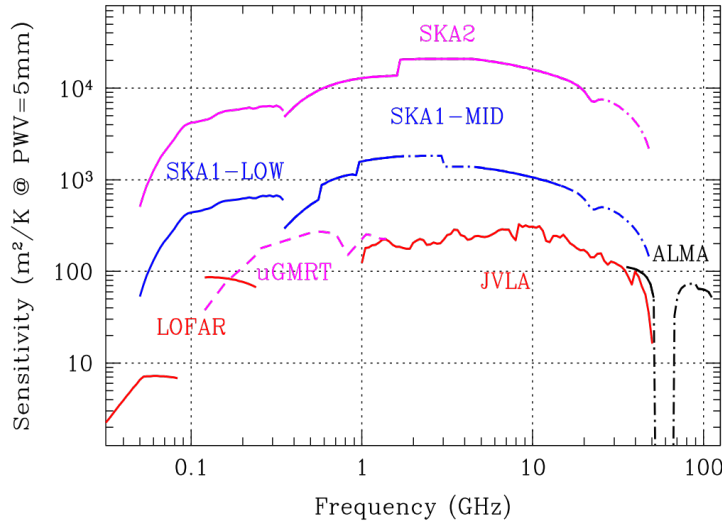


Figure 3.2: Sensitivity of SKA1-LOW ( This figure is taken from [Braun et al.](#))

SKA-Low is sensitive to low-frequency radio signals 50 MHz to 350 MHz ( $z=27-3$ ). The sensitivity of a radio telescope depends on the collecting area available to capture signals and the system temperature  $T_{\text{sys}}$ . SKA-Low provides 400,000m<sup>2</sup> of collecting area. From the above figure, we can see the sensitivity of SKA1-LOW is much higher than LOFAR and uGMRT for lower frequency regions for PWV (precipitable water vapour)=5mm.

### 3.2.2 Primary beam pattern for some specific channels:

By analyzing the primary beam fits file, we observe that the primary beam gets narrower as the channel number increases. So, for higher frequency the resolution increases ( $\theta \approx \frac{\lambda}{D}$ ). It is noticeable that as the channel number increases, the side lobes rise slowly.



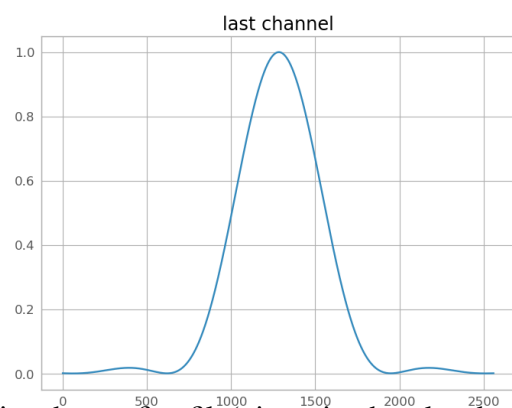
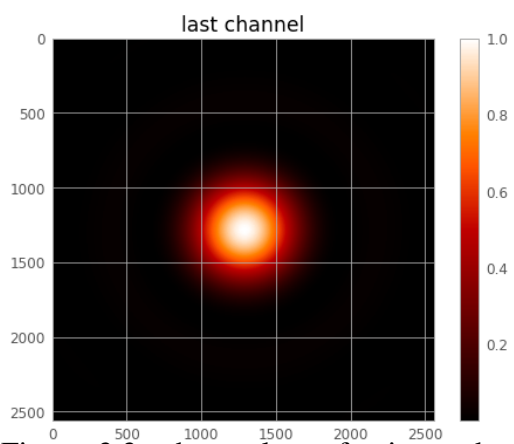
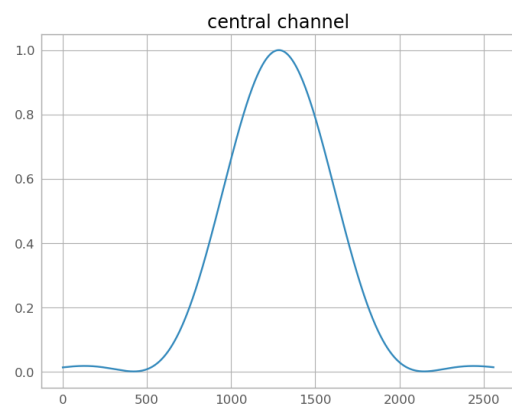
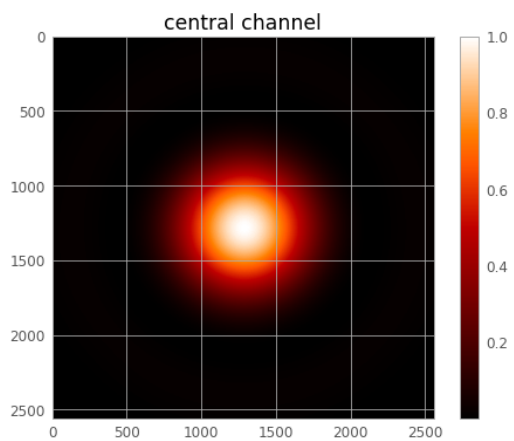
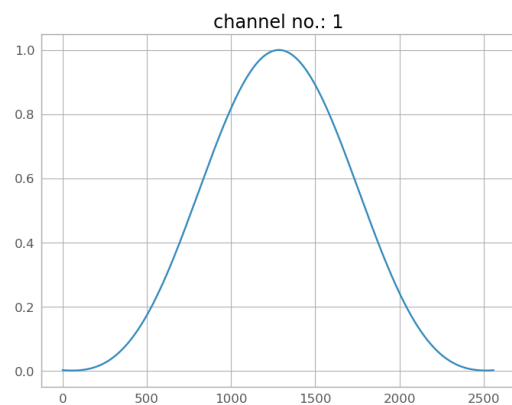
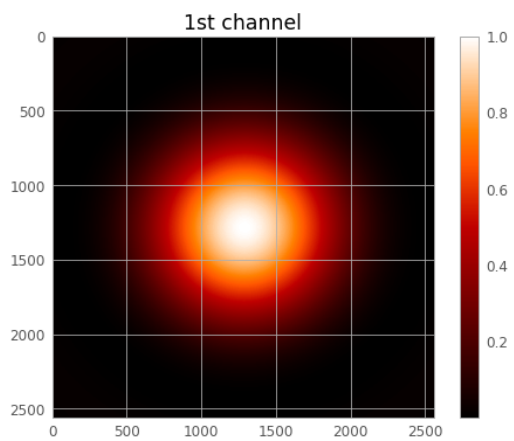


Figure 3.3: these plots of primary beam is from the station beam fits file(given in the ska data challenge)

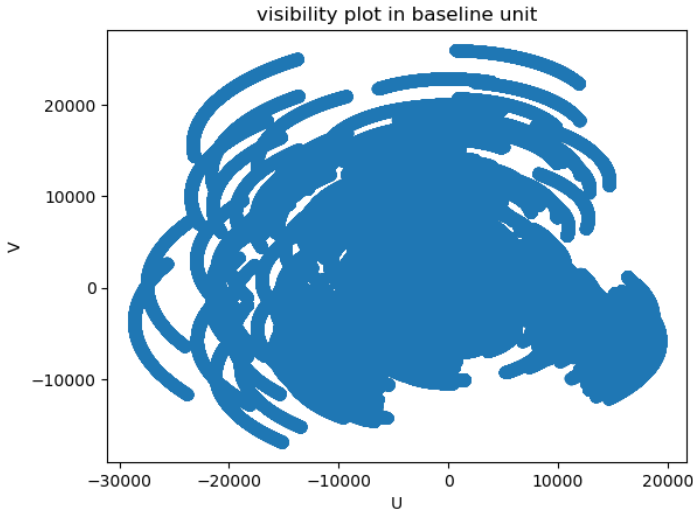
# Chapter 4

## Data Processing Steps and Results

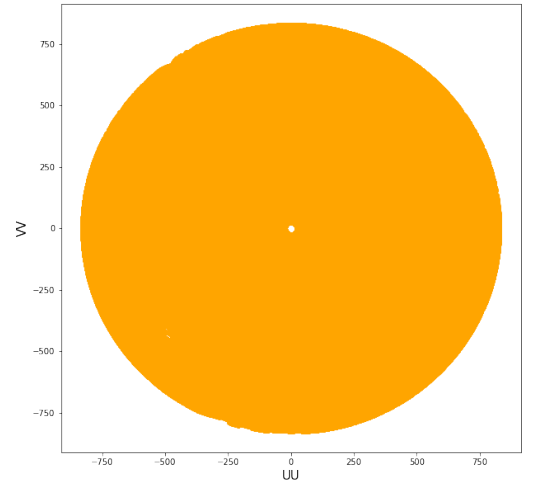
### 4.1 Data Reduction:

#### Extracting 1.5 km region:

We had initially acquired 150 UVfits test files for the frequency range 166 to 181 MHz, each file being 8.5 GB in size. To facilitate data management, we extracted a 1.5 km region (830 in wavelength unit) from the original visibility plane using the "split" task in the CASA software application, resulting in a reduction of each file's size to 3.5 GB.



(a) Total uv coverage



(b) Extracted uv coverage

#### Perform Averaging:

Due to storage limitations, we further reduced the file size by averaging over 120 seconds using the "split" task in CASA. This averaging process resulted in a file size reduction to 272 MB per file. The averaging operation required approximately 7 hours and 41 minutes for each file on our local server.

### **Perform concatenation:**

To address insufficient RAM issues, we transitioned our operations to the SKA server, where we concatenated the 150 .ms files (ms stands for measurement set) into a single file, resulting in a file size of 8.1 GB.

### **Perform ms transform:**

Subsequently, we performed a ms transform to break down the spectral windows between the channels and converted them into a single UVfits file format (using the CASA task "ms transform").

### **Perform the same for true data:**

We attempted to perform the same processing steps for the true dataset; however, the averaging process for each file took approximately 35 hours and 55 minutes. To optimize time efficiency, we parallelized the execution by creating multiple script files and ran them concurrently. As a result, we successfully combined the first 150 files out of the total 900 UVfits files.

### **Extracting 1st 150 Station Primary beam:**

We have successfully extracted out 1st 150 primary beams from the station beam fits file using CASA for 106 to 121 MHz frequency range for later usage.

### **Further plan for processing:**

Our plan was to create 6 frequency bands by combining the 900 UVfits files, with each band containing 150 UVfits files. Currently, we have successfully created one band spanning the frequency range from 106 to 121 MHz, resulting in a file size of 8.1 GB.

## **4.2 Processing the Combined UVfits file:**

### **Creating 10 realizations:**

We create 10 realizations for the test and the true data by simply copying the actual data into 10-fits files.

### **Applying ngrf and visgridfits:**

These two code is required for Mg calculation. ngrf code is designed for generating simulated diffuse images with specified characteristics, such as size, randomness, and frequency properties, and saving them in FITS format for further analysis or visualization. Here, we have created 10 Mg\_r1.fits to Mg\_r10.fits files for using in visgridfits code.

This program essentially takes a multichannel FITS image and a UVFITS template, performs interpolation of gridded visibilities from the image to the UV baselines, and updates the UVFITS data with the interpolated visibilities. It involves FFT operations, data reading/writing, and interpolation techniques.

### Applying TGE for data and the realizations:

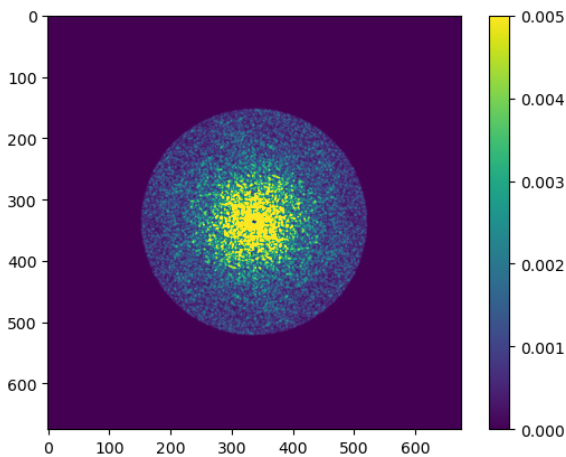
We applied the Tapered Gridded Estimator (TGE) to the Combined test data and the 1st band. The inputs used for the estimation are as follows: Number of cells ( $nc$ ) was set to 151, Maximum baseline length ( $U_{\max}$ ) was set to 1500.0, Full Width at Half Maximum (FWHM) was measured at 266.2 arcmin, Tapering factor ( $f$ ) was set to 0.8, Grid spacing ( $dU$ ) was 4.57 wavelengths, and the total number of baselines within the uv range was determined to be 4668308. The size of the output Gridded Visibilities (GV) is 675 x 675 x 151. Furthermore, the TGE produces one binary file, which we have read using Jupyter Notebook, and plotted the phase and the amplitude of the visibilities.

### Applying readgv.py:

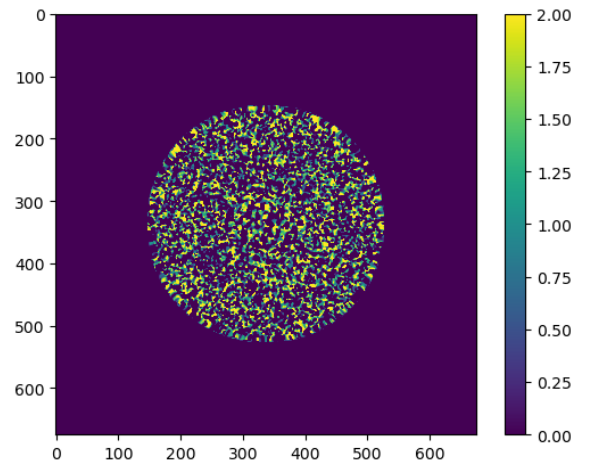
We apply readgv.py to read the output binary files of the TGE code. There is only XX polarization in our dataset and we successfully extracted the visibility and visibility square term from the binary files of the data and the realizations and saved it successfully to compute the angular power spectrum.

### Phase and Amplitude of the gridded visibility plots for test data (for channel no. 10):

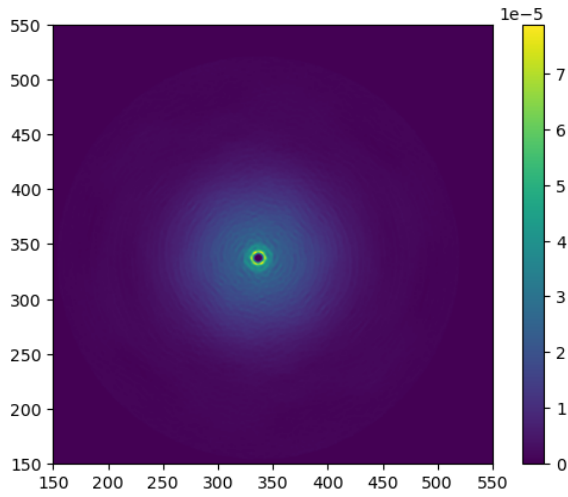
Visibility is a complex quantity. So, we can decompose it into amplitude and phase parts. visibility square is real quantity. So, it has only amplitude, it doesn't have phase.



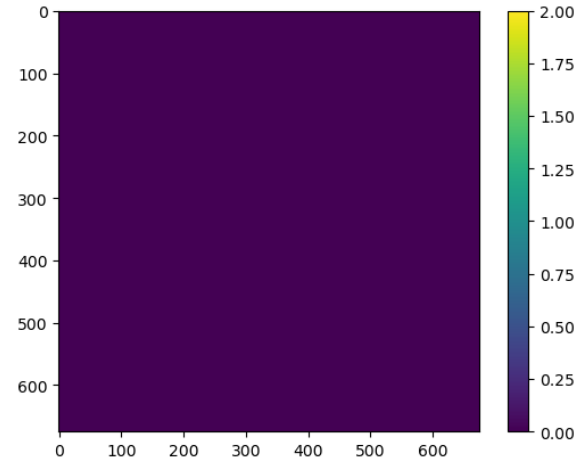
(a) Amplitude of Visibility (here, the plots are shown in uv (baseline) coordinates)



(b) Phase of Visibility



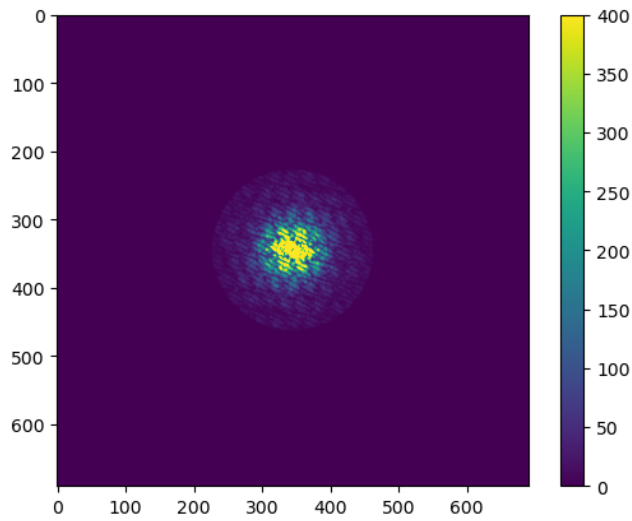
(a) Amplitude of Visibility square (here, the plots are shown in uv (baseline) coordinates)



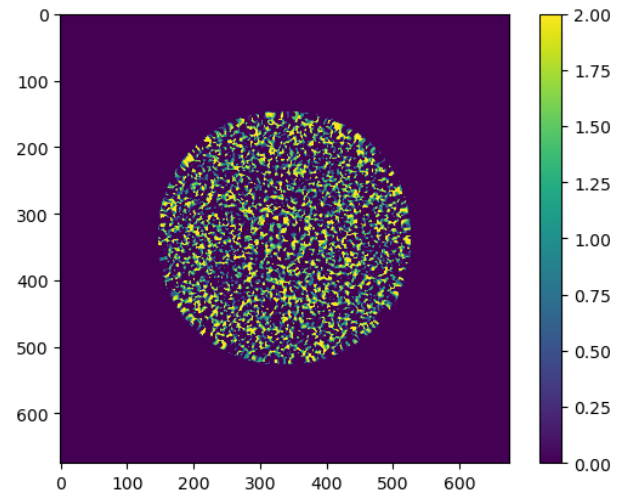
(b) Phase of Visibility square

Figure 4.3: Visualization of Visibility Properties

### Phase and Amplitude of the gridded visibility plots for true data:



(a) Amplitude of Visibility (here, the plots are shown in uv (baseline) coordinates)



(b) Phase of Visibility square

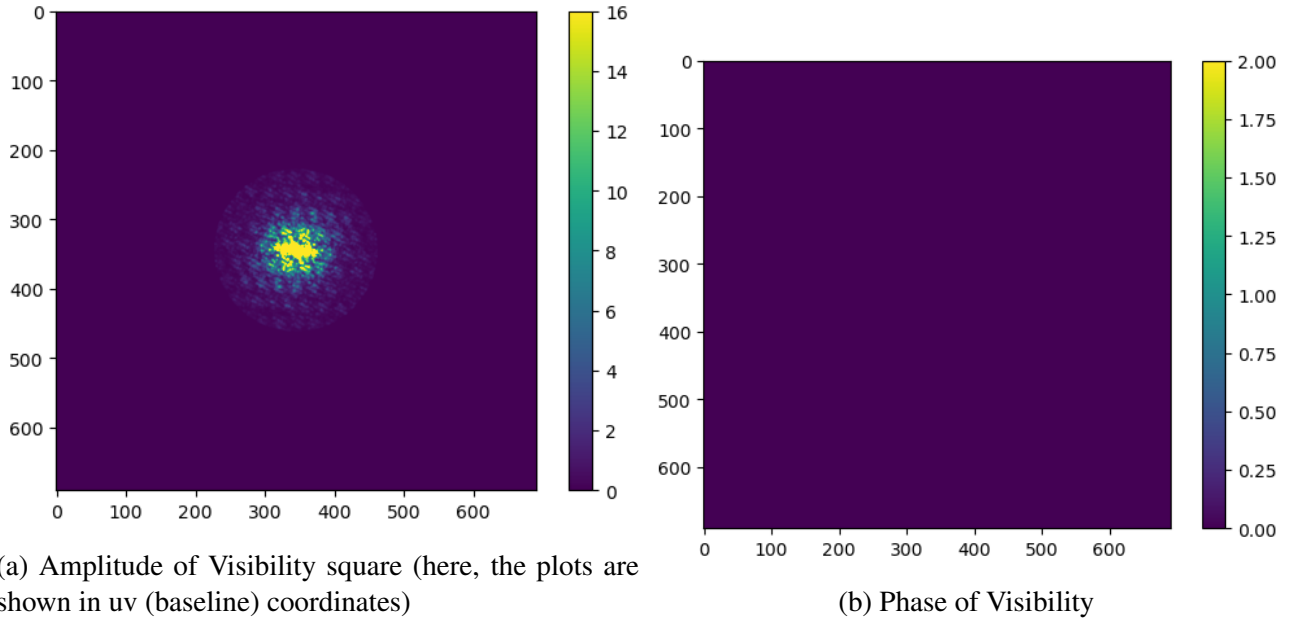


Figure 4.5: Visualization of convolved tapered gridded Visibility Properties (both the axes are in baseline unit, The uv Plane is shown here) )

#### 4.2.1 Angular power spectrum $cl(\Delta\nu)$ :

$\Delta\nu$  is 0.0 to 15 MHz the interval between them is 0.1 MHz.

##### For test data:

Here, we have shown the angular power spectrum of test data for six angular modes ( $\ell = 724, 1123, 1529, 1939, 2344, 2761$ ).

We observe that the 21-cm signal is localized within a small range of frequency separations. As the frequency separation increases, the signal declines rapidly. The test data for validation is free from noise and foreground, so we expect a smooth nature of the  $C_\ell$  as depicted in the figure. Additionally, as the angular mode  $\ell$  increases, the correlation decreases rapidly from Figure 4.6, which is noticeable. Thus, for higher angular modes, detection of the 21 cm signal is less probable. Similar properties are discussed in the paper [Bharadwaj et al.](#). Furthermore, some unknown strange behaviour is observed for higher frequency separations, where the correlation diverges, as shown in Figure 4.7.

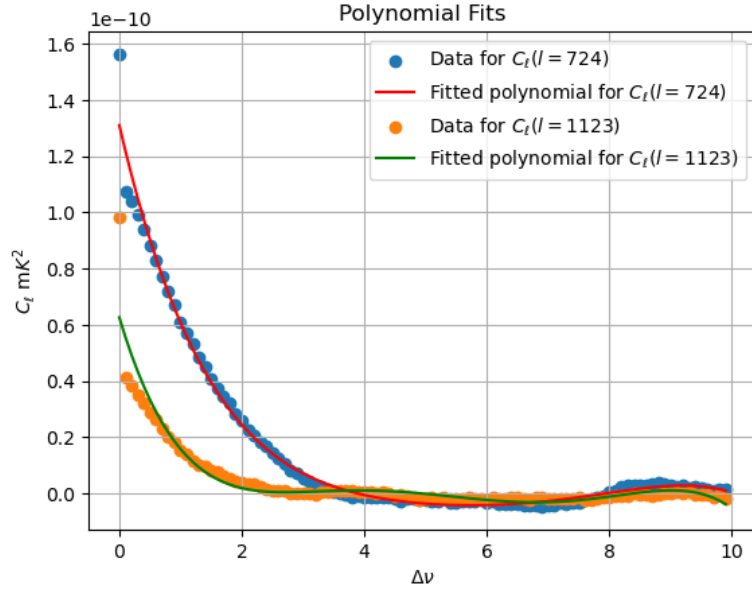


Figure 4.6: plot of  $cl(\Delta\nu)$  for two different values of  $\ell$  ( $\Delta\nu$  is in MHz unit) and we can see a de-correlation as frequency separation increases.

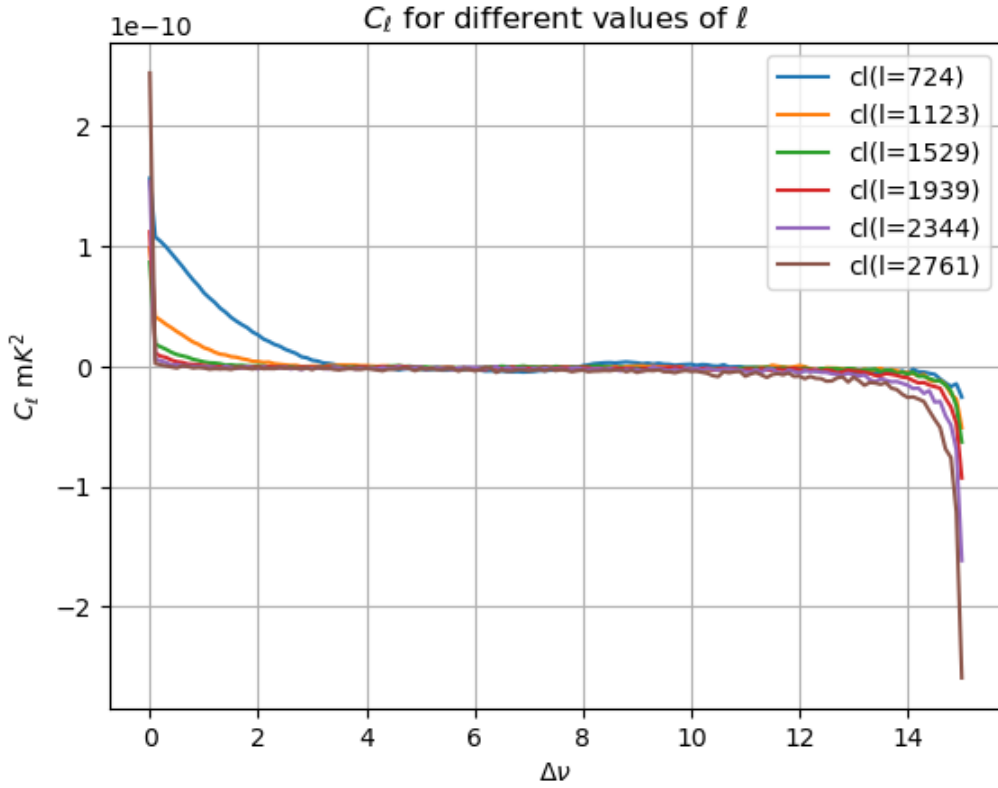


Figure 4.7: plot of  $cl(\Delta\nu)$  for different value of  $\ell$  ( $\Delta\nu$  is in MHz unit). Here, as the frequency separation increases beyond 10 MHz the fluctuation in  $cl(\Delta\nu)$  increases (this behaviour is unknown).

**For true data:**

below we have shown the angular power spectrum of test data for six angular modes ( $\ell = 703, 1102, 1502, 1916, 2357, 2771$ ).

True data was dominated by noise and foregrounds as the amplitude of  $C_\ell$  is much higher than the test data and  $C_\ell$  doesn't de-correlate as frequency separation increases, unlike the test data.

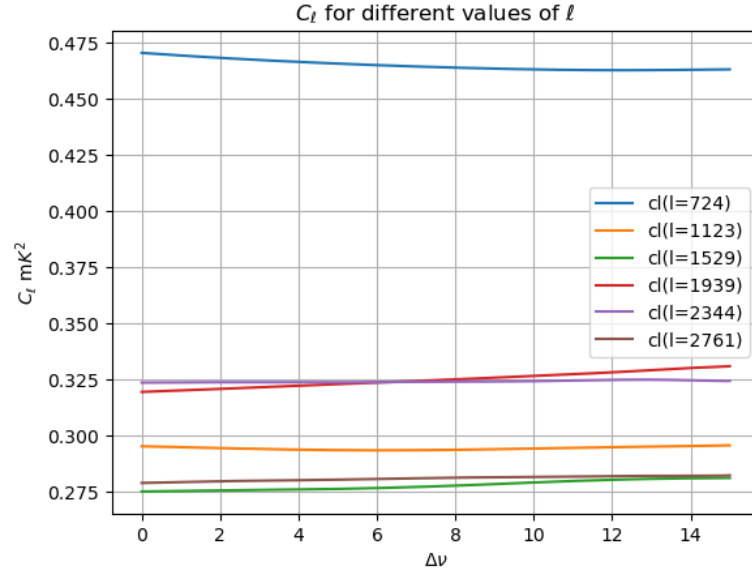


Figure 4.8: plot of  $cl(\Delta\nu)$  for different value of  $\ell$  for the 1st band of true data ( $\Delta\nu$  is in MHz unit). Here, there is no de-correlation, this suggests that the data is noise-dominated.



### 4.2.2 Estimating cylindrical power spectrum of the test data for the validation of the TGE estimator:

We have the following numerical values:  $r = 9.448626 \times 10^3$  Mpc,  $r' = 1.773767 \times 10^1$  Mpc/MHz, and  $\Delta\nu = 0.1$  MHz. We partition the data into 6 bins of linearly spaced intervals along  $k_\perp$ , each containing  $N_c = 151$  values along  $k_\parallel$ . Additionally, we subdivide each bin along  $k_\parallel$  into 10 equal linear bins. The  $k_\perp$  bin values of the model test data are 0.05, 0.1, 0.15, 0.2, 0.25, and 0.3, while the estimated  $k_\perp$  bin values of the test data are 0.0744, 0.1166, 0.1587, 0.2028, 0.2495, and 0.2933. Similarly, the  $k_\parallel$  bin values of the model test data are 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.40, 0.45, and 0.5, while the estimated  $k_\parallel$  bin values of the test data are 0.0531, 0.1003, 0.1476, 0.2007, 0.2539, 0.3011, 0.3483, 0.3956, 0.4487, and 0.5018.

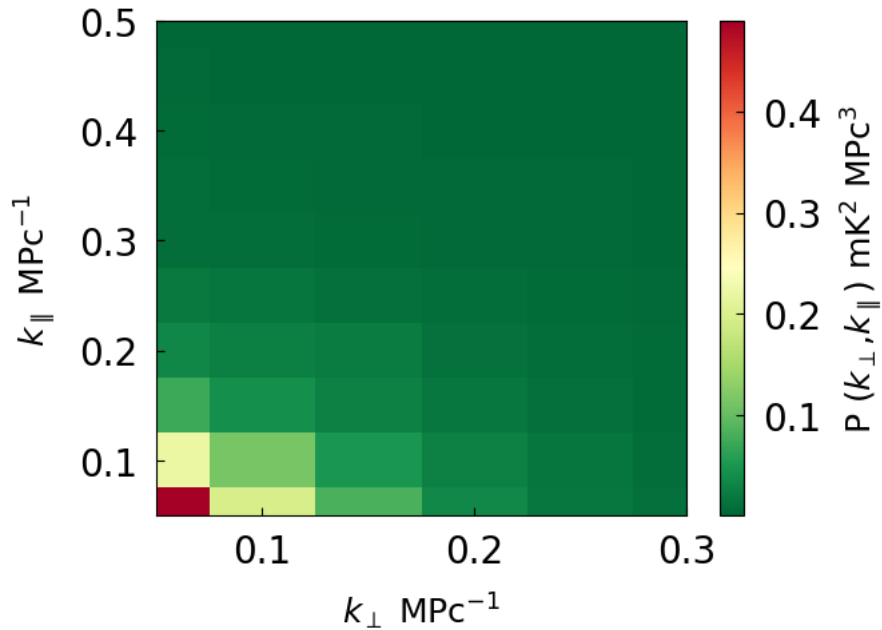


Figure 4.9: Given cylindrical power spectrum for the validation.  $k_\perp$  bin values of the model test data are 0.05, 0.1, 0.15, 0.2, 0.25, and 0.3. for  $k_\parallel$  bin values of the model test data are 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.40, 0.45, and 0.5. The colour bar shows the magnitude of the  $P(k_\perp, k_\parallel)$  and it is in the millikelvin square unit (as sky brightness temperature is measured in millikelvin).

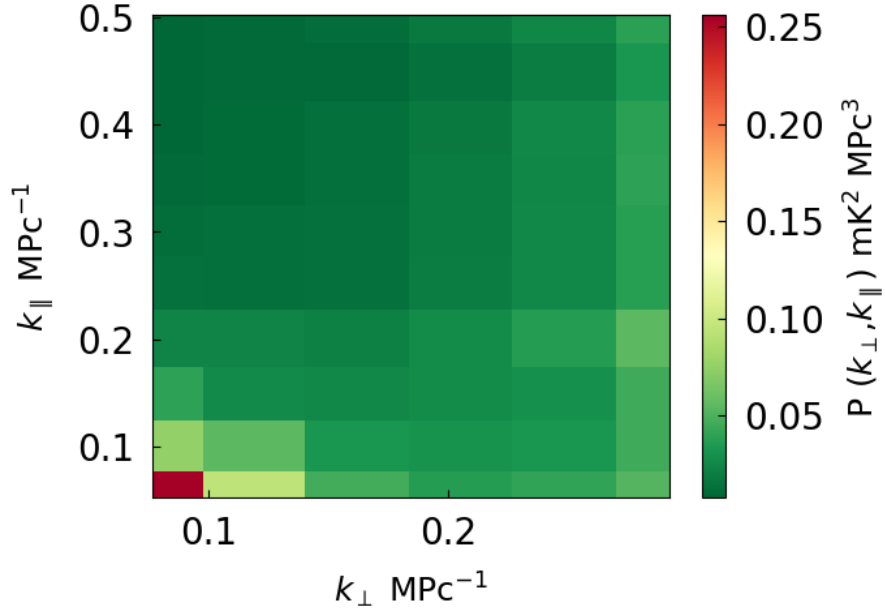


Figure 4.10: Generation of cylindrical power spectrum from the test data. the estimated  $k_{\perp}$  bin values of the test data are 0.0744, 0.1166, 0.1587, 0.2028, 0.2495, and 0.2933 and estimated  $k_{\parallel}$  bin values of the test data are 0.0531, 0.1003, 0.1476, 0.2007, 0.2539, 0.3011, 0.3483, 0.3956, 0.4487, and 0.5018. The colour bar shows the magnitude of the  $P(k_{\perp}, k_{\parallel})$

### 4.2.3 Percentage deviation of the generated Power Spectrum from the model power spectrum:

For the higher values of k-parallel and k-perpendicular, the percentage deviation is much higher. For lower k-parallel and k-perpendicular values, we have successfully recovered the input Cylindrical power spectrum with very less deviation.

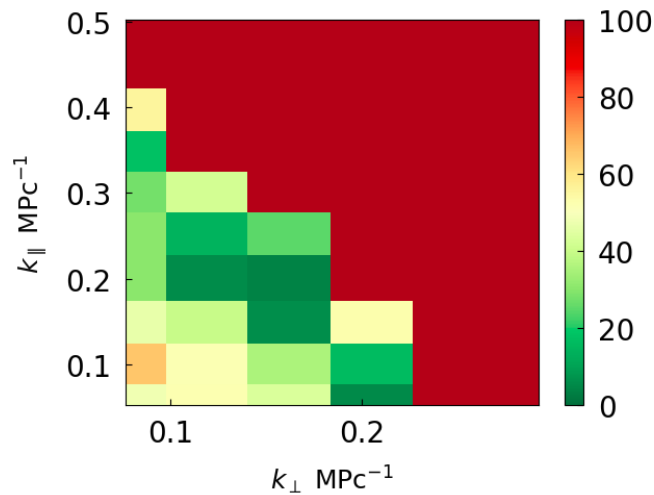


Figure 4.11: percentage deviation (a comparison between estimated power spectrum and input power spectrum). The colour bar shows the intensity of the deviation.

#### 4.2.4 Cylindrical power spectrum for true data:

- estimated  $k_{\perp}$  bin values of test data = (0.0744, 0.1166, 0.1590, 0.2028, 0.2495, 0.2933)
- estimated  $k_{\parallel}$  bin values of test data = (0.0531, 0.1004, 0.1476, 0.2007, 0.2539, 0.3011, 0.3483, 0.3956, 0.4487, 0.5018)
- This is the cylindrical power spectrum without a reduction of foregrounds and noise.

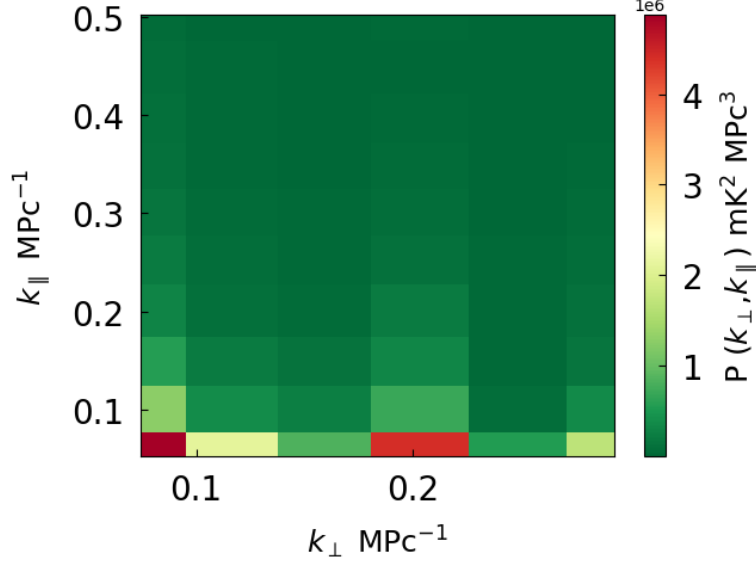


Figure 4.12: Generation of cylindrical power spectrum from the true data. The  $P(k_{\perp}, k_{\parallel})$  is noise-dominated.

### 4.3 Conclusion:

From the Angular Power spectrum,  $C_{\ell}$  we can see the 21-cm signal is localized within a small range of frequency separations. As the frequency separation increases, the signal declines rapidly and also as the angular mode increases, the correlation decreases rapidly. For lower  $k_{\parallel}$  and  $k_{\perp}$  values, we have successfully recovered the input Cylindrical power spectrum. But  $C_{\ell}$  for test data is noise-dominated, As a future work we have to eliminate the noise to get the correct power spectrum for further analysis.

# Bibliography

- [1] Jonathan R Pritchard and Abraham Loeb. 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901, July 2012. ISSN 1361-6633. doi: 10.1088/0034-4885/75/8/086901. URL <http://dx.doi.org/10.1088/0034-4885/75/8/086901>.
- [2] A. Richard Thompson and M. T. Moran. *Interferometry and Synthesis in Radio Astronomy*. Wiley, 1986.
- [3] Samir Choudhuri, Somnath Bharadwaj, Abhik Ghosh, and Sk. Saiyad Ali. Visibility-based angular power spectrum estimation in low-frequency radio interferometric observations. *Monthly Notices of the Royal Astronomical Society*, 445(4):4351–4365, November 2014. ISSN 0035-8711. doi: 10.1093/mnras/stu2027. URL <http://dx.doi.org/10.1093/mnras/stu2027>.
- [4] Samir Choudhuri, Somnath Bharadwaj, Suman Chatterjee, Sk. Saiyad Ali, Nirupam Roy, and Abhik Ghosh. The visibility-based tapered gridded estimator (tge) for the redshifted 21-cm power spectrum. *Monthly Notices of the Royal Astronomical Society*, 463(4):4093–4107, October 2016. ISSN 1365-2966. doi: 10.1093/mnras/stw2254. URL <http://dx.doi.org/10.1093/mnras/stw2254>.
- [5] Somnath Bharadwaj, Srijita Pal, Samir Choudhuri, and Prasun Dutta. A tapered gridded estimator (tge) for the multifrequency angular power spectrum (maps) and the cosmological hi 21-cm power spectrum. *Monthly Notices of the Royal Astronomical Society*, 483(4):5694–5700, December 2018. ISSN 1365-2966. doi: 10.1093/mnras/sty3501. URL <http://dx.doi.org/10.1093/mnras/sty3501>.
- [6] Robert Braun, Anna Bonaldi, Tyler Bourke, Evan Keane, and Jeff Wagg. Anticipated performance of the square kilometre array – phase 1 (ska1), 2019.