Hydrogen Epoch of Reionization Array Experiment (HERA):

Effect of Radio Frequency Interference Excision on 21-cm Power Spectrum

Ankur Dev
MSc (Astrophysics) – June 2020
Queen Mary University of London, UK

Queen Mary
University of London

CMB and Dark Ages

- Recombination : observed as the Cosmic Microwave Background (CMB)
- The universe was opaque before the recombination.
- After the Universe became neutral, it became transparent.
- ISM – predominantly consists of HI at this early epoch
- Dark ages began – no light sources

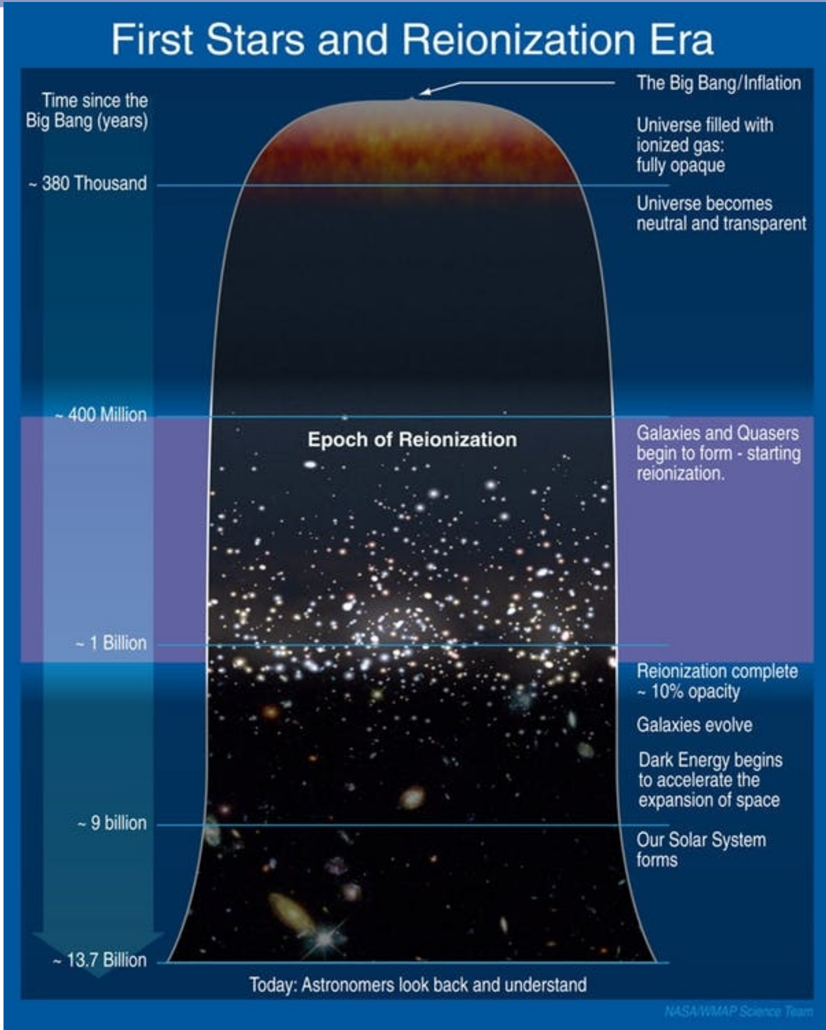


Image: From NASA/WMAP Science Team

Epoch of Reionization

- Ignition of the first stars marks the end of the Dark Ages and the beginning of “Cosmic Dawn”
- Epoch of Reionization (EoR) - neutral intergalactic medium ionized by the emergence of the first luminous sources
- Reionization complete redshift z ~ 6.5
- Exact nature of the first reionizing sources not precisely known

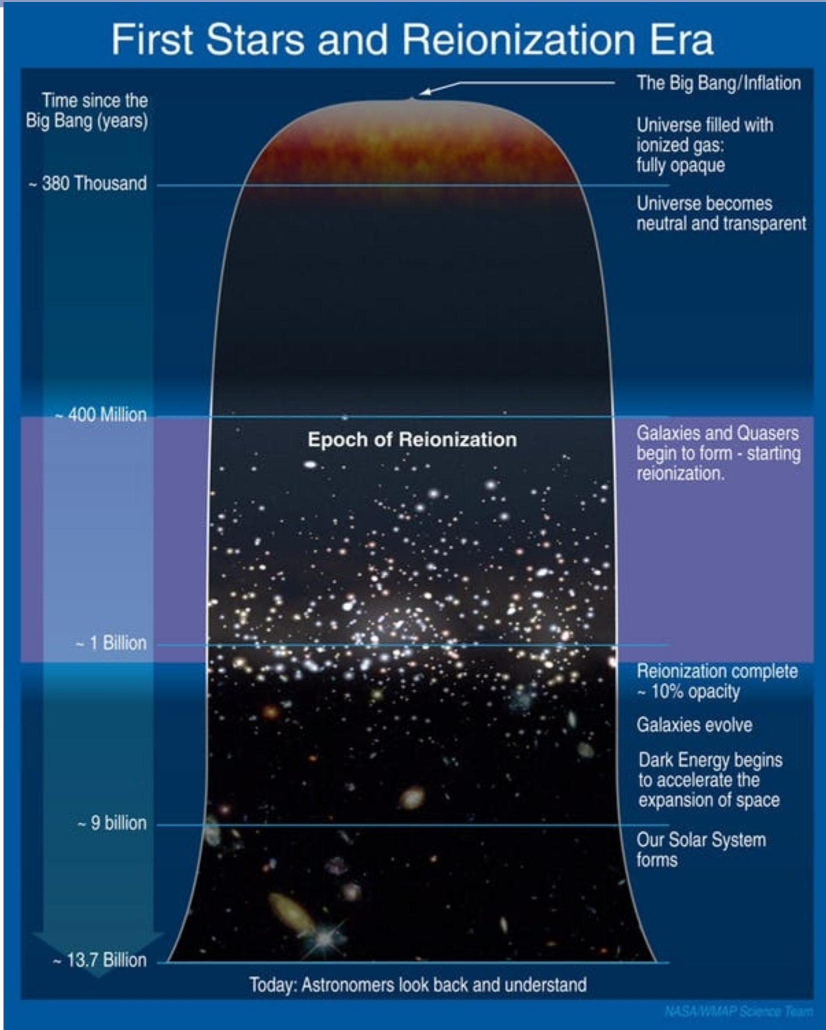


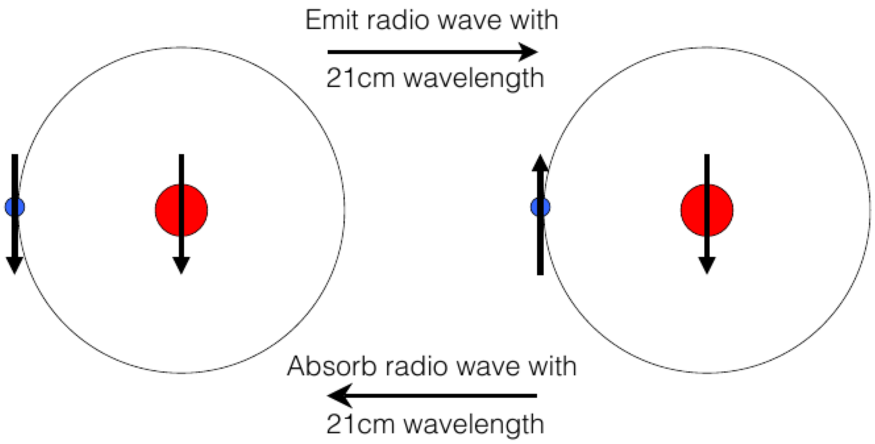
Image: From NASA/WMAP Science Team

21-cm line of Neutral Hydrogen

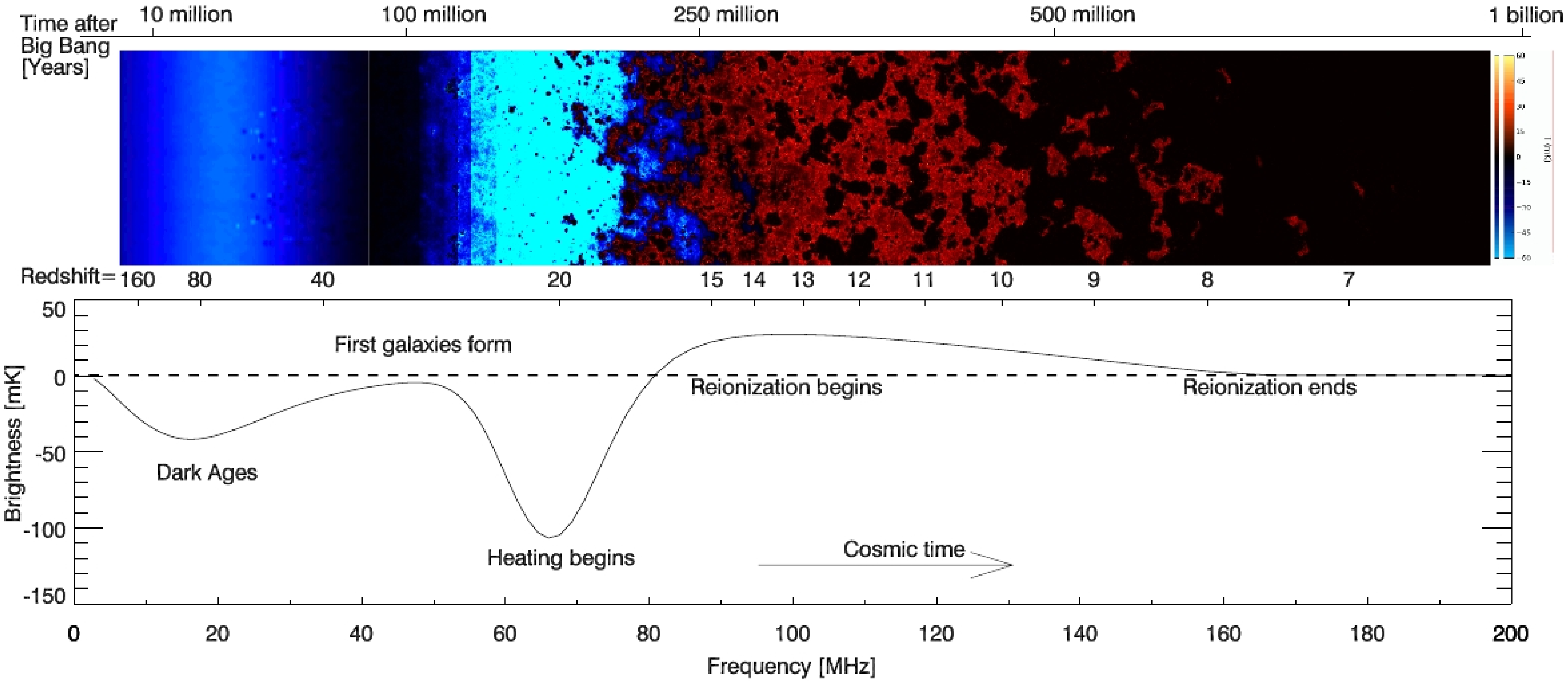
- 21 cm line neutral H line - created by a change in the energy state of neutral hydrogen atoms.

- Spin flip - resulting transition has a rest frame frequency of 1.4 GHz, and a wavelength of 21 cm.

- 21 cm line – redshifted to radio frequencies – used as cosmological probe.



Spin-flip transition in neutral Hydrogen. Created by a change in the energy state of neutral hydrogen atoms.



Evolution of intergalactic Hydrogen as a function of redshift. Red colour represents the 21-cm radiation relative to the CMB.
Black curve below shows expected global 21-cm signal, shown here as a sky-averaged spectrum. Adapted from: Pritchard and Loeb (2012)

HERA

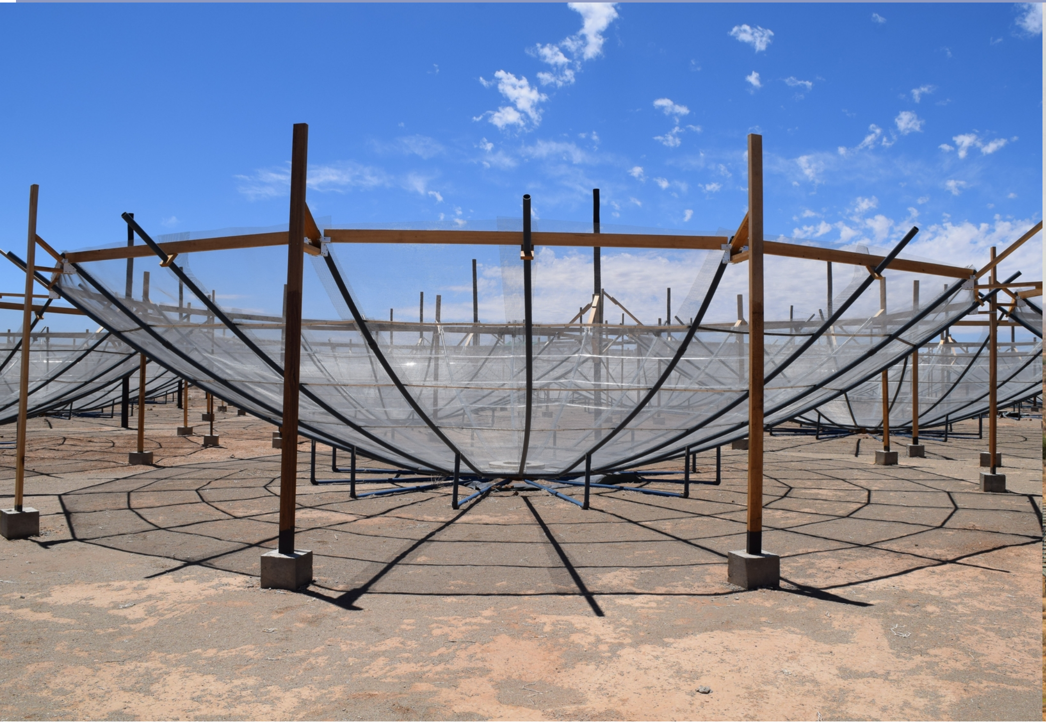
- Hydrogen Epoch of Reionization Array (HERA) experiment to measure 21 cm emission from the primordial intergalactic medium (IGM) throughout cosmic reionization ($z = 6 - 12$)

- Radio telescope - Located in the Karoo Desert of South Africa - 350-element interferometer – 14m non-tracking parabolic dishes observing from 50 to 250 MHz.

HERA dishes in the Karoo Desert. Image from February 2016

Ref: http://reionization.org/

HERA



Left: Single element - Parabolic HERA dish. Right: HERA photographed in March 2020.
(Image Credit: Dara Storer, HERA Collaboration).

Measuring the EoR

- Two categories of 21-cm experiments:
  - one for measuring the “global” sky averaged signal Example: EDGES, SARAS
  - the other for measuring the brightness temperature fluctuations. Examples: Experiments such as MWA, PAPER, LOFAR and HERA.

- Measure the fluctuation in δT_b (brightness temperature of 21-cm line measured against the CMB) caused by the Lyman-α and x-ray heating fluctuations.

- Measuring the intensity at different frequencies corresponding to different redshifts we can obtain a statistical measure of the fluctuations.

- Foreground signal is predominantly spectrally smooth because the electrons emit synchrotron or free-free emission over a broad range of frequencies

- EoR Power Spectrum: two techniques – Delay-Spectrum Approach $P(\boldsymbol{k})$ and map-making (imaging)

Challenges in EoR detection

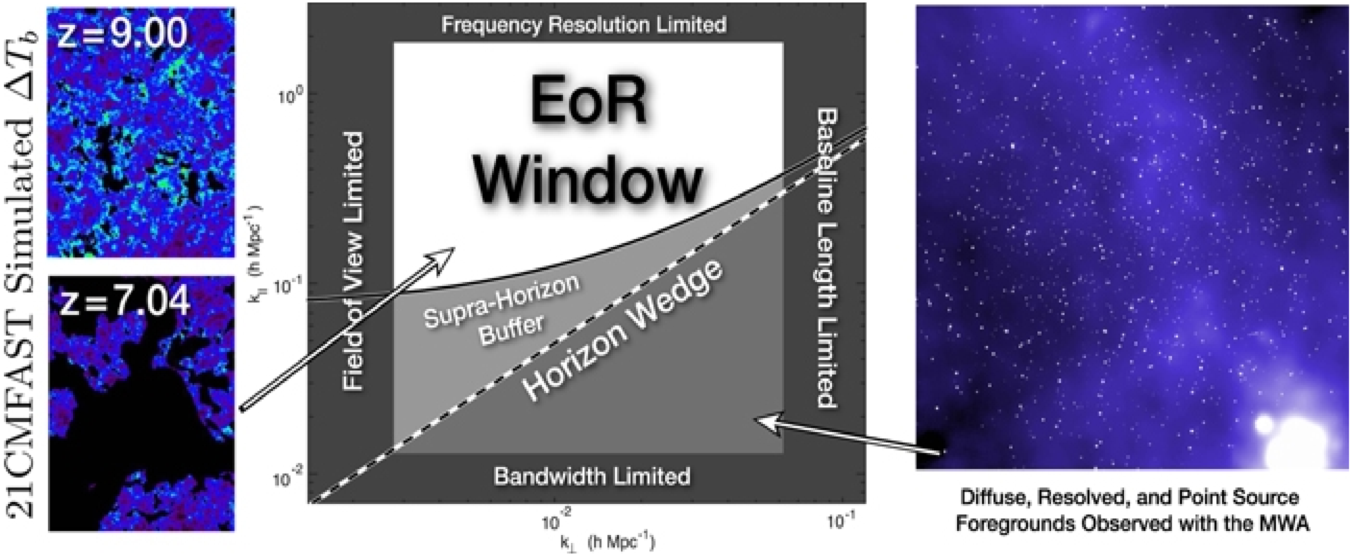
- Foreground contamination - foreground contribution $\sim 5-6$ orders of magnitude brighter than the 21 cm signal
  - Diffuse Galactic synchrotron radiation
  - Supernova remnants
  - Extragalactic radio sources

- Radio frequency interference (RFI) : from FM radio, TV, Satellites
- Instrument calibration and stability
- Key issue to measuring the EOR power spectrum is to understand and minimize systematics.
- Goal - measure a statistical power spectrum of the signal over the sky, rather than direct imaging

Dealing with Foregrounds

- Foreground subtraction method. Issues: Over-fitting, residuals
- Alternative: **Foreground avoidance technique**. Used by HERA and PAPER.
- Leave out the parameter space that is mostly occupied by the strong foregrounds.
- Smooth-spectrum foregrounds interact with antenna beam chromaticity to produce a characteristic “wedge” of foreground leakage in Fourier space outside of which the 21 cm signal dominates (the “EOR window”). (DeBoer et al. (2017))

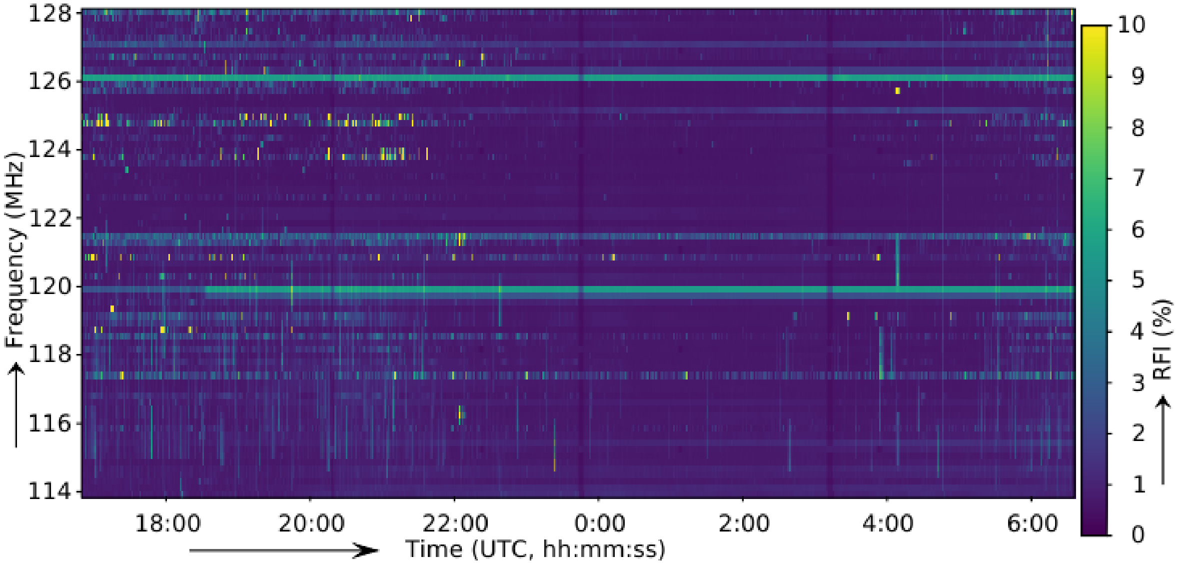
- Wavenumber: $\boldsymbol{k} = k_{\perp} + k_{\parallel}\hat{z}$
- “Contaminated” phase space is a wedge-shaped region in $k_{\perp} - k_{\parallel}$ space.
- Flat spectrum foregrounds cannot contaminate the EoR window beyond the horizon wedge.



The Problem - RFI

- Foreground signals are smooth spectrum sources
- EoR is expected to be rough : made up of non-ionized regions which are randomly distributed over a wide range of redshifts.
- RFI detection causes samples to be sporadically flagged and rejected.
- Fourier transform of the signal – sharp endpoint discontinuity. Spectral ringing contaminates the power spectrum

Maintaining spectral smoothness is key to separating foregrounds from 21cm signal.



Example of RFI occupancy over time and frequency obtained for a LOFAR observation of the North Celestial Pole.

(Image from: Offringa et al. (2019))

Simulations

- Simulate visibilities for ideal sky:
  - add diffuse foreground, point sources, noise, simulated EoR signal (noise-like) – referred as **unflagged data**

- Add random RFI flags – contaminate the data. Gives us **flagged data**

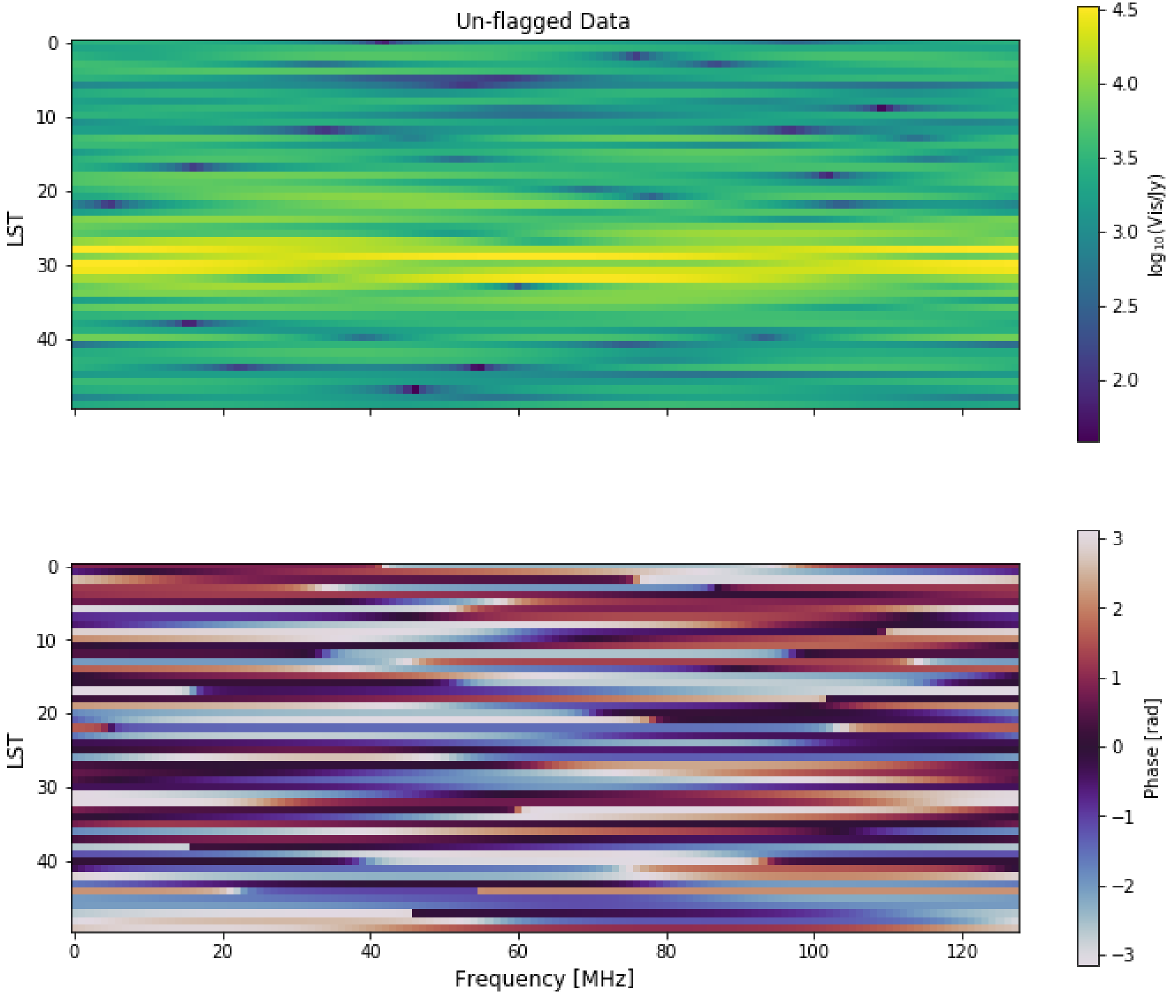
- Fit linear least square filter to RFI flagged data : **in-painting** process
  - The best-fit smooth model is fit to flagged data, using a linear least-squares solver.
  - The model is a Fourier series up to a specified order.

- Compare power spectrum of in-painted data to the unflagged power spectrum – how does it depend on RFI occupancy?
- *Python:* Used and worked on *hera_pspec*, *hera_sim* packages

Results – Unflagged (No RFI)

Shows the visibility (**amplitude -**Upper panel and **phase** - Lower panel)) of the simulated data.

50 LST (Local Sidereal Time) bins and 128 frequency bins. 16 hours of total observation time



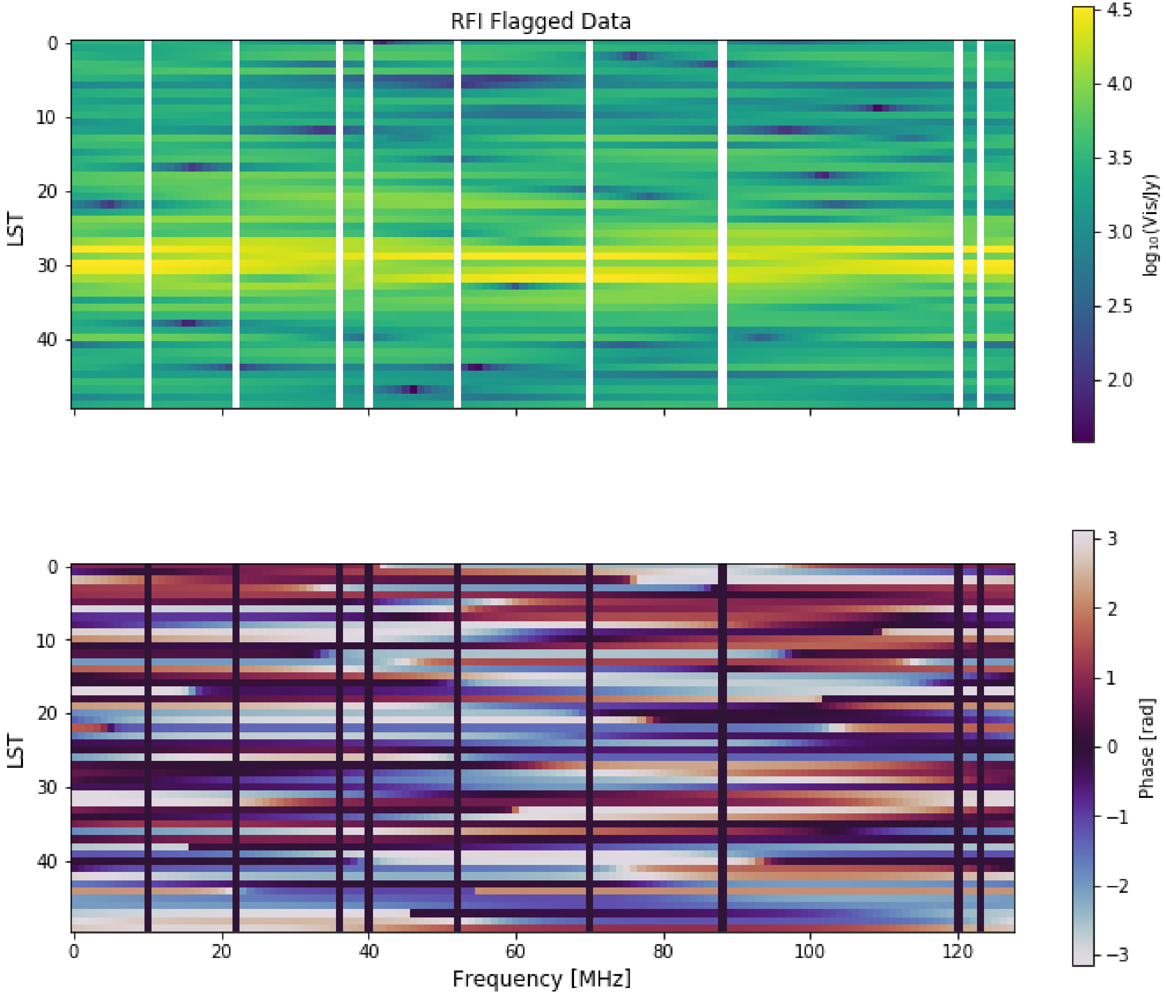
Simulated data without any flags. Ideal case scenario. Un-flagged 'ideal' sky visibilities after including diffuse foreground, 500 point sources, noise and EoR signal.

Results – Flagged (With RFI)

Simulated data with RFI flags. RFI flags are randomly added to the previous ideal sky data.

Same RFI flags are applied for all LST bins. Fraction of RFI flagged channels here is 7%

Shows the visibility (amplitude and phase) of the RFI flagged data.

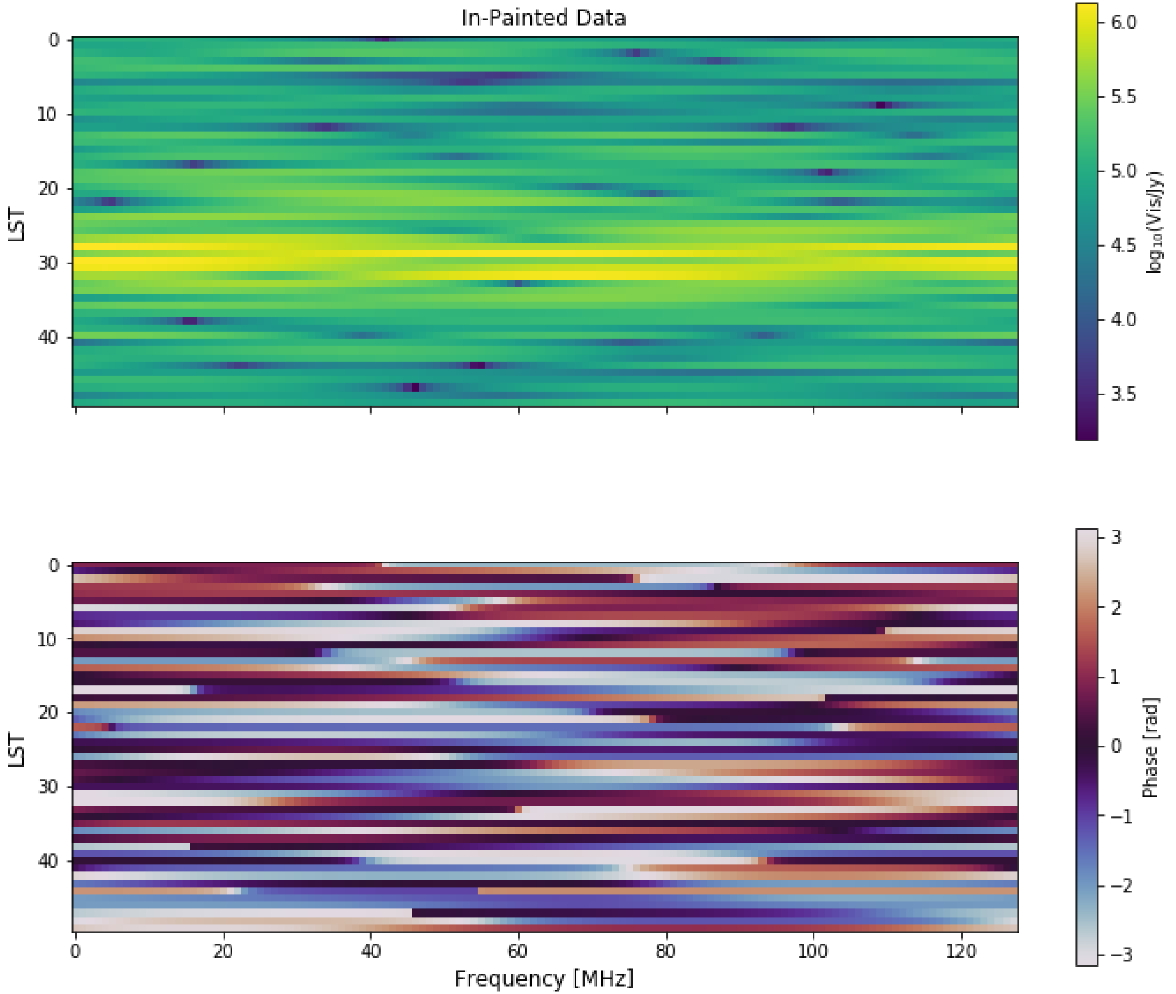


Results – In-Painted Data

Post best-fit model fitting to previous RFI flagged data.

Visibilities after applying least squares filter are shown.

The visibilities are similar to the ideal sky data

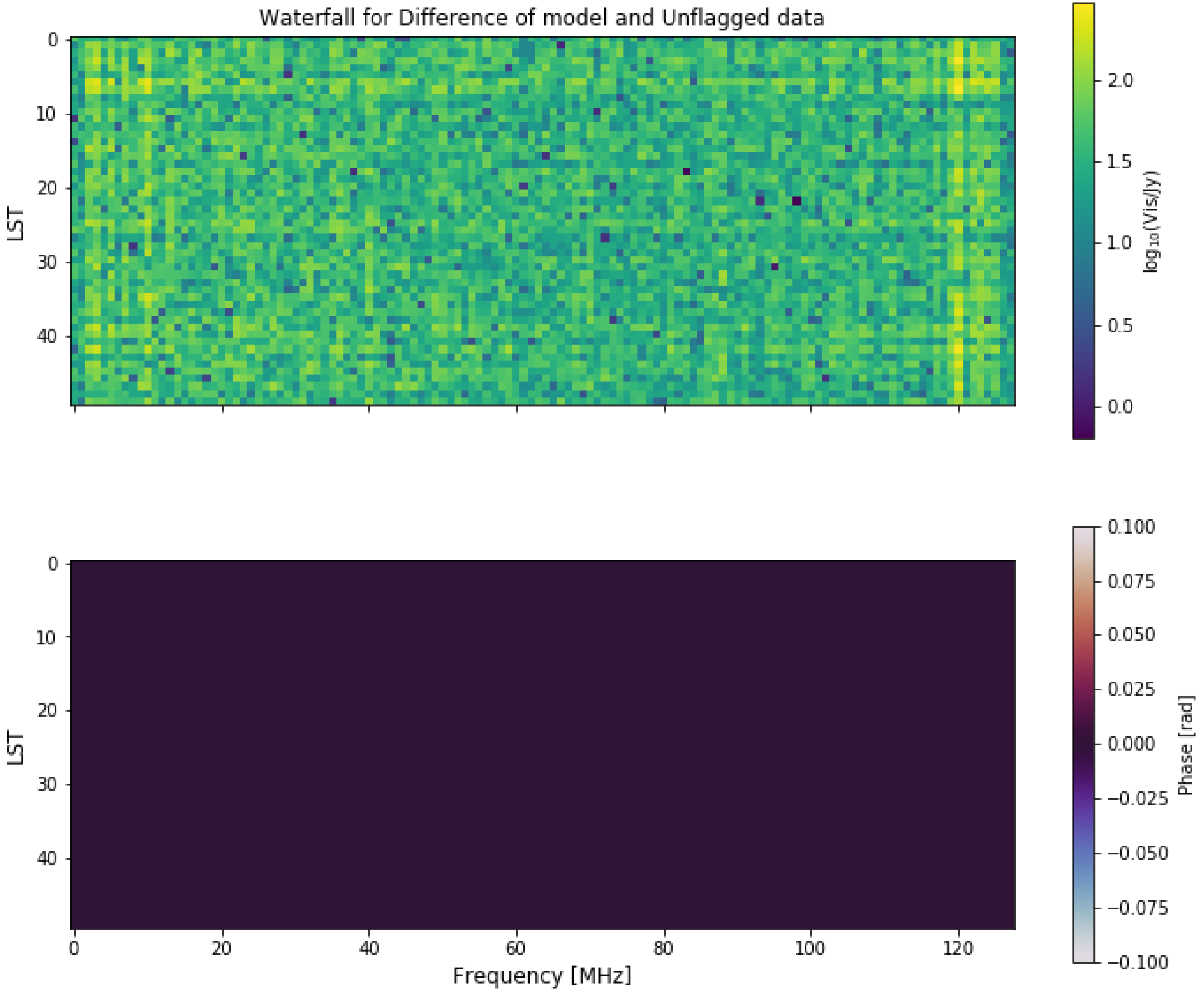


Model and Ideal Data Difference

Visibilities after taking difference of smooth fitted model and un-flagged data.

This shows the noisy residual in the data, which is at the same level as the noise introduced earlier in the simulation.

The model is able to recover the ideal case scenario from the RFI flagged dataset.

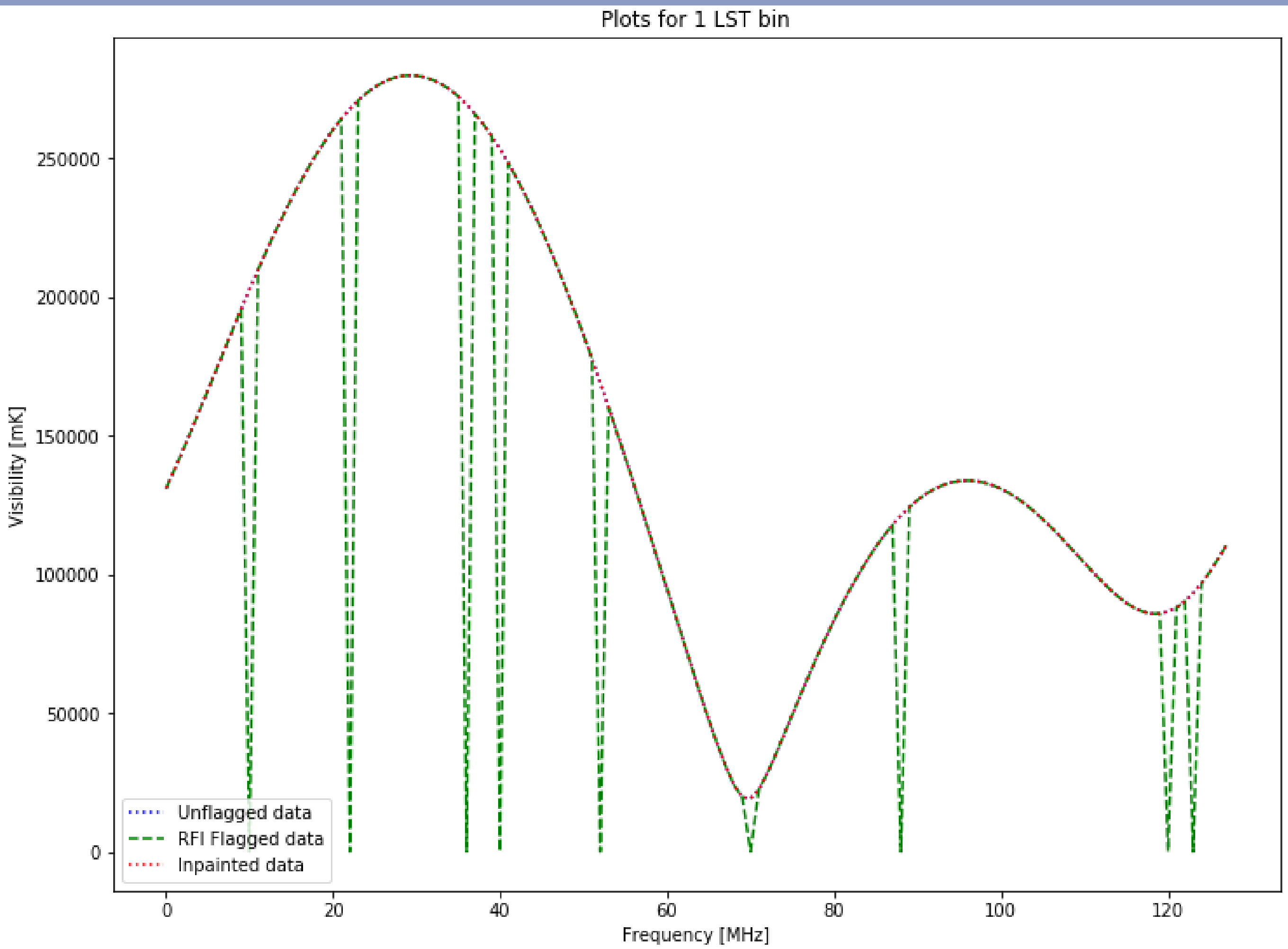


Comparing Visibility data for 1 LST bin

Visibilities for the same LST (Time) bin from un-flagged, flagged and in-painted data are shown.

The in-painted data has 'filled-in' the gaps in the RFI flagged regions.

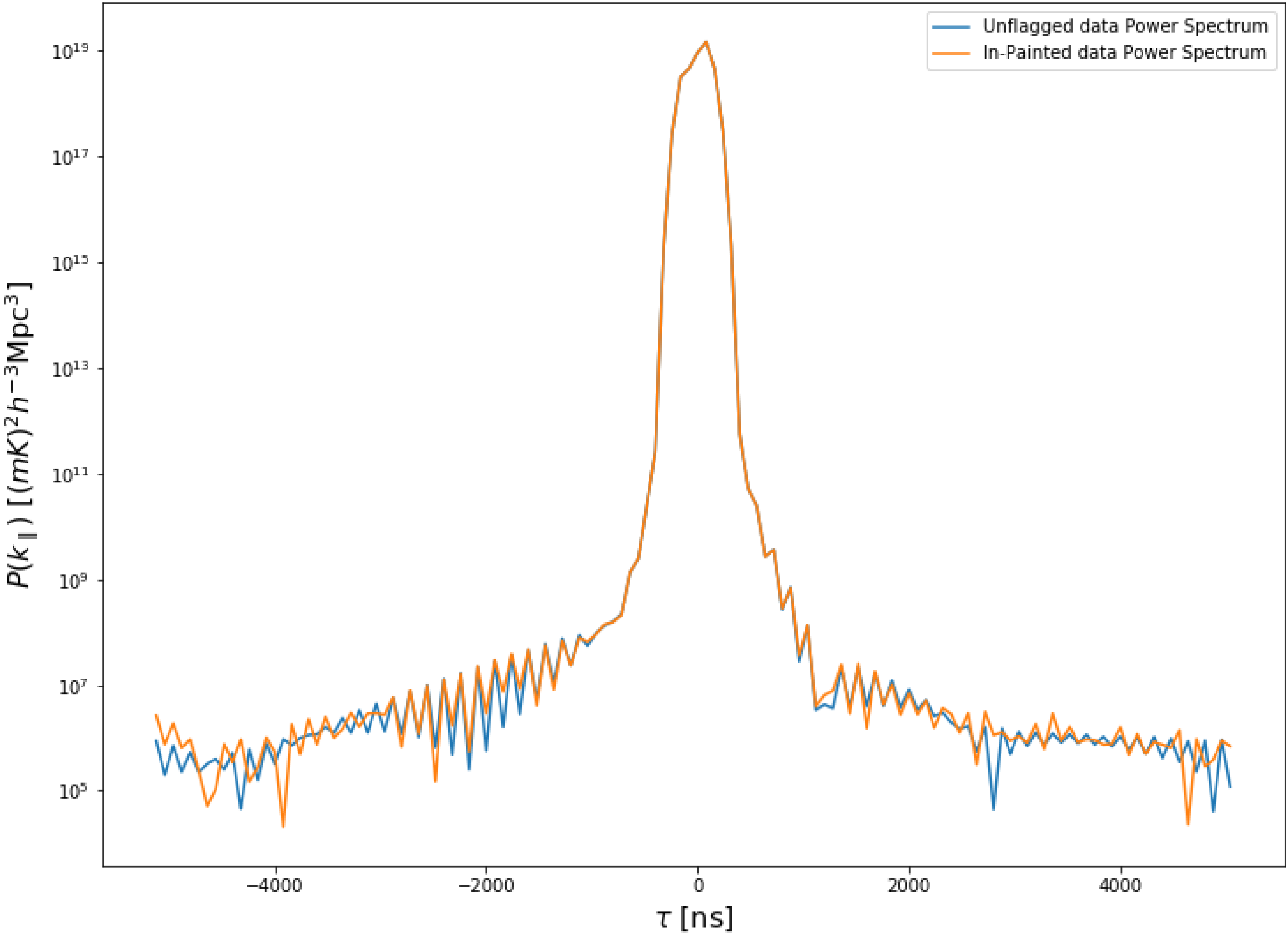
Shows that the in-painting process is able to recover the smooth features of the original unflagged ideal data from the RFI flagged data.



Results

Delay power spectrum [$mK^2\ h^{-3}\ Mpc^3$] averaged over Time and Baseline for in-painted and un-flagged data.

Power spectrum after fitting the smooth model recovers the smooth features of ideal case scenario well.



Conclusion

- Power spectrum comparison for unflagged and in-painted data shows the model performs well. Smooth features of power spectrum recovered.
- Differences in the two datasets are within the noise levels.
- RFI mitigation method used in this project works well for cases where fraction of RFI flagged frequencies is < 15%
- Excess power increases as number of flagged channels increase.

Ref: https://github.com/Ankurdev-astro/HERA-RFI

https://github.com/HERA-Team/hera_pspec

https://github.com/HERA-Team/hera_sim

Appendix

Instrument Design Specification	*Observational Performance*
Element Diameter: 14 m	**Field of View:** 9°
Minimum Baseline: 14.6 m	**Largest Scale:** 7.8°
Maximum Core Baseline: 292 m	**Core Synthesized Beam:** $25'$
Maximum Outrigger Baseline: 876 m	**Outrigger Synthesized Beam:** $11'$
EOR Frequency Band: 100–200 MHz	**Redshift Range:** $6.1 < z < 13.2$
Extended Frequency Range: 50–250 MHz	**Redshift Range:** $4.7 < z < 27.4$
Frequency Resolution: 97.8 kHz	**LoS Comoving Resolution:** 1.7 Mpc (at $z = 8.5$)
Survey Area: $\sim 1440\ \mathrm{deg}^2$	**Comoving Survey Volume:** $\sim 150\ \mathrm{Gpc}^3$
$\mathbf{T_{sys}}$: $100 + 120(\nu/150\ \mathrm{MHz})^{-2.55}$ K	**Sensitivity after 100 hrs:** $50\ \mu\mathrm{Jy\ beam}^{-1}$

Table: From DeBoer et al. (2017) , HERA-350 design parameters and their observational consequences.

Thank You!