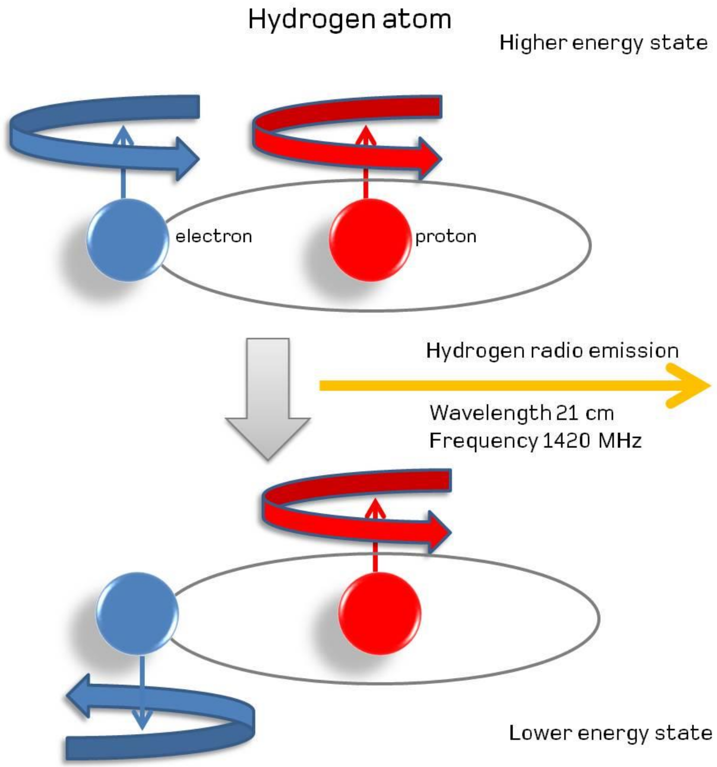
SKA Data Challenge 3

Guided by Dr. Samir Choudhuri

Santanu Das (PH22C040)

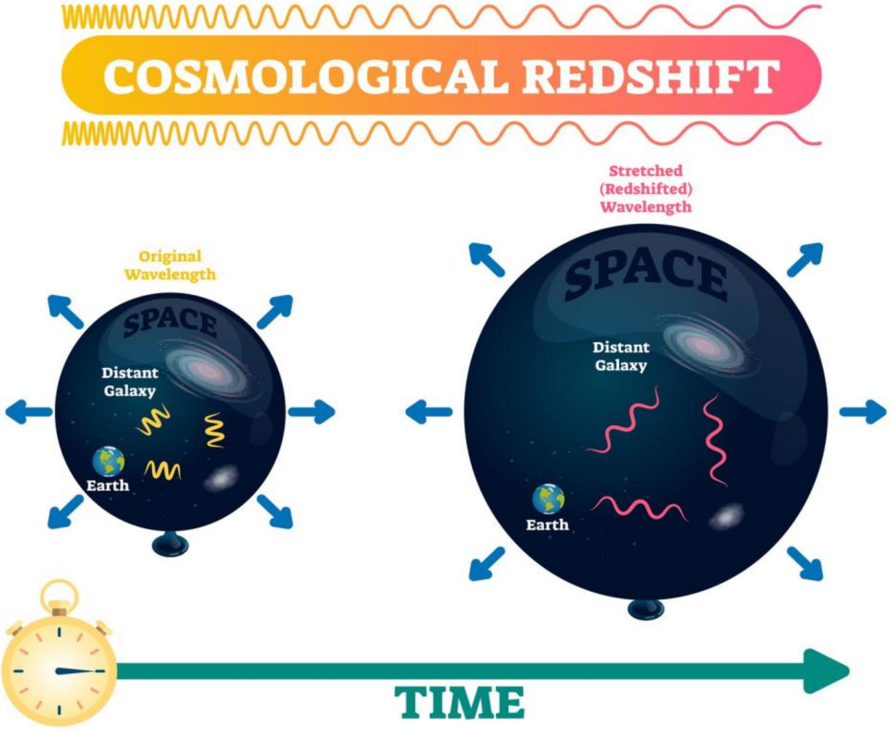
21 –cm signal:



Source:https://images.app.goo.gl/1RWmYuFvVaXQkN8i8

- 75% of all baryons is hydrogen

Concept of redshift and measuring the 21-cm signal at low frequency



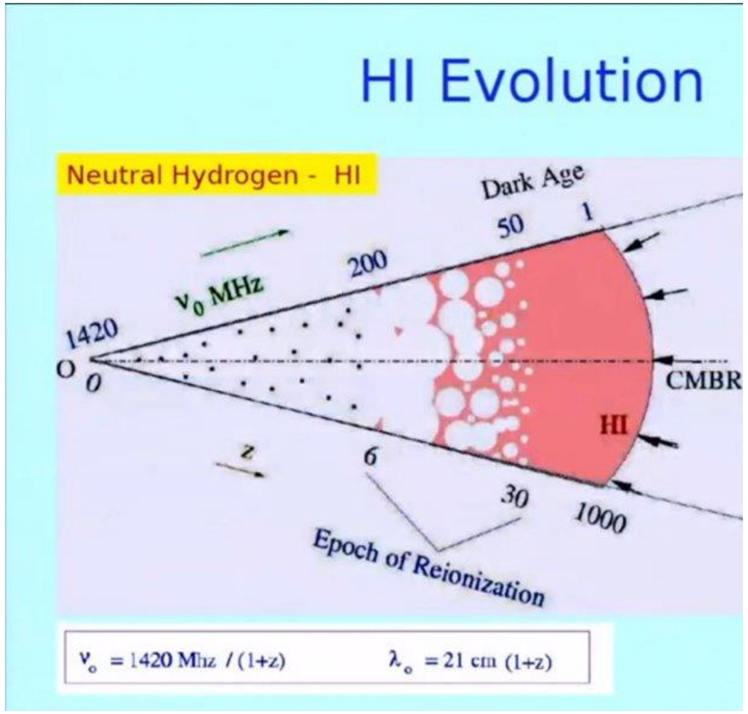
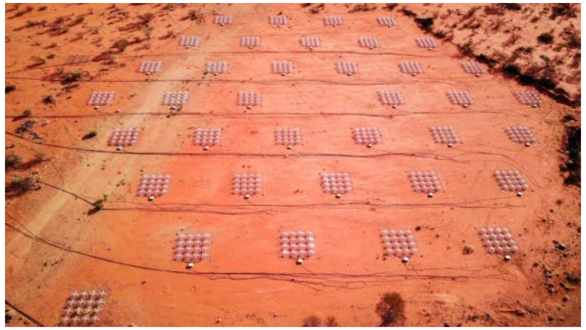


Image Source: https://www.thoughtco.com/what-is-redshift-3072290
https://youtu.be/CbJLJ5FB08k?si=8Lm_xXkH3U7SsJoY

Radio telescopes:(for detecting Redshifted 21-cm signal)

GMRT

LOFAR



MWA Telescope

SKA-mid

SKA-low

Source: Google

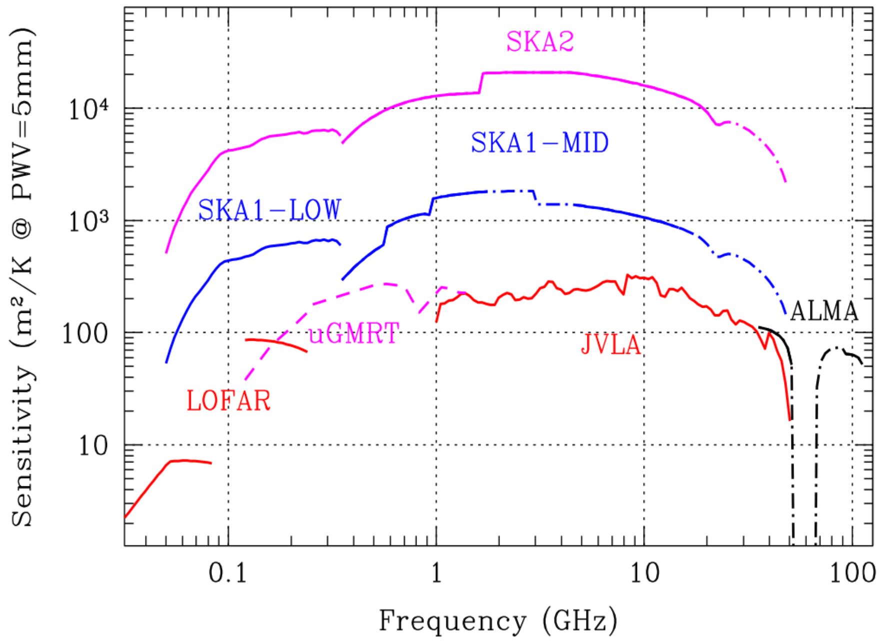
SKA1-LOW: Square kilo Meter Array

- **it is situated in the western part of Australia.**

- 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. SKA-Low provides **400,000m^2 of collecting area.**

$$S = \frac{A_{\rm eff}}{T_{\rm sys}}$$



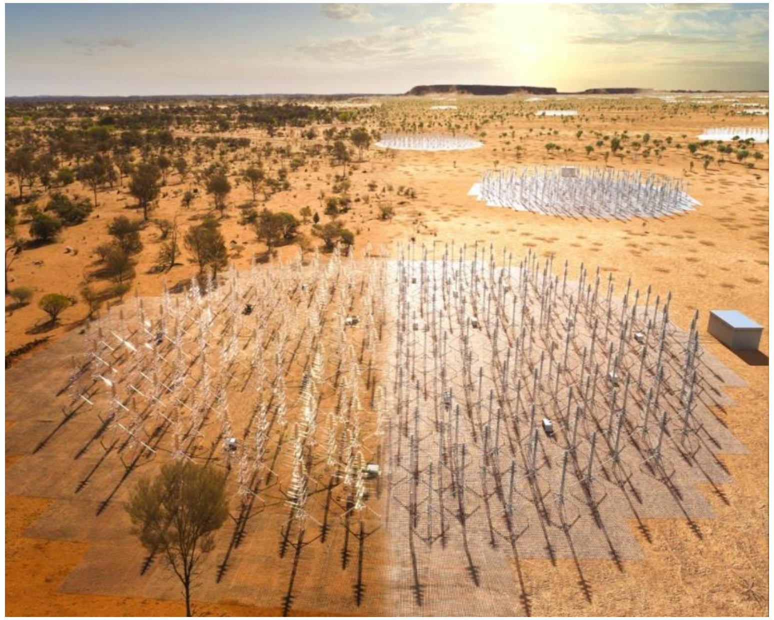
Source:arxiv.org/pdf/1912.12699

SKA1-LOW: Square kilo Meter Array

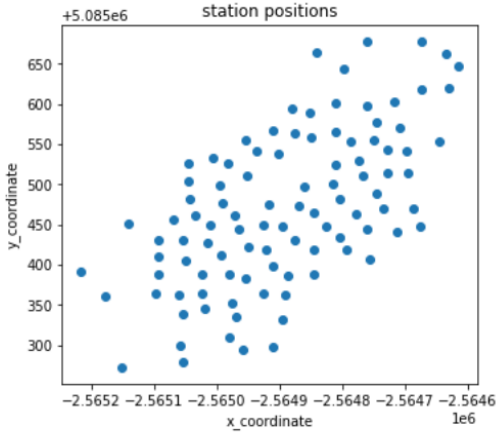
- The SKA1-Low has **512 stations, each containing 256 dual linearly polarized antennas,** so total **131,072** antennas will be configured with a **dense central core of antennas**, surrounded by **three spiral arms.**

- the **longest distance** (or baseline) between antennas will be **65km**. longer baselines mean the **higher resolution.** (resolution θ~λ/D)

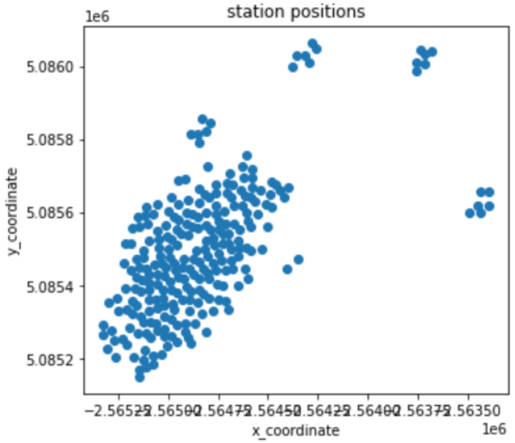
- SKA-Low uses **beamforming** technique to digitally "point" the telescope anywhere in the sky by combine signals from some or all of the antennas in the array.



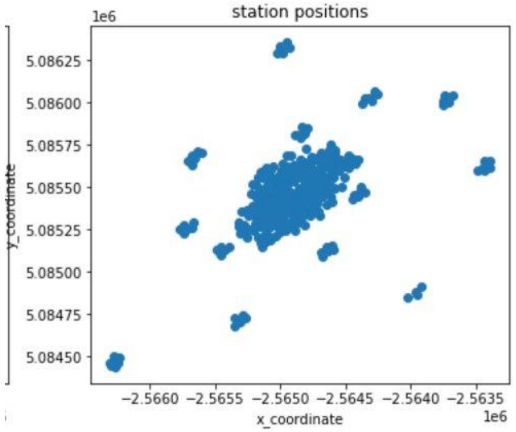
Source:https://www.space.com/square-kilometre-array-observatory-skao



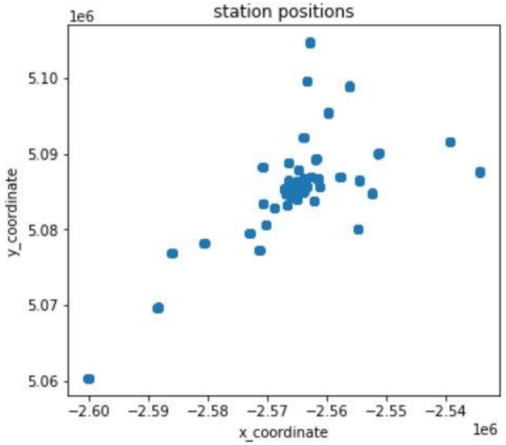
100 stations



200 stations



300 stations

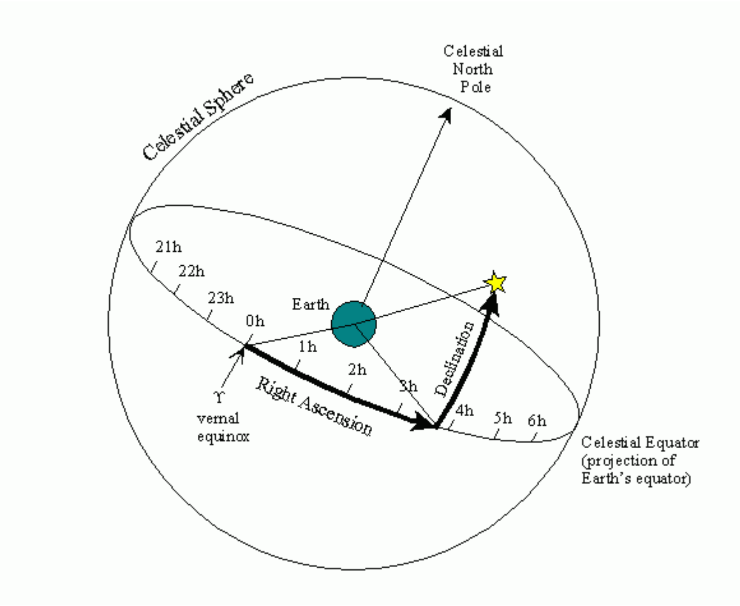


512 stations

SKA data specification:

- **Test data:**
- 150 uvfits test files was given
- This data was given to check that the estimator we are using is giving correct power-spectrum.
- Frequency coverage 166 – 181MHz

- **Main data**:
- 900 uvfits files
- 7.5TB
- Station beam image file(for 900 channels) (22.4GB)
- Field of View: 5°×5°in sky at RA, Dec = 0h, -30deg
- Frequency coverage 106 – 196MHz (z=12.39-6.24)
- Channel width: 100kHz (900 channels)
- Observation time: 4 hours
- Integration time: 10s

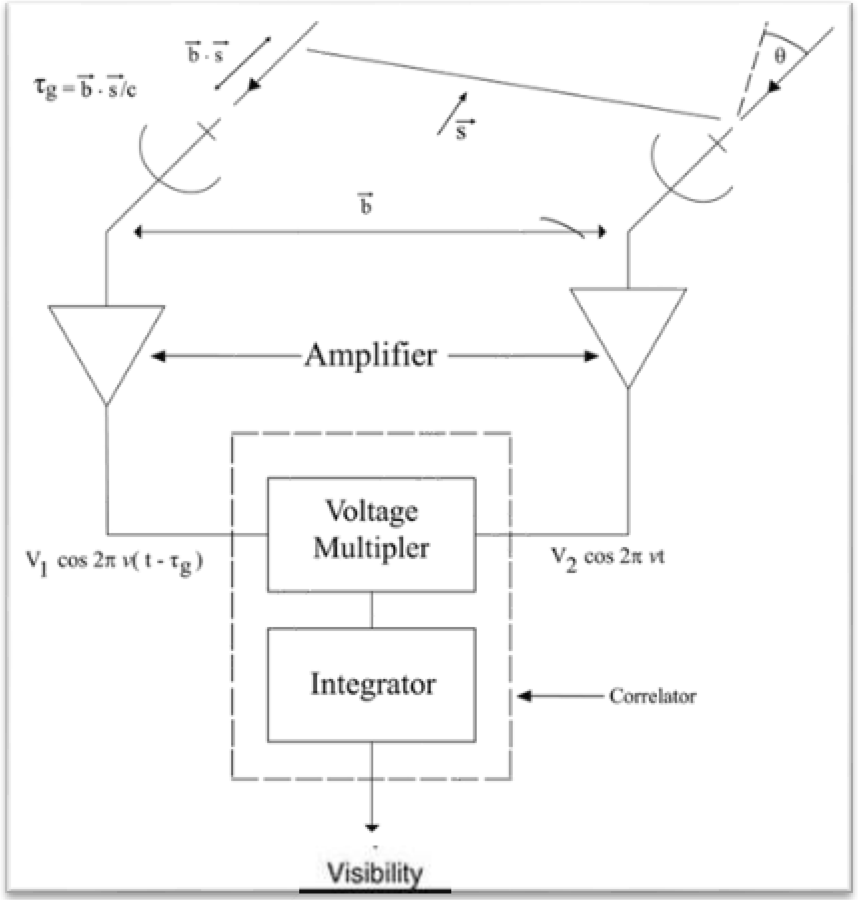


Source:https://www.cloudynights.com/topic/346535-explain-right-ascension-and-declination/

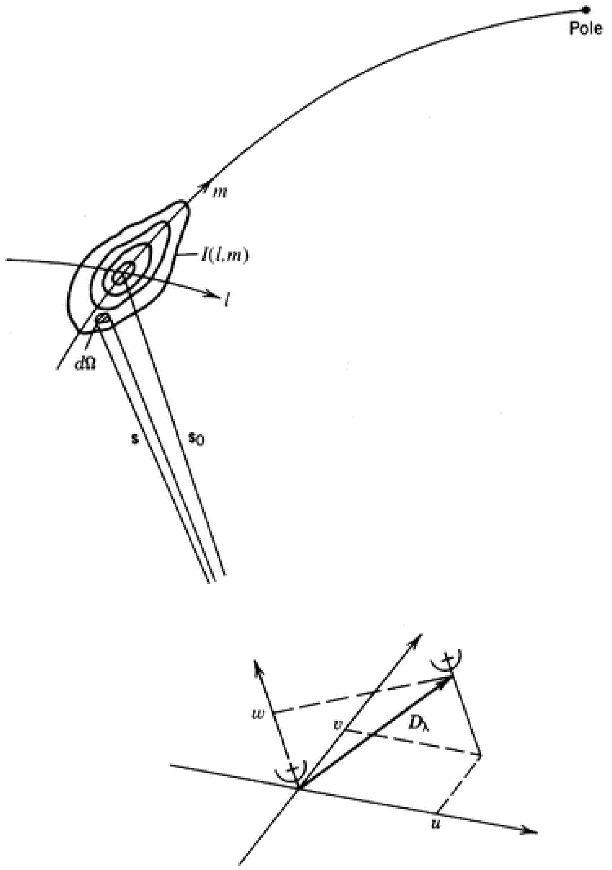
Tools we used to process the data:

- IIT-M Aqua cluster, SKA provided server and local server
- Astropy, numpy, matplotlib etc. (python package to read the and plot the data)
- **CASA** (common astronomy software application) to process the data of UVfits files.

Concept of Visibility:



Source:https://www.researchgate.net/figure/Correlator-circuit-for-two-elements-interferometer_fig1_330038494

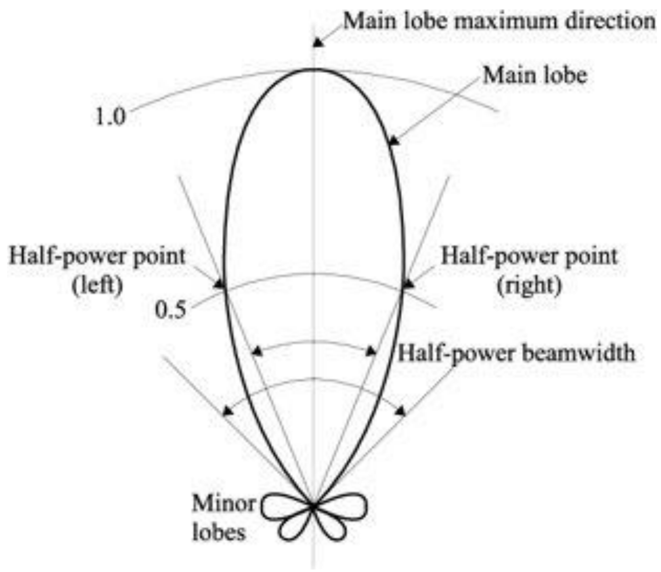


Source: "Interferometry and the synthesis in radio astronomy by A.Thompson and M.Moran"

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

- So, for **small field of view** the visibility is the 2D Fourier transform of the source intensity pattern/the sky brightness.

Antenna beam pattern:

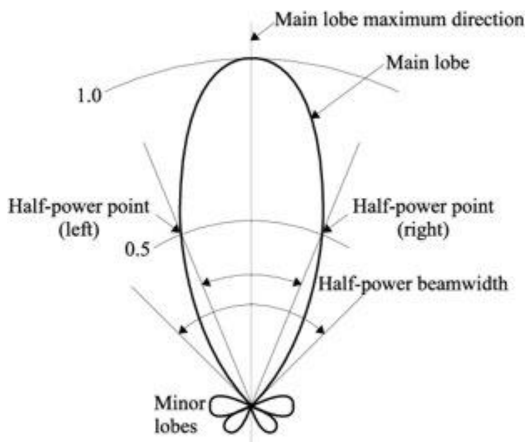


Source:https://images.app.goo.gl/MVyhWd9BfBBn8jqG9

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

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

Antenna beam pattern:



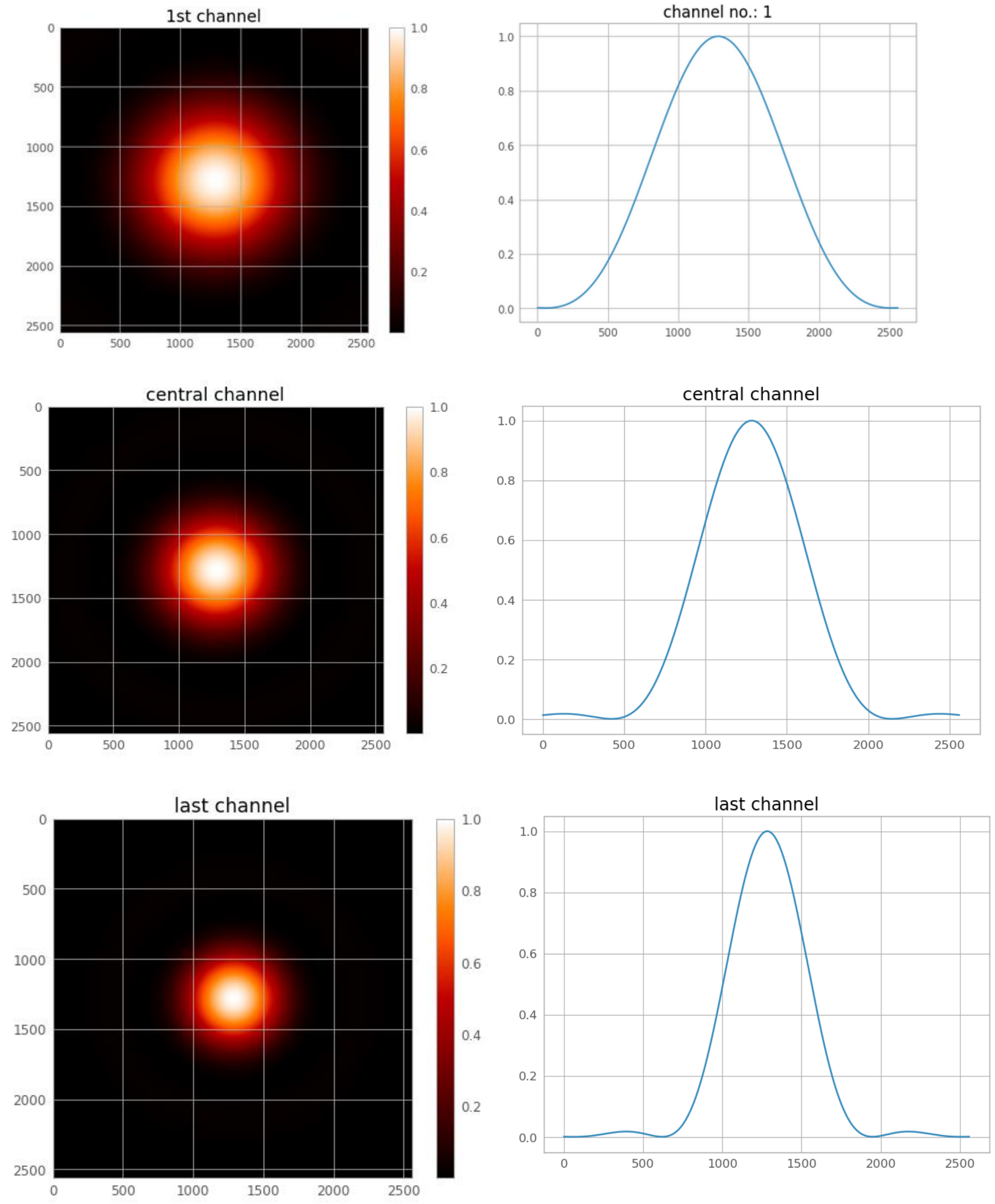
Source:https://images.app.goo.gl/MVyhWd9BfBBn8jqG9

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

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

- primary beam gets narrower. So, for higher frequency resolution increases.
- Side lobe contribution increases for higher frequency.

Some Plots of Station Beam pattern (from the Station Beam fits file)

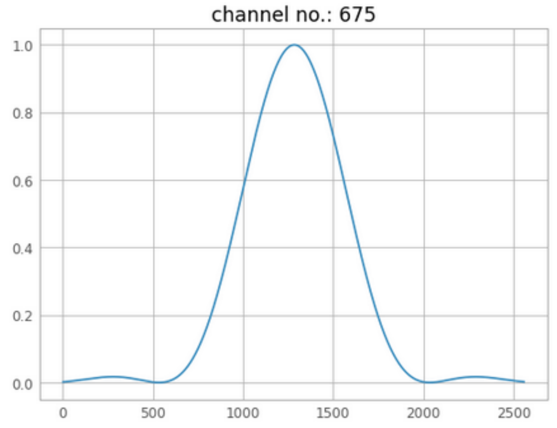


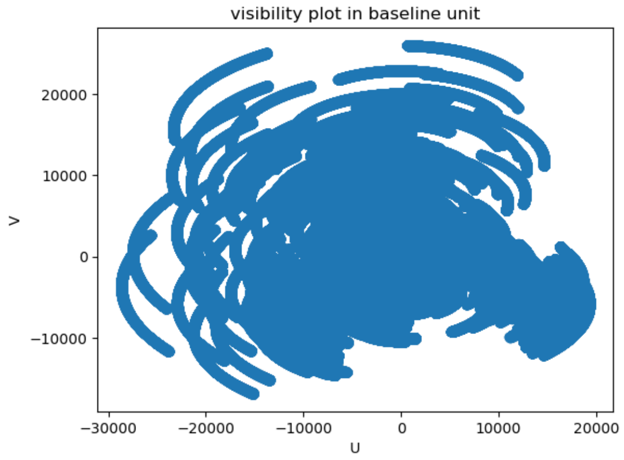
TGE(Tapered Gridded Estimator):

- Point source presented near the first null and the side lobe add foreground noise to the signal.

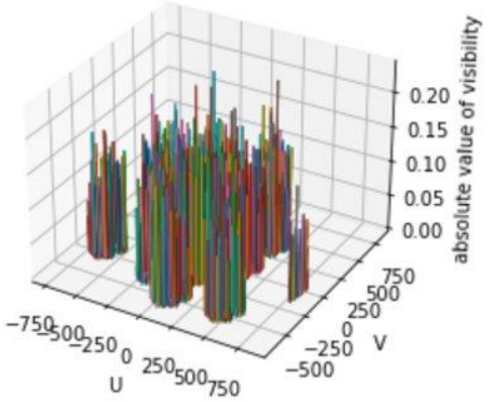
- To get the 2D sky image and the power spectrum we have to perform the **Fourier transform**(fast Fourier transform to individual visibility point). But which is **computationally very expensive**.(188375040 visibility points shown in second side figure)

- So we will use The visibility-based tapered gridded estimator (TGE) to generate the power spectrum for further analysis.





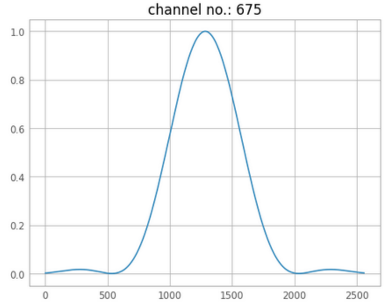
TGE(Tapered Gridded Estimator):



For 10000 visiblity points

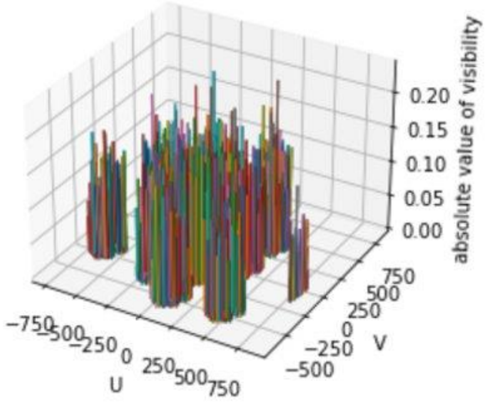
Tapering → Gridding → Apply the estimator → Angular Power spectrum → Cylindrical Power spectrum

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



source:(10.1093/mnras/stu2027)

TGE(Tapered Gridded Estimator):

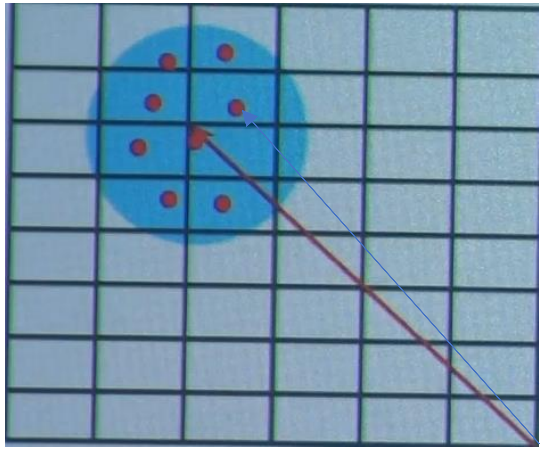


For 10000 visiblity points

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

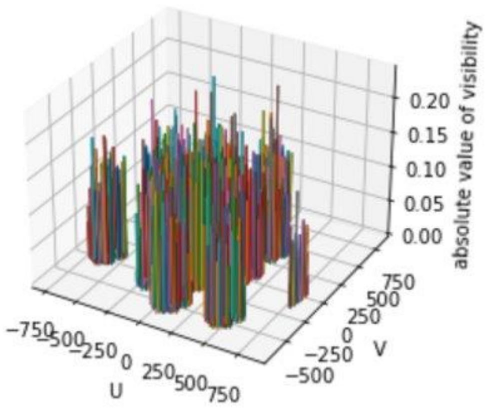
Tapering → Gridding → Apply the estimator → Angular Power spectrum → Cylindrical Power spectrum

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



source:(10.1093/mnras/stu2027)
https://youtu.be/r14nisJc4pQ?si=8dfn98g1liMwwFTW

TGE(Tapered Gridded Estimator):



For 10000 visiblity points

Tapering → Gridding → Apply the estimator → Angular Power spectrum → Cylindrical Power spectrum

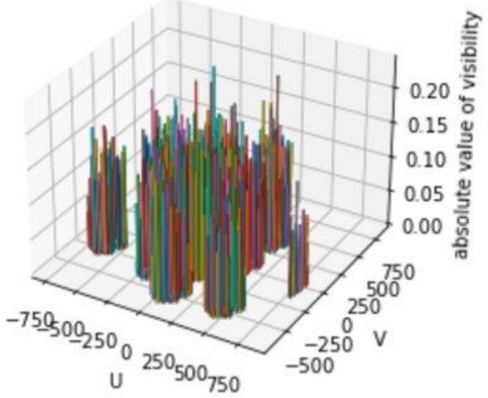
$$V_{cg} = \sum_i \tilde{w}(U_g - U_i)V_i$$

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

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

source:(10.1093/mnras/stu2027)

TGE(Tapered Gridded Estimator):



For 10000 visiblity points

Tapering → Gridding → Apply the estimator → Angular Power spectrum → Cylindrical Power spectrum

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

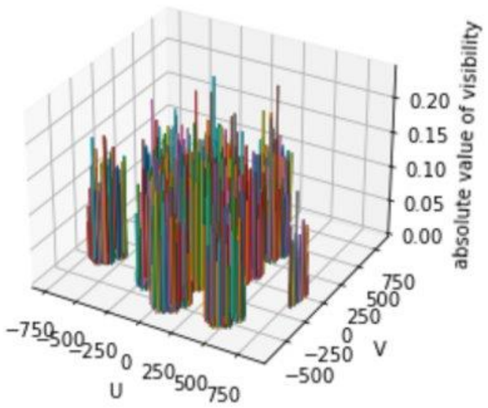
$$\langle \hat{E}_g \rangle = C_{\ell_g}$$

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

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

source:(10.1093/mnras/stu2027)

TGE(Tapered Gridded Estimator):



For 10000 visiblity points

Tapering → Gridding → Apply the estimator → Angular Power spectrum → Cylindrical Power spectrum

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

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

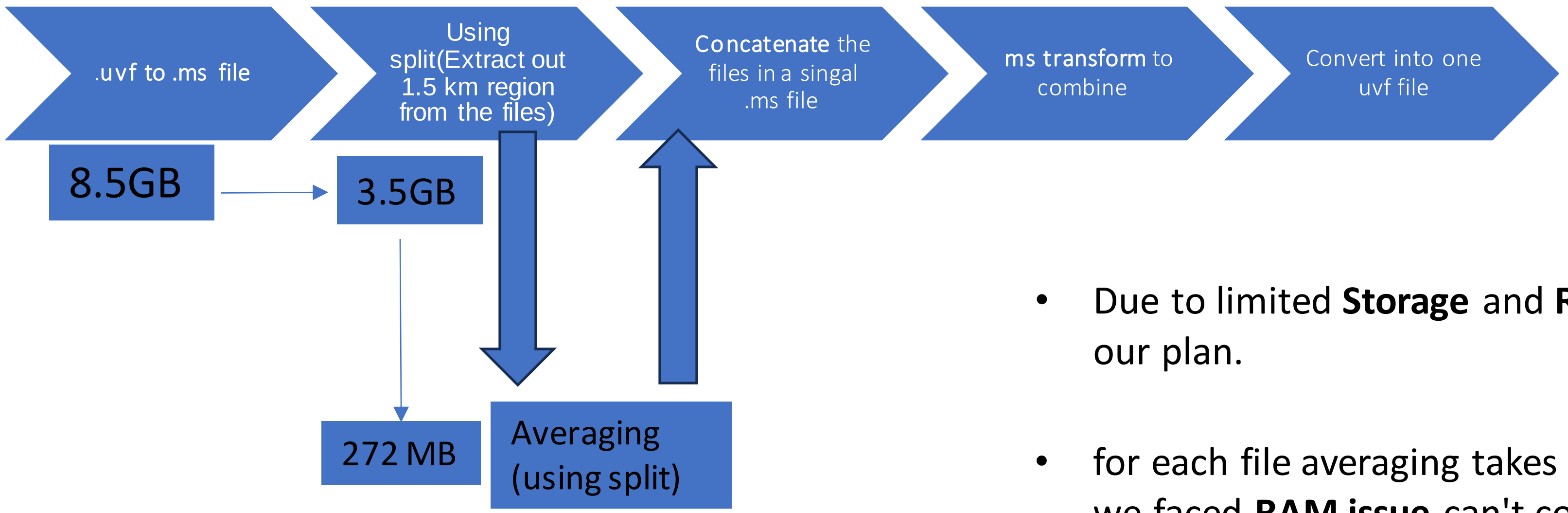
$$\hat{E}_g = M_g^{-1} \left(|V_{\rm cg}|^2 - \sum_i |\tilde{w}(U_g - U_i)|^2 |V_i|^2 \right)$$

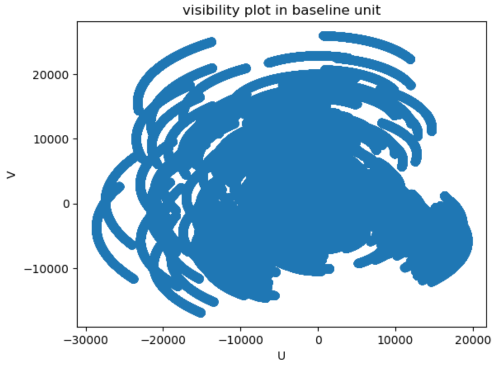
$$\langle \hat{E}_g \rangle = C_{\ell_g}$$

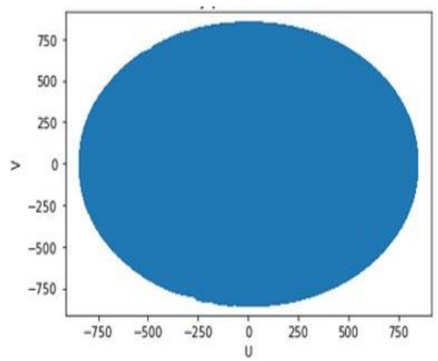
$$P(k_\perp, k_\parallel) = r^2 r' \int_{-\infty}^{\infty} d(\Delta\nu) e^{-ik_\parallel r' \Delta\nu} C_\ell(\Delta\nu)$$

source:(10.1093/mnras/stu2027)
(10.1093/mnras/sty3501)

Data processing:







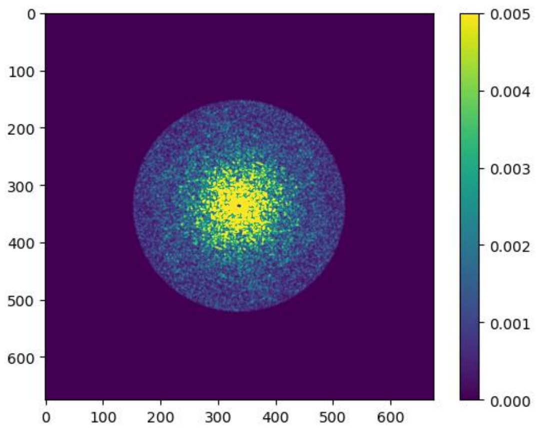
- Here we extract out 1.5 km region (or 830 in wavelength unit) from the original visibility plane.(188375040, 1, 1, 1, 1, 1, 3) to (56007728, 1, 1, 1, 5, 1, 3)

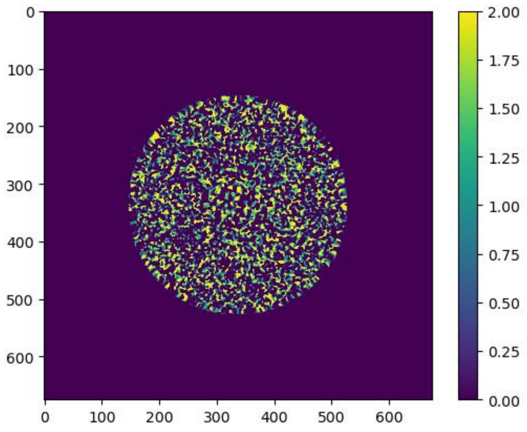
- Due to limited **Storage** and **RAM** in local server we had to change our plan.
- for each file averaging takes **7 hours 41 mins** in our local server but then we faced **RAM issue** can't combine the files**.**
- We move to the **SKA server** to average out 150 files and combine them in a single uvf file frequency range of 166-181 MHz.
- But when we try to work main data the averaging of each file takes **35 hours 55 mins.**

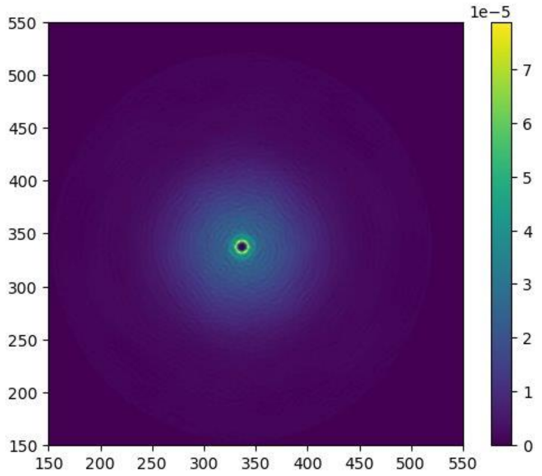
Data Processing:

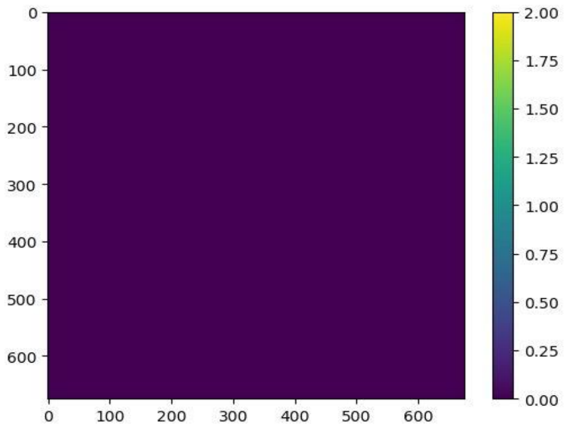
- We use the **time very efficiently**(by running multiple script files) for averaging.
- Process out 1st 150 uvfits files out of **900** and combine them to make one **band** of frequency range 106-121 MHz.
- We make 10 realizations of our combined test data and convolve with the primary beam pattern to get the convolved visibilities.
- Then, we apply the TGE for the combined test data and the 10 realizations.
- As a output we get binary files.
- We make a python code to the read the binary files and extracted out the convolved tapered gridded visibilities and save them in two separated files for further processing.

<u>**Amplitude** and **Phase** of the visibility after applying TGE for test data for channel no. 10: (56007728)--> (455625)</u>



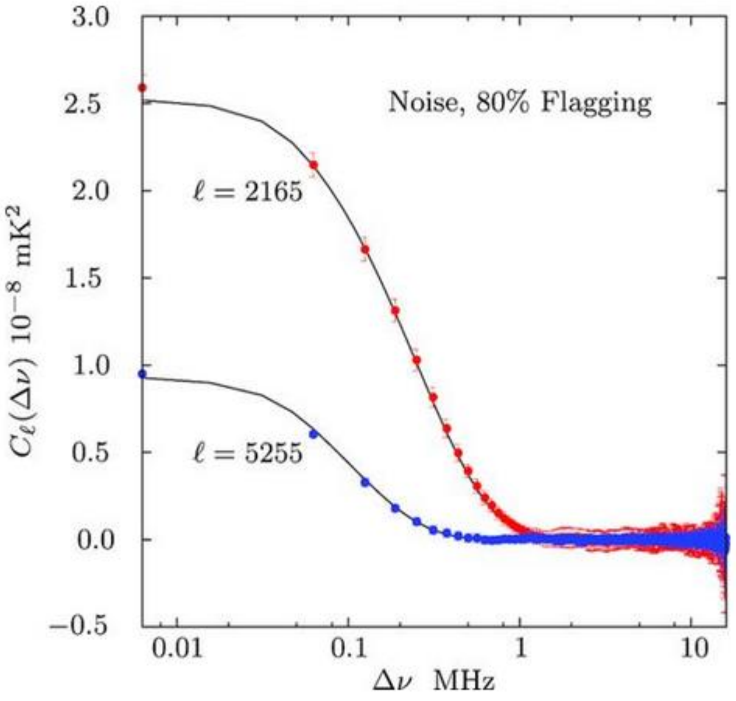


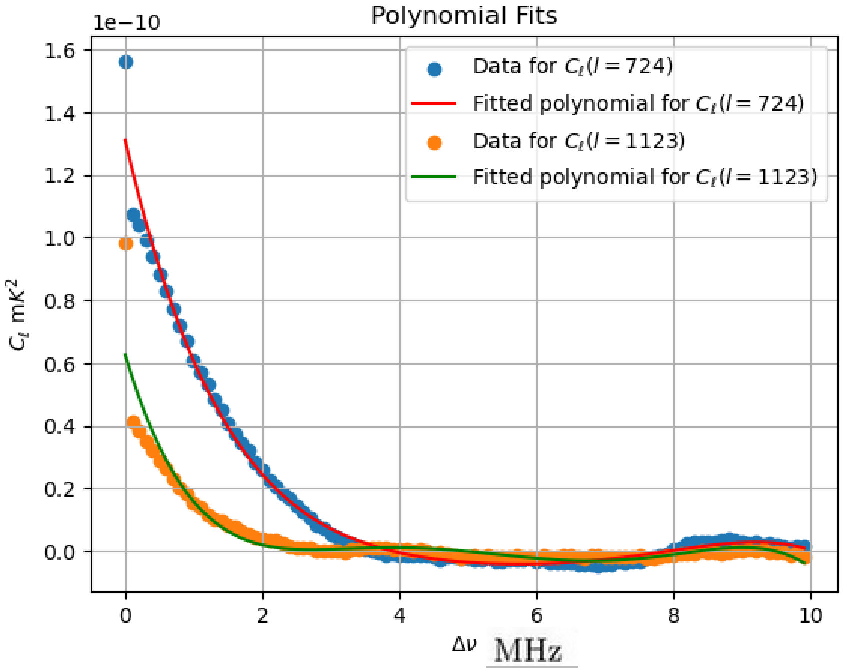




the axes are in baseline unit.

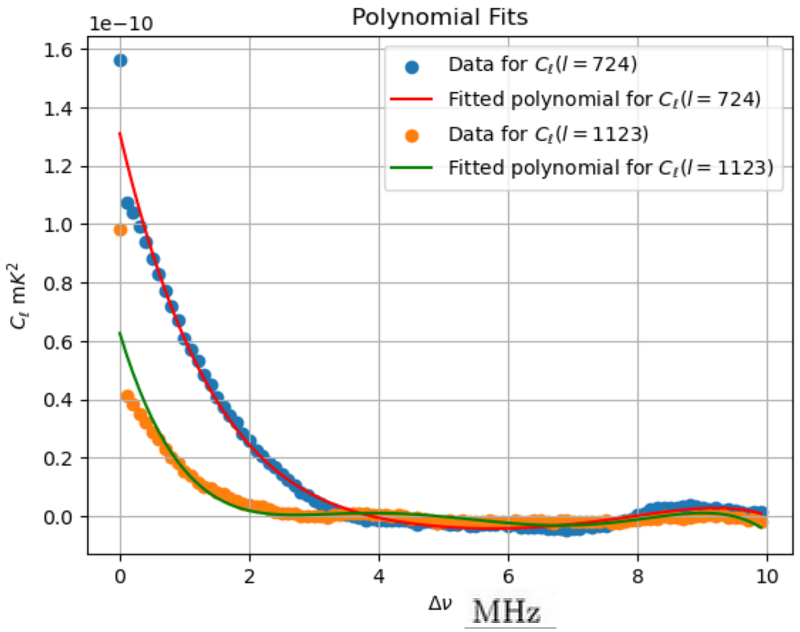
The angular power spectrum(test data):

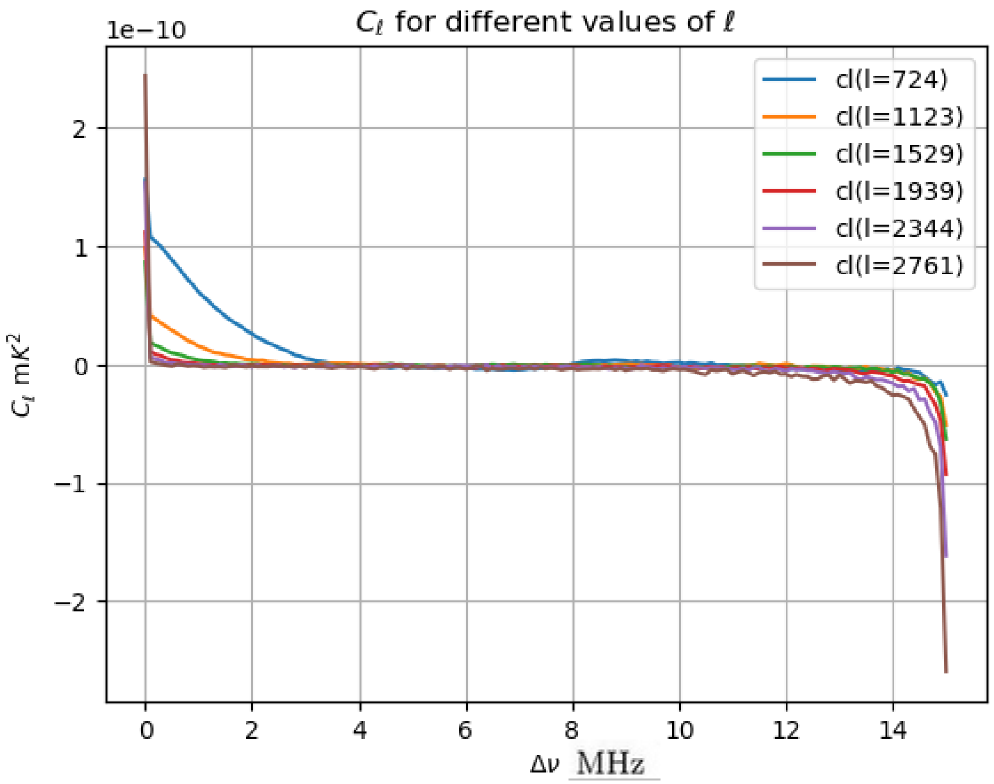




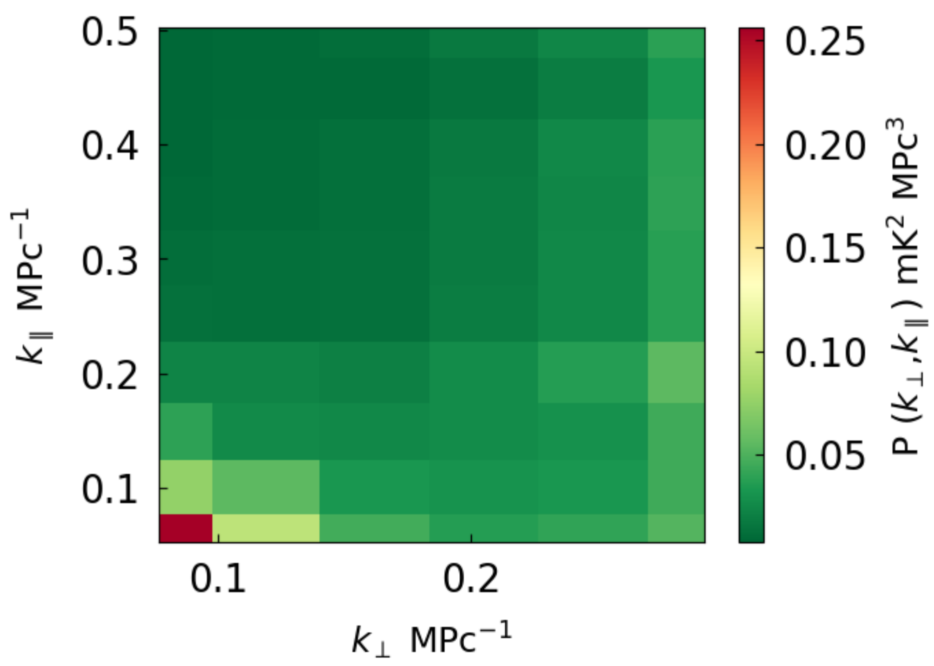
Somnath Bharadwaj, Srijita Pal, Samir Choudhuri, Prasun Dutta, A Tapered Gridded Estimator (TGE) for the multifrequency angular power spectrum (MAPS) and the cosmological H i 21-cm power spectrum, *Monthly Notices of the Royal Astronomical Society*, Volume 483, Issue 4, March 2019, Pages 5694–5700, https://doi.org/10.1093/mnras/sty3501

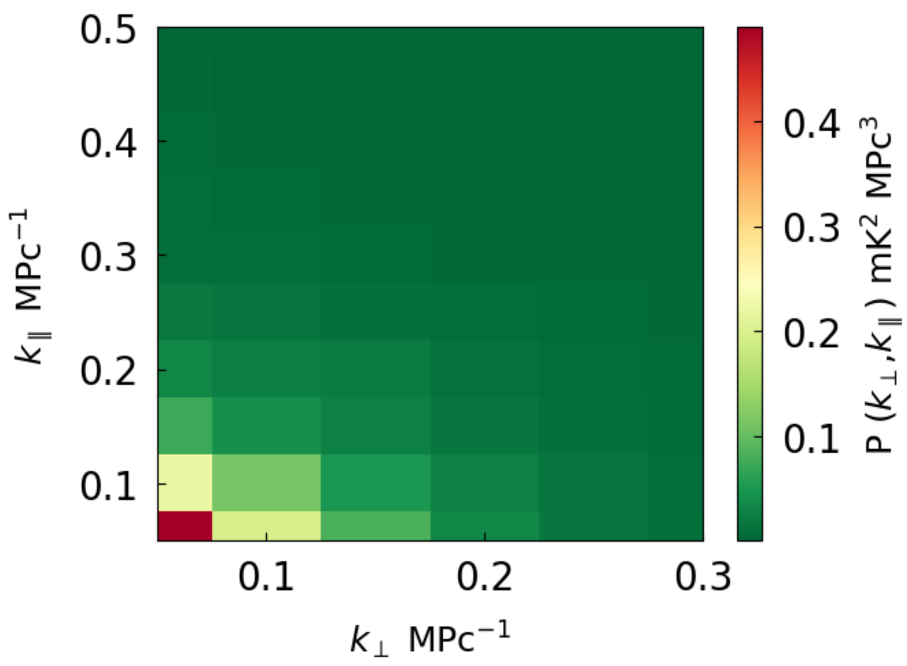
The angular power spectrum(test data):



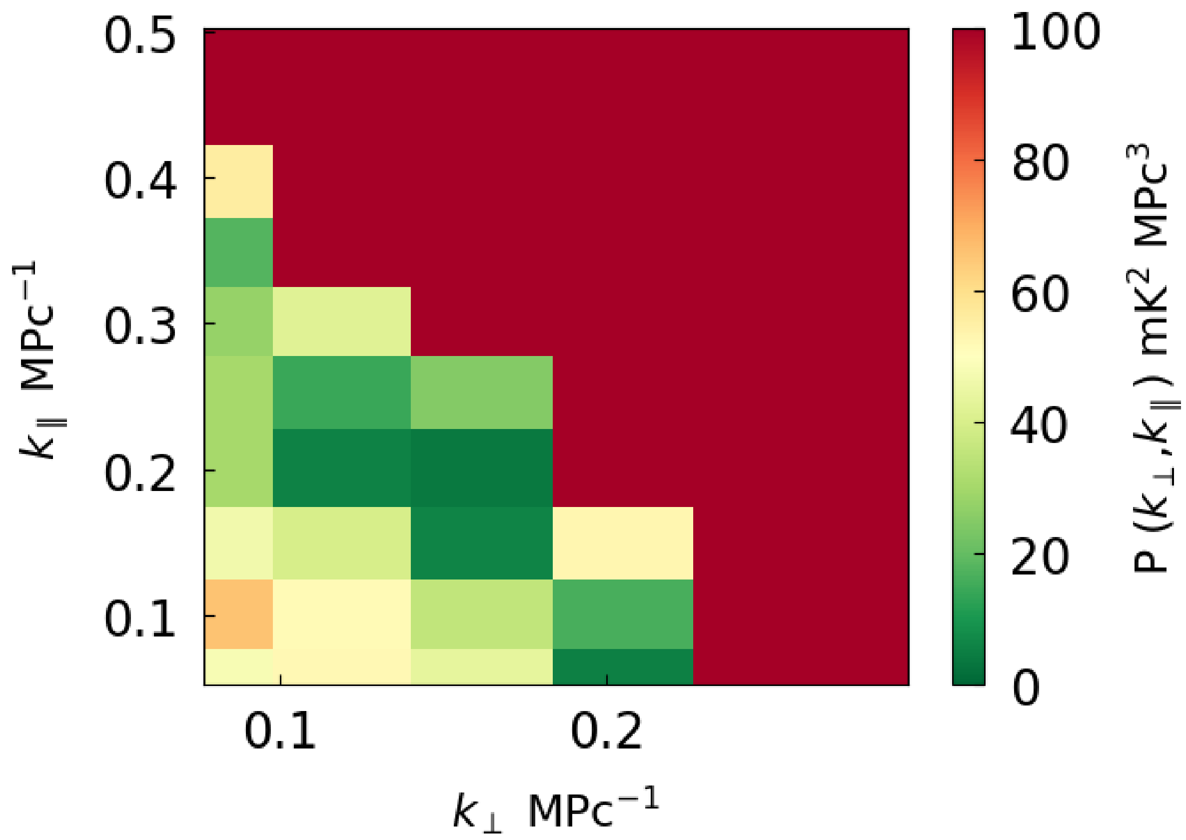


Estimated Power Spectrum(left) and Input model Power Spectrum(right)

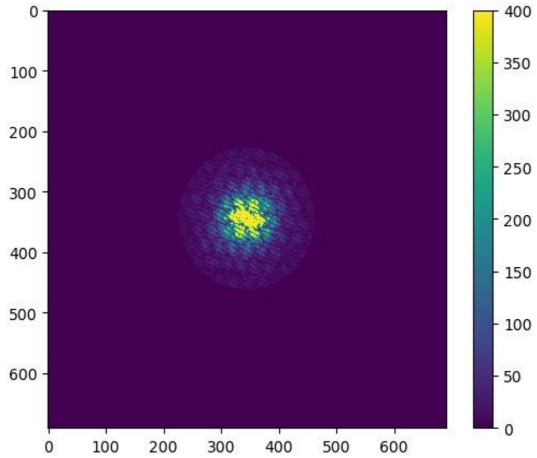


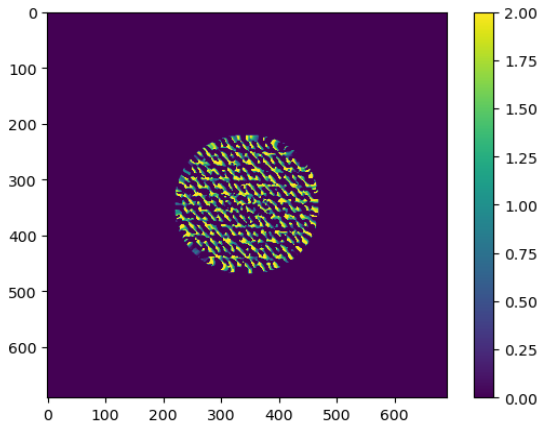


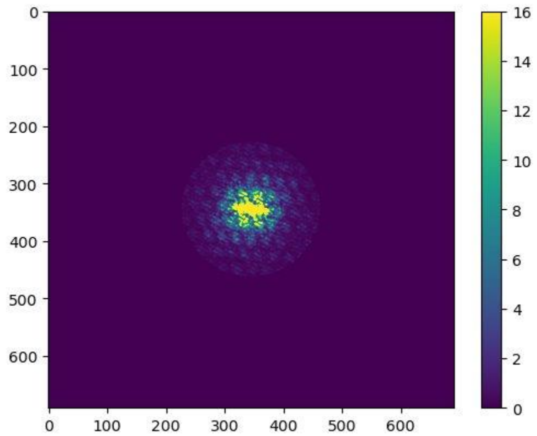
Percentage deviation:

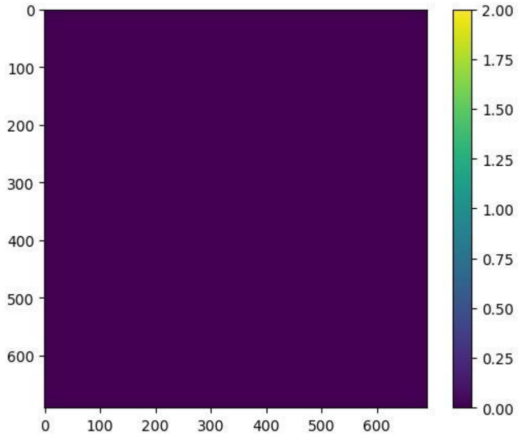


Amplitude and **Phase** of the visibility after applying TGE for true data(1st band) for channel no. 10: (56007728)--> (455625)

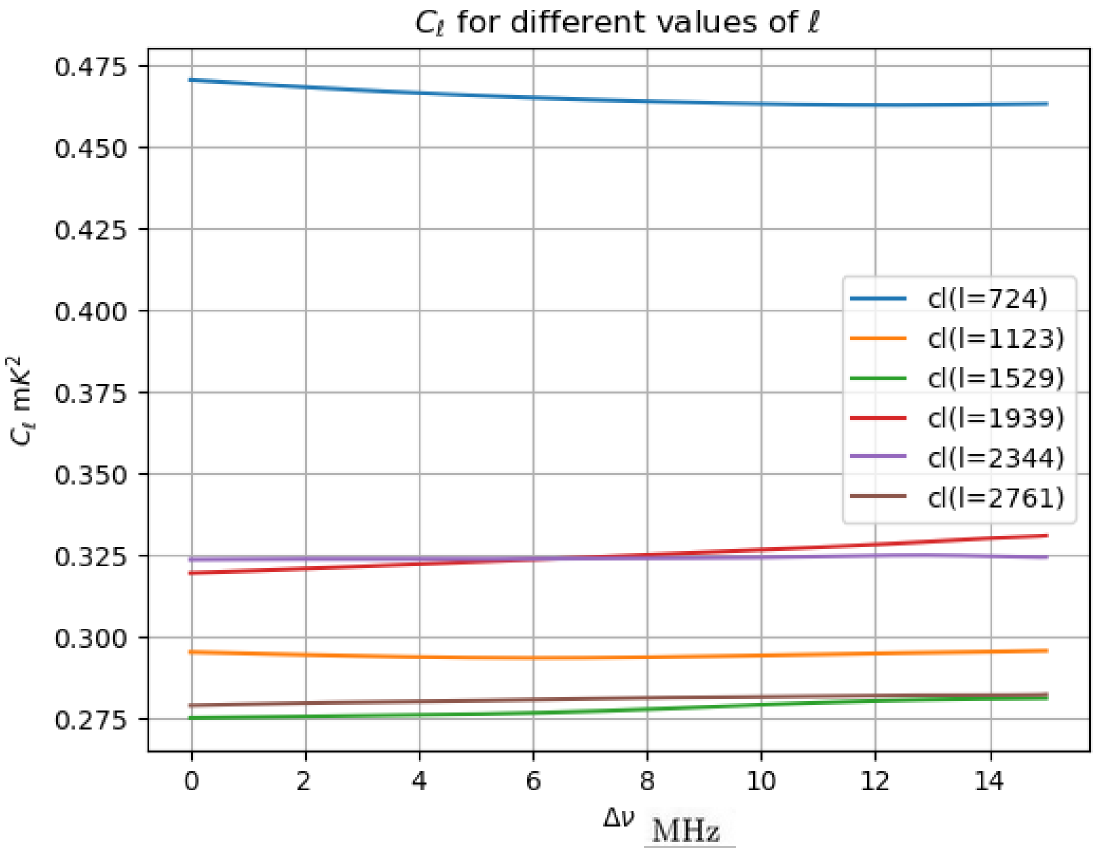


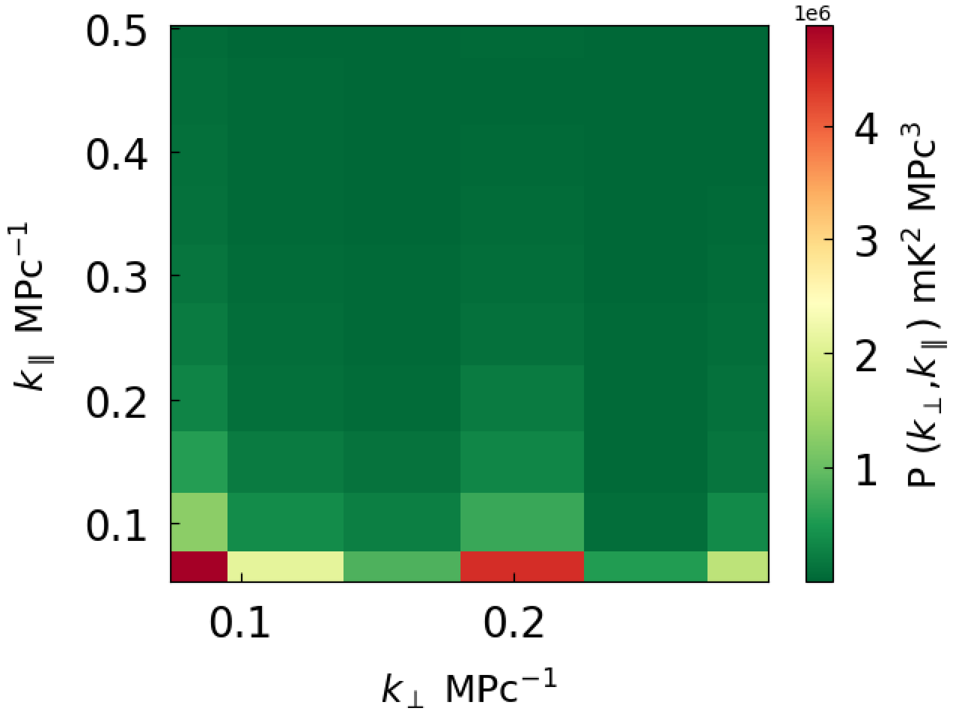






Angular and Cylindrical Power spectrum for true data (1st band of 106-121 MHz)





Thank you