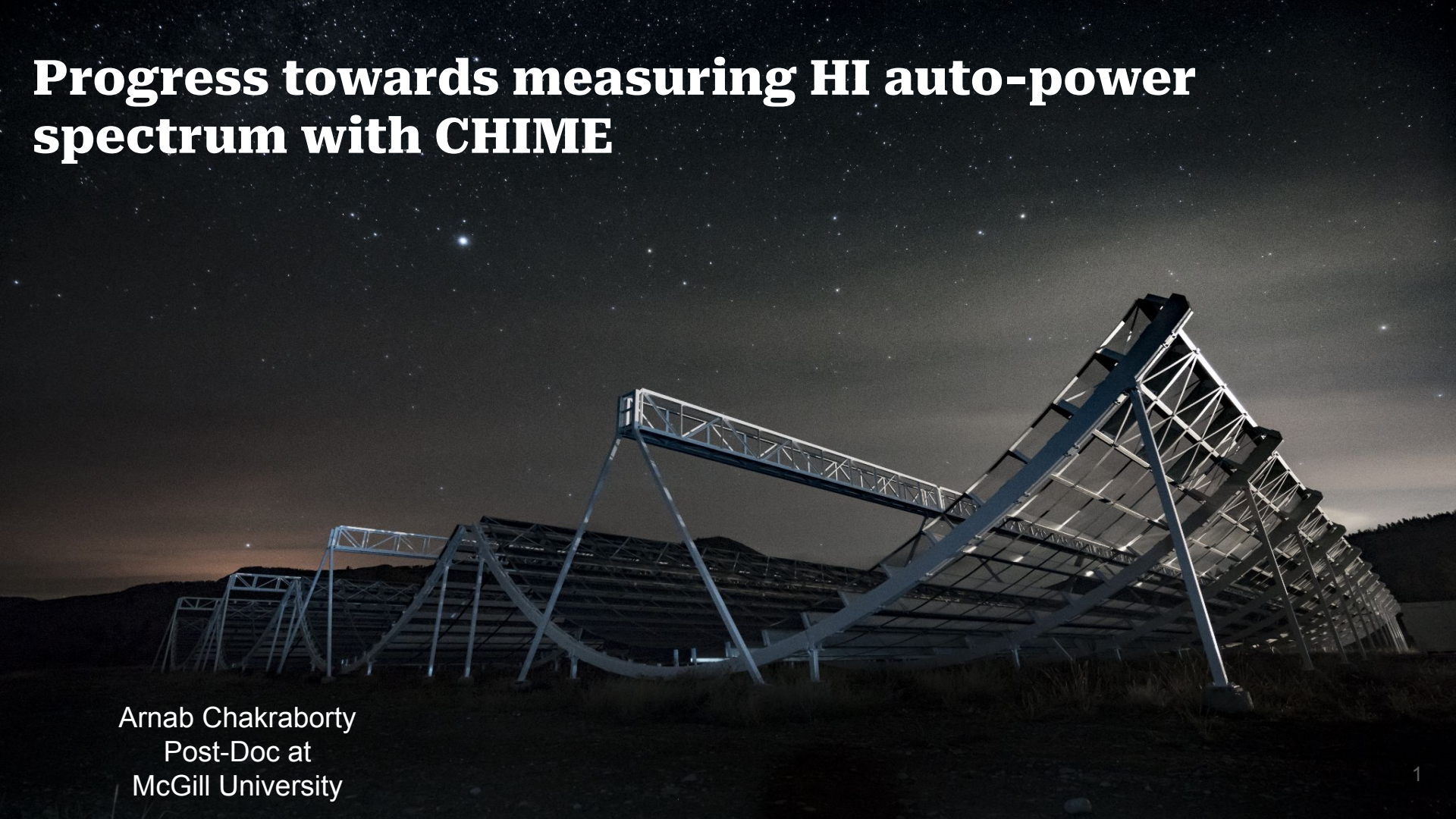
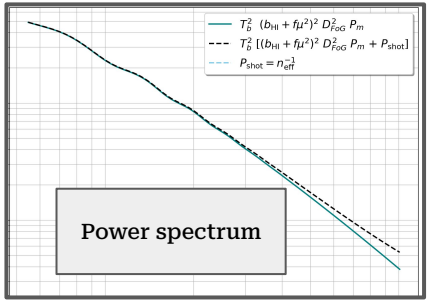
# Progress towards measuring HI auto-power spectrum with CHIME



Arnab Chakraborty
Post-Doc at
McGill University

# Outline


Improvements in the Data processing pipeline




Future Outlook


a collaboration between





with partners at



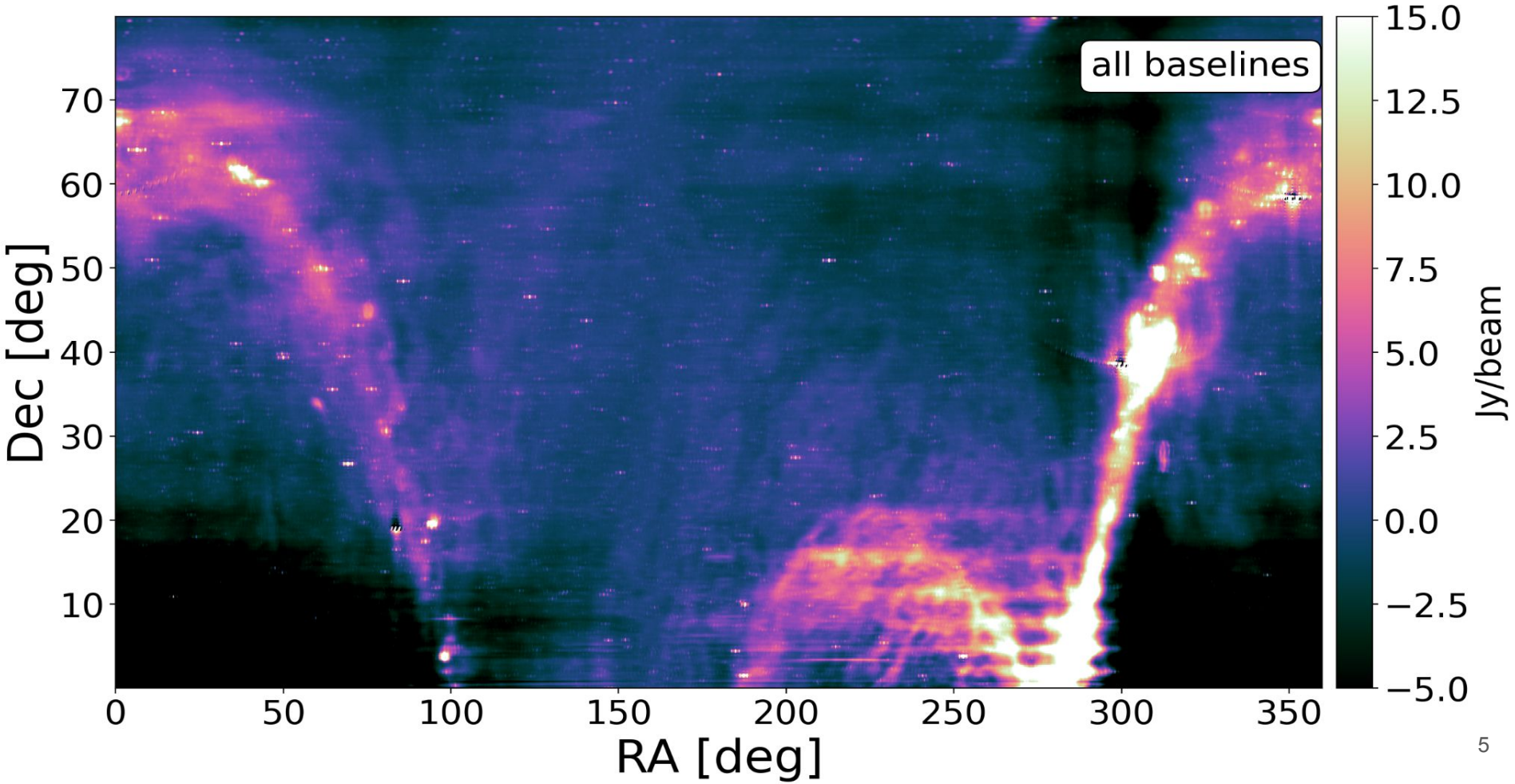



# Canadian Hydrogen Intensity Mapping Experiment (CHIME)

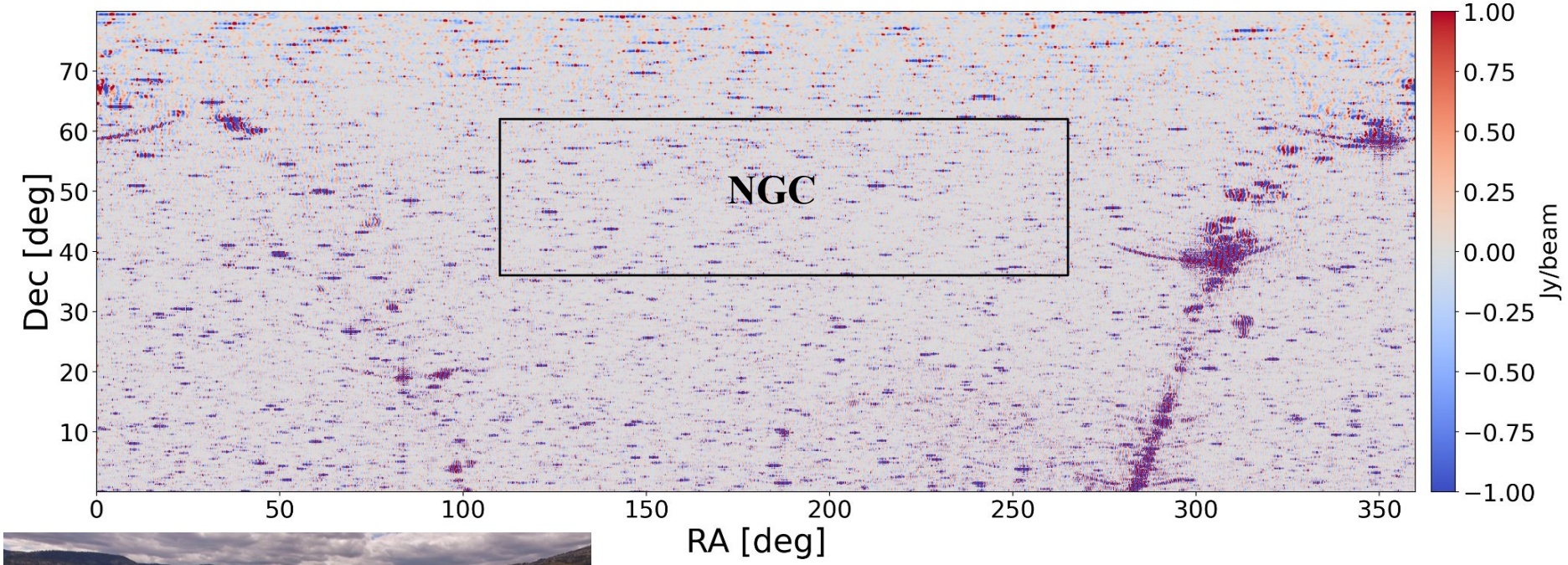
- Transit radio interferometer (no moving parts).
- 4 cylindrical reflectors, 20m x 100m each.
- Observe between 400-800 MHz, corresponding to redshift of 2.5-0.8 at the 21 cm line.
- 1024 dual polarized antennas.
- Maps the northern radio sky every day.

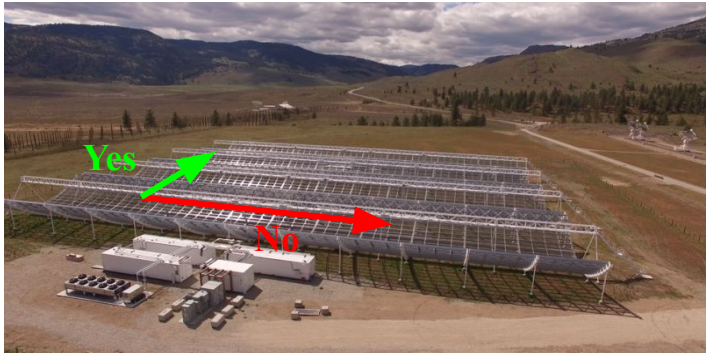
Beam:
~120 x 2 deg

# Map of the sky at a single frequency using all baselines only



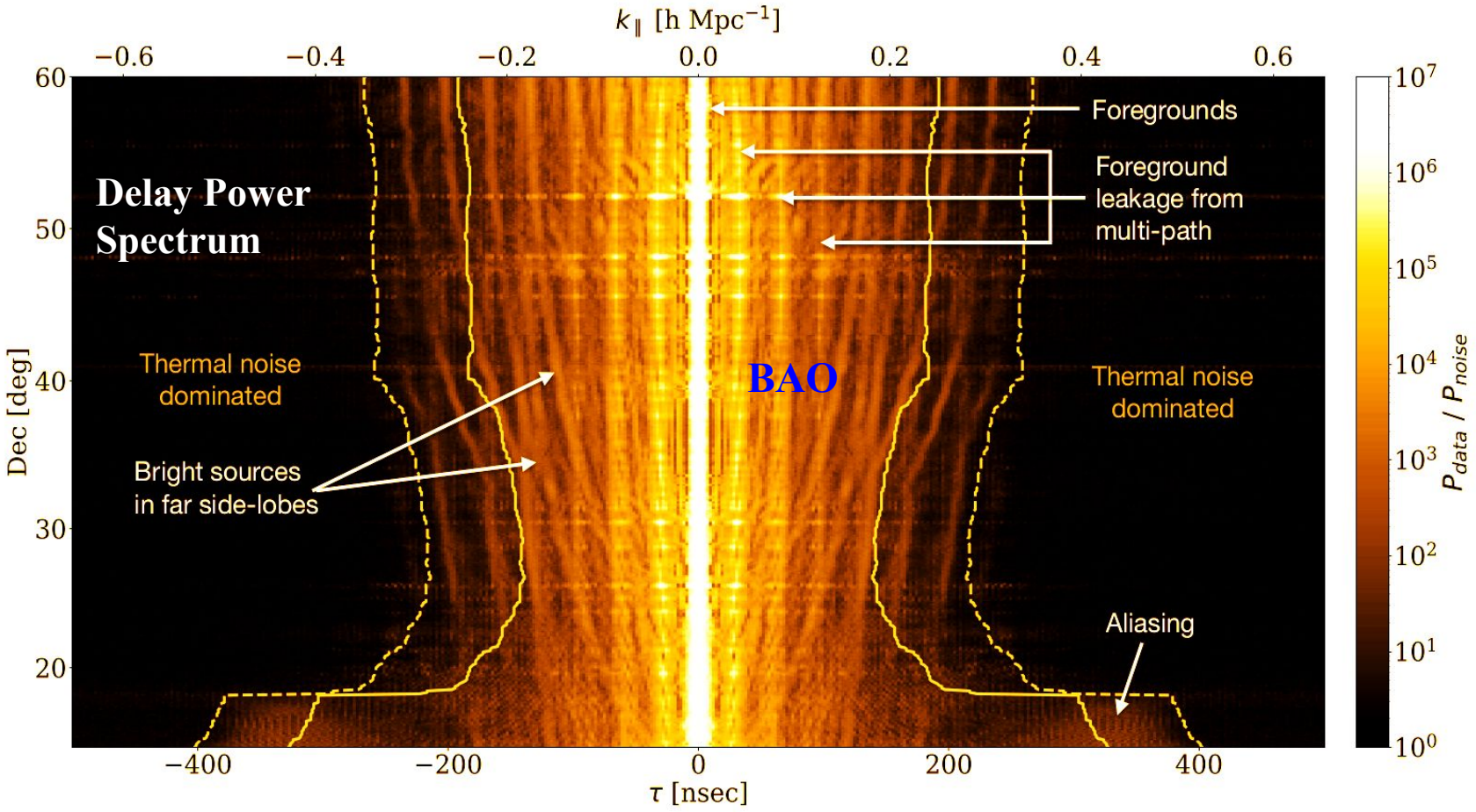
# Map of the sky at a single frequency using inter-cylinder baselines only



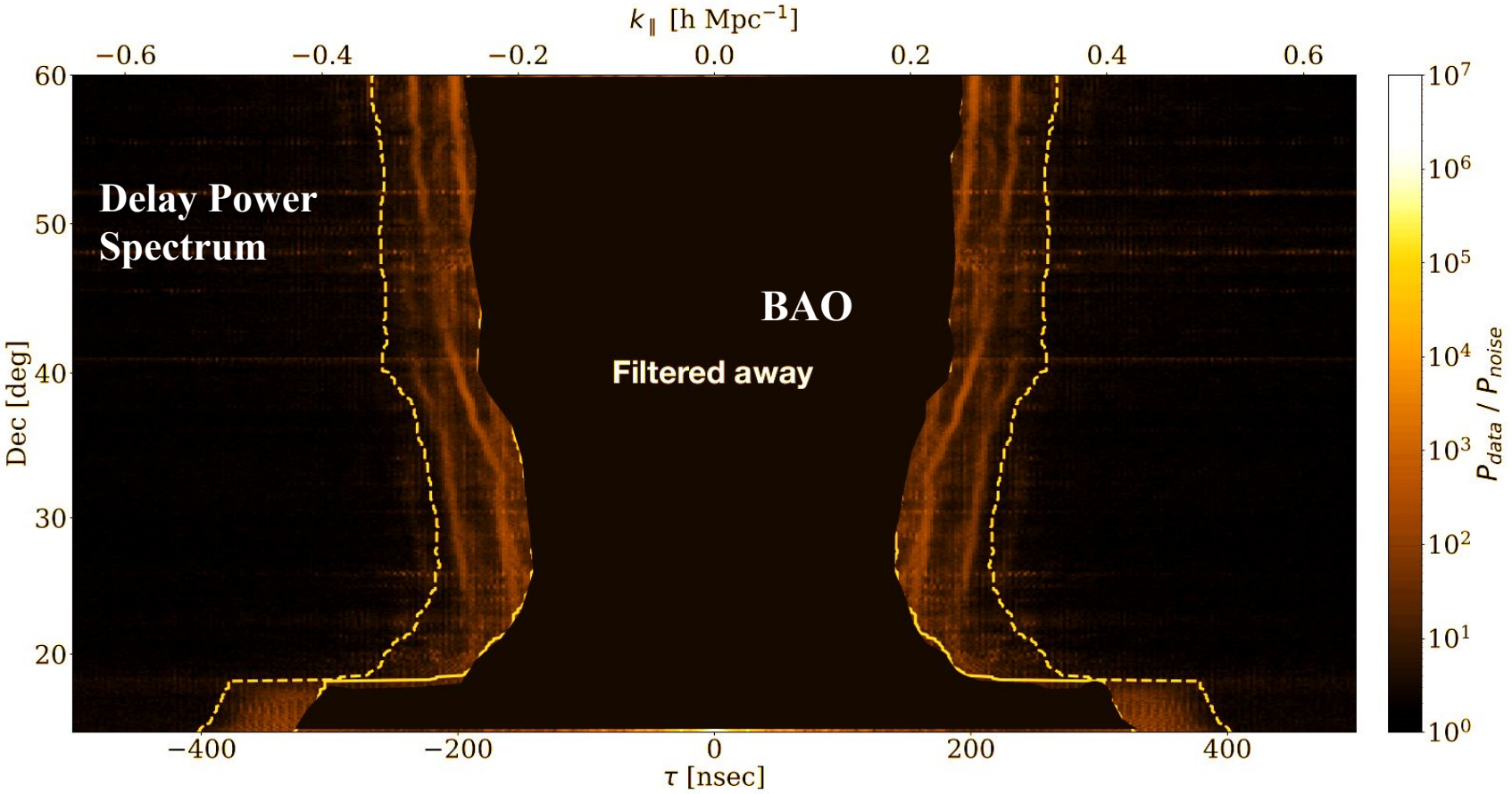




The map constructed using inter-cylinder baselines only, which resolve out diffuse Galactic emission and leaving primarily emission from extragalactic point sources.

# Delay power spectrum of the CHIME map : Variance over RA

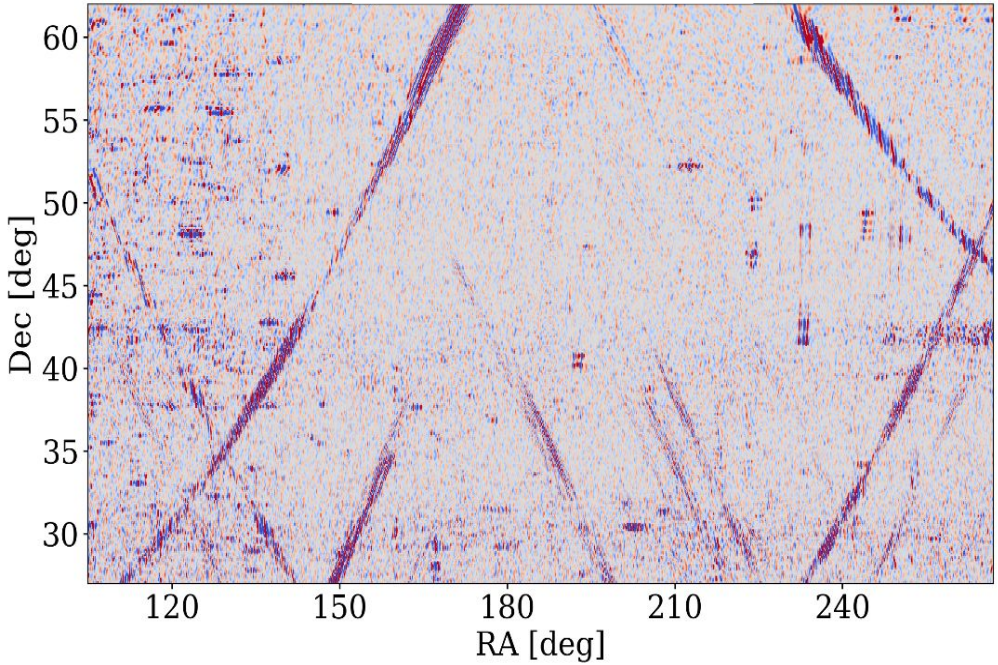


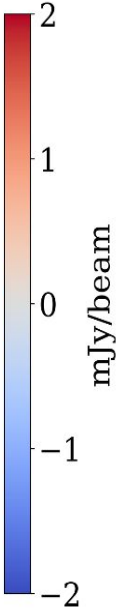
# Delay power spectrum of the CHIME map : Variance over RA



# Progress on data analysis to clean the CHIME data to measure Auto-Power Spectrum

Data cube from cross-corr : **2022**




**CHIME col. 2022**

# Progress on data analysis to clean the CHIME data to measure Au

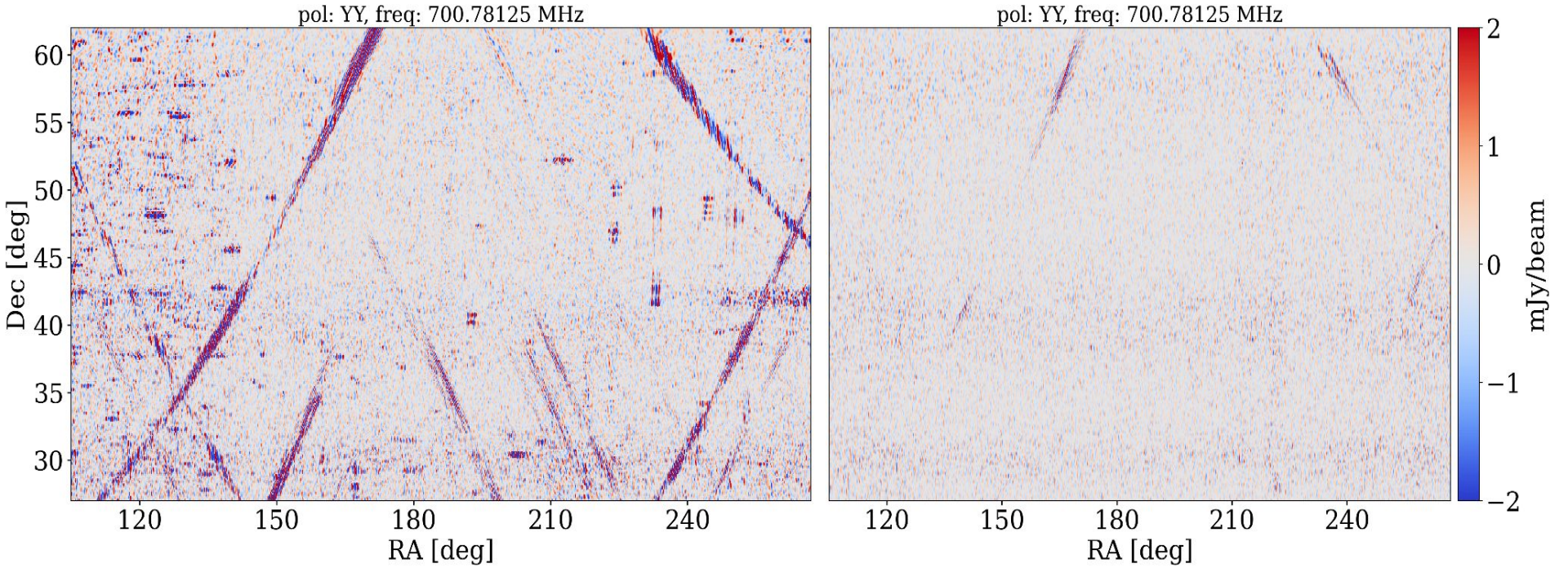


Liam Gray, UBC

Seth Siegel, Perimeter/McGill

Data from cross-corr : **2022**

Data after new processing: **10/2024**




**How we go from left to right ?**

# Improved analysis pipeline to measure the auto-powerspectrum of HI signal with CHIME

**Analysis pipeline**

## RFI masking algorithms

- Masking transient RFI
- Identify outliers in the "system radiometer test"
- Identify outliers in FG filtered data - $\chi^2 - \text{test}$
- Flag time-ranges with large amount of rainfall.
- Identify bad days and remove it from stack.

**Foreground filtering the raw time series data of individual days**

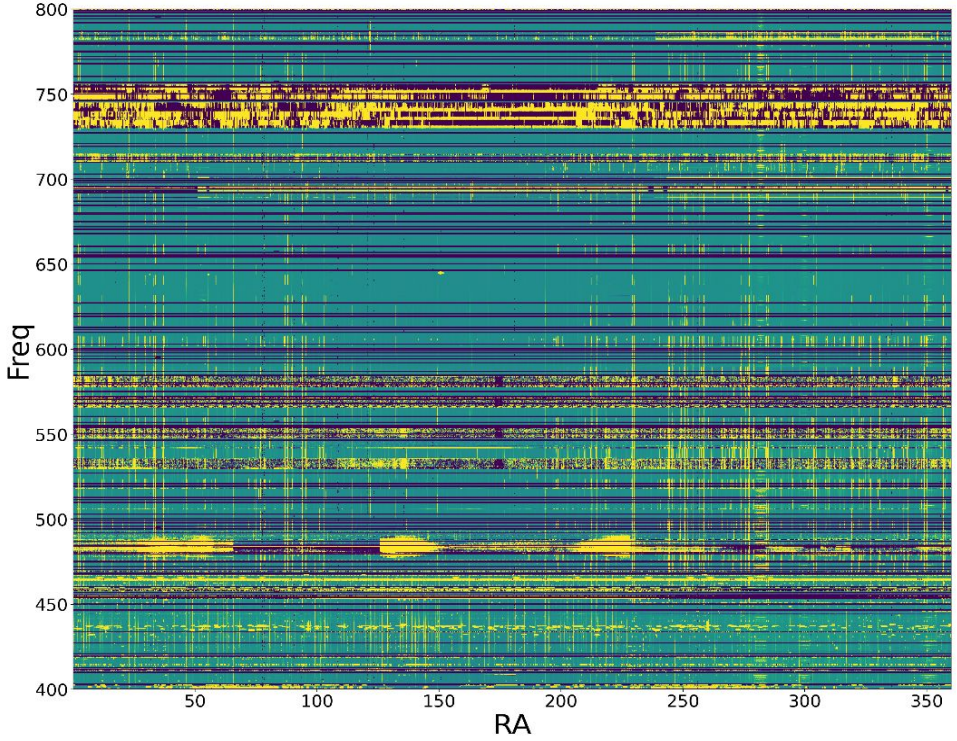
# RFI flagging : Radiometer test

- Identify outliers in the “system radiometer test”:

$$R_{t\nu} = \Delta t \Delta \nu \frac{\sum_{ij} \mathrm{Var}\left(V_{ij,t\nu}\right)}{\sum_{ij} V_{ii,t\nu}\, V_{jj,t\nu}}$$

Sample variance from even-odd difference at 30 msec

Expectation for variance based on radiometer equation



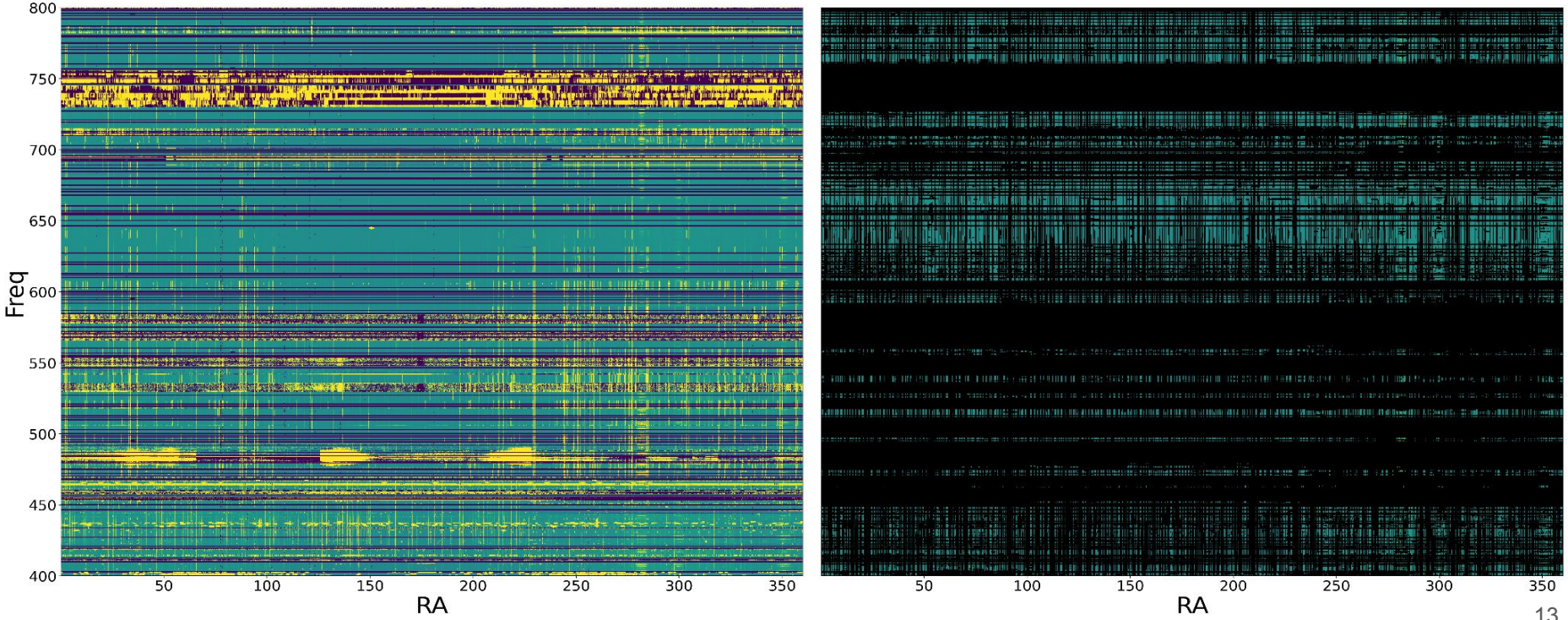
# RFI flagging : Radiometer test

- Identify outliers in the “system radiometer test”:

$$R_{t\nu} = \Delta t \Delta \nu \frac{\sum_{ij} \mathrm{Var}\left(V_{ij,t\nu}\right)}{\sum_{ij} V_{ii,t\nu}\, V_{jj,t\nu}}$$

Sample variance from even-odd difference at 30 msec

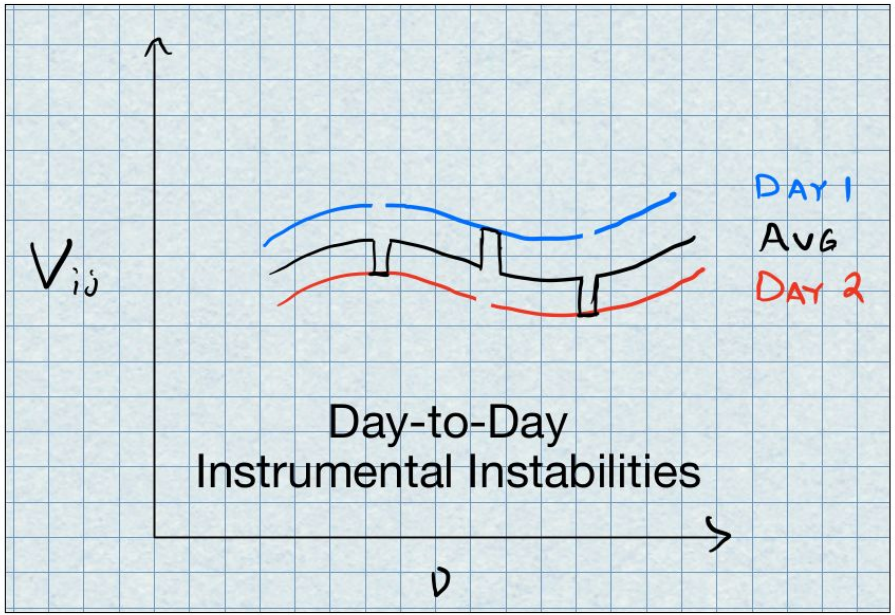
Expectation for variance based on radiometer equation




**71.23% total flagging for this particular night**

# Systematics : Subtle issues with data averaging

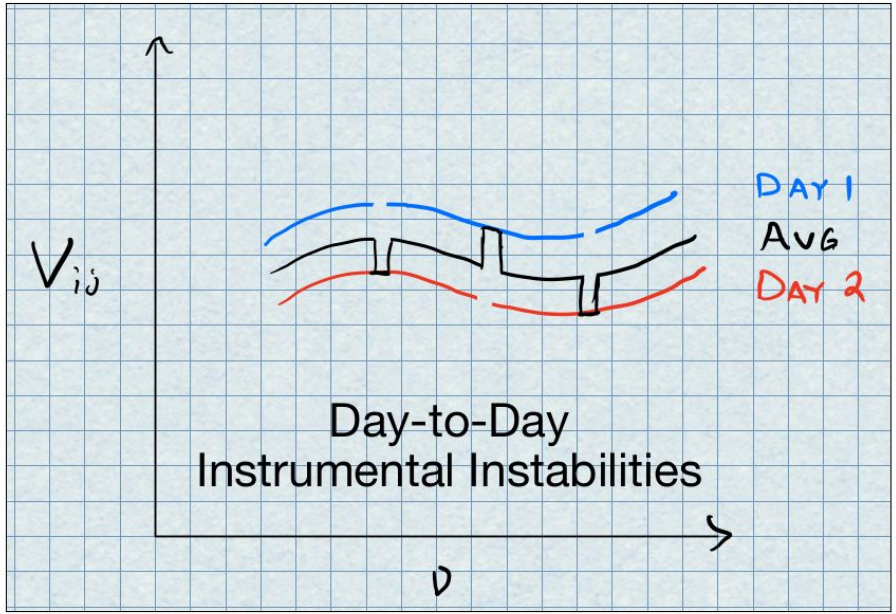
Integrating over visibilities that are changing as a function of time using frequency-dependent weighting leaks foreground power to high delays.

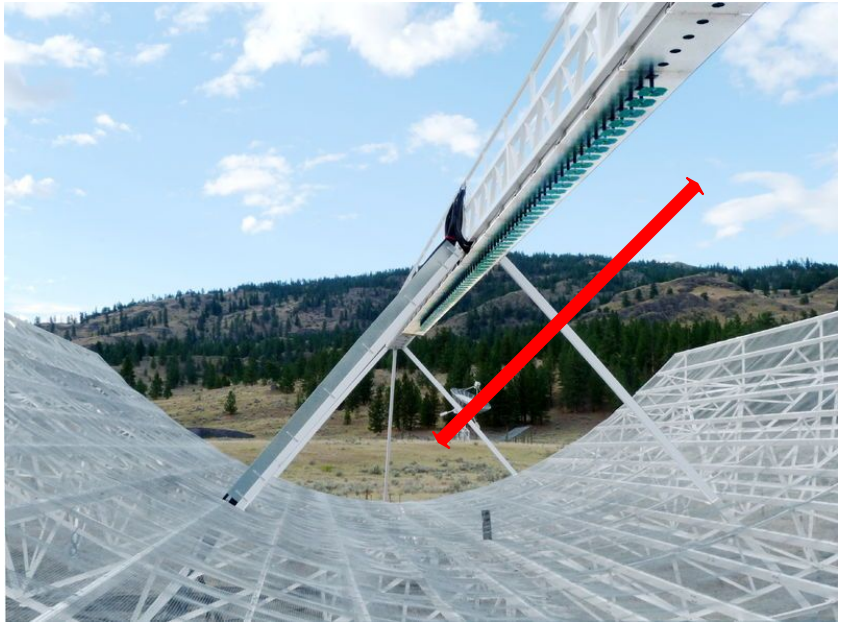



5

# Systematics : Subtle issues with data averaging

Integrating over visibilities that are changing as a function of time using frequency-dependent weighting leaks foreground power to high delays.

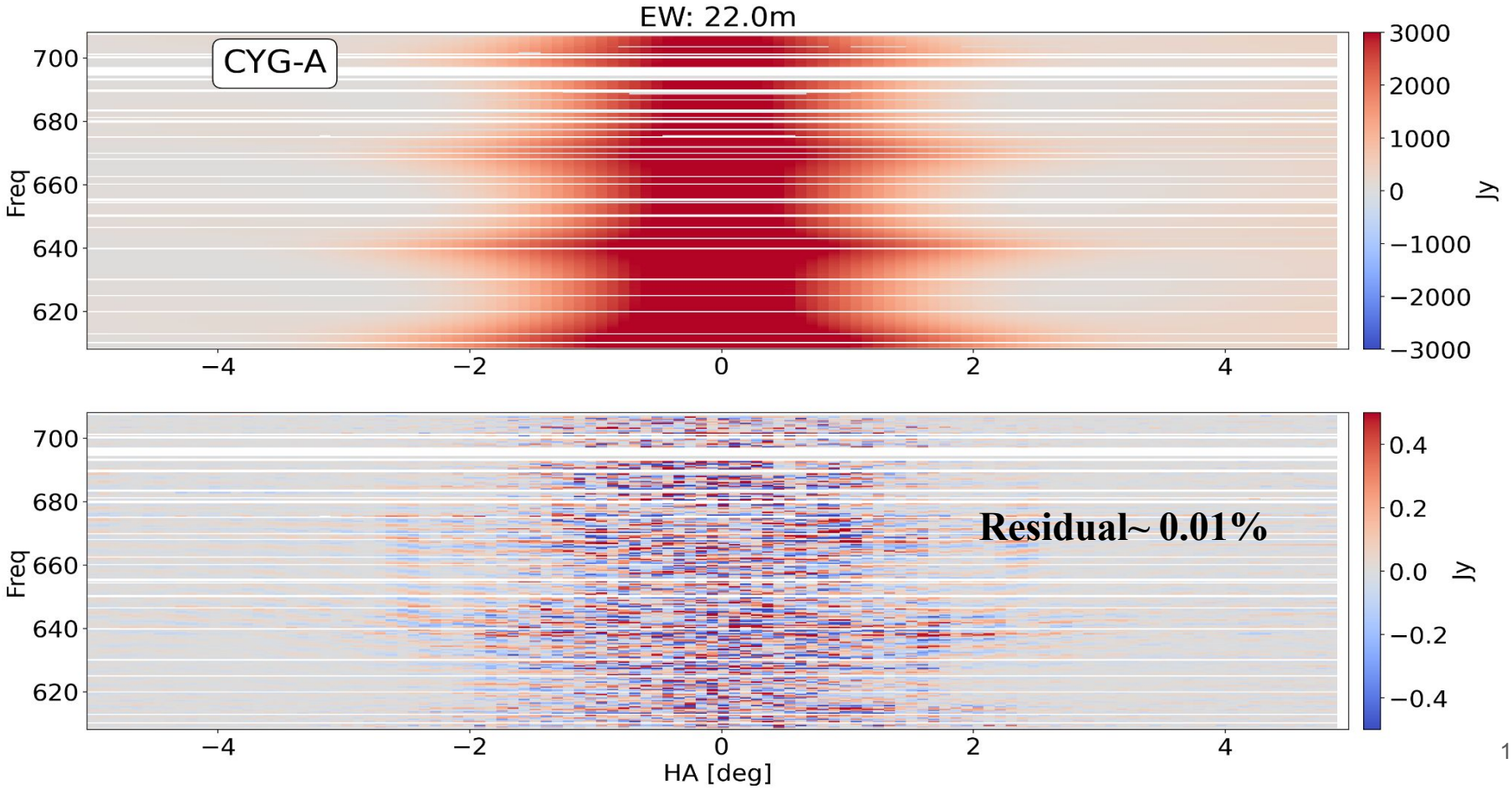


5

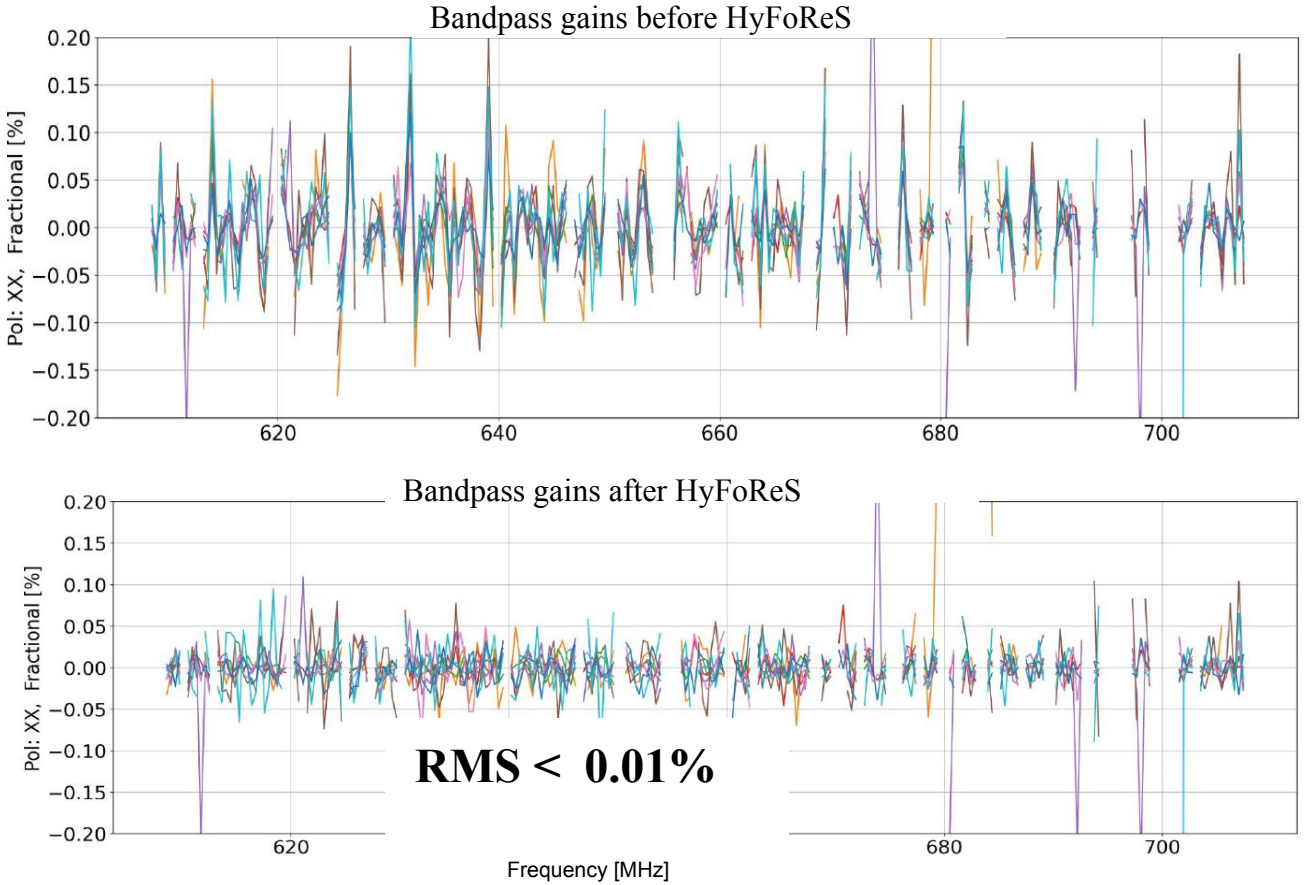
**Solution :** We take the **timestream data**, then beamformed along declination axis for each east-west baseline separation and use a constant FG filter cutoff for all declination and all baselines

# Foreground residual and quantification of residual error: Beamformed to CYG-A



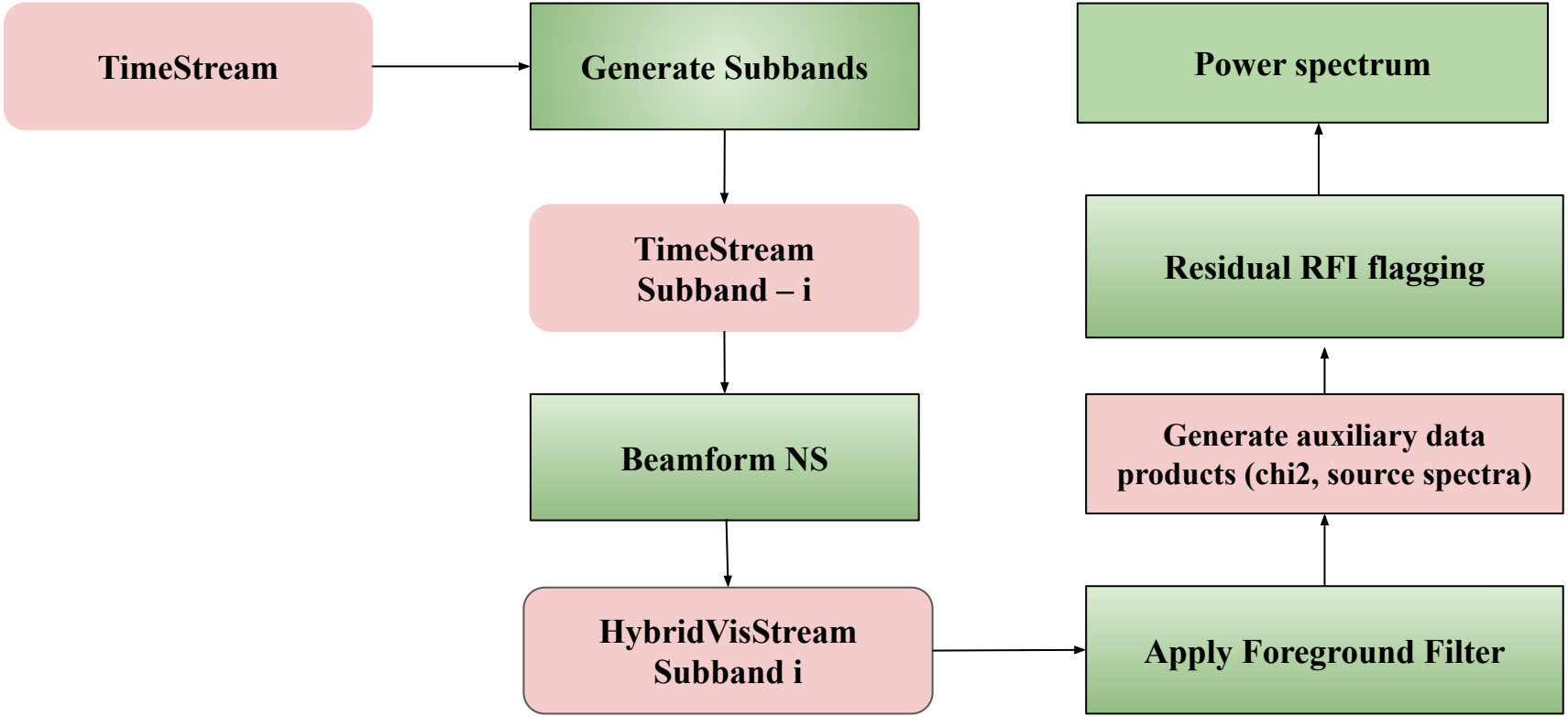
# Improvements on gain calibration: HyFoReS Wang, et. al (2022,2024) [arxiv: 2203.07184, 2408.08949]

Haochen Wang (MIT)

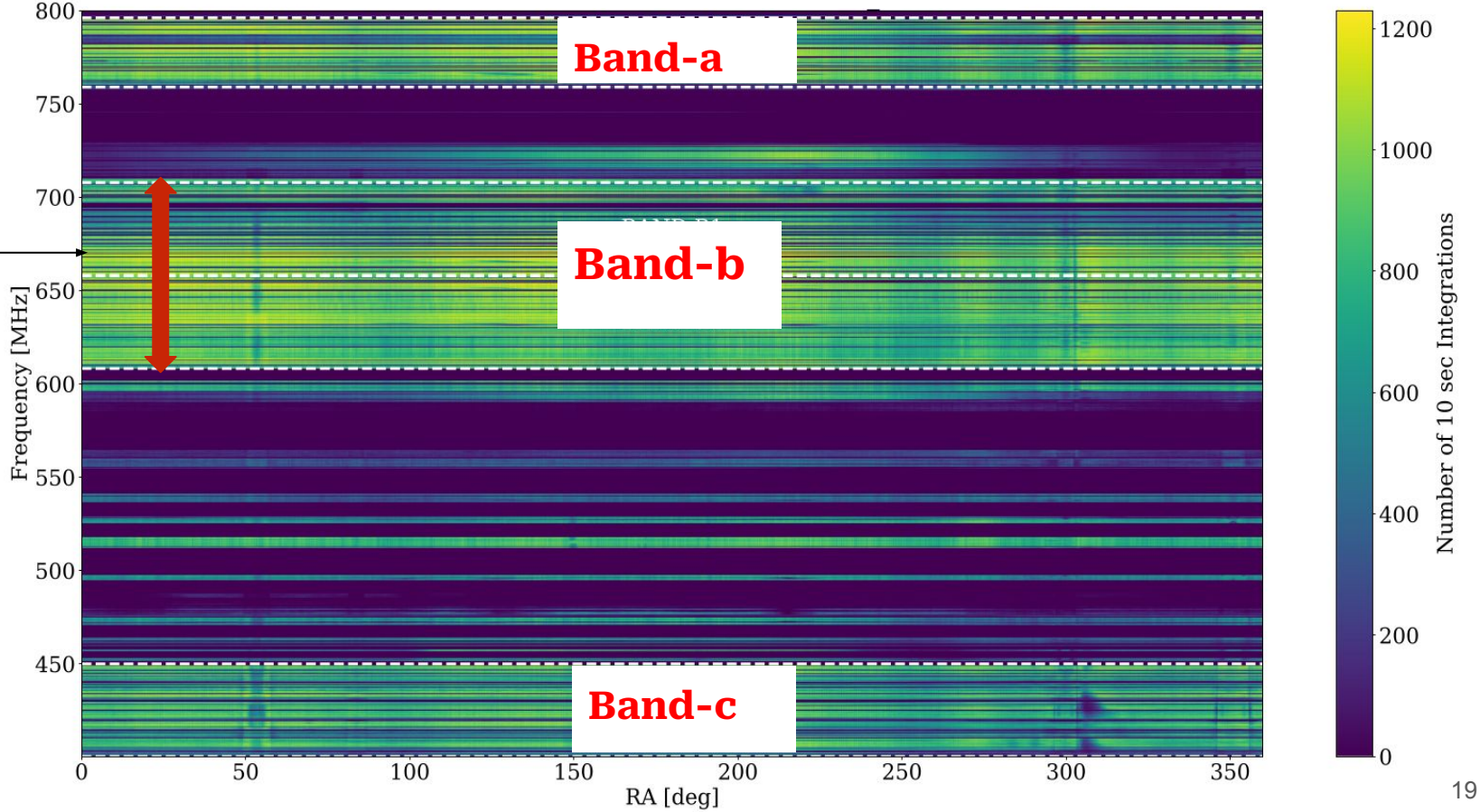



- Bandpass gains from 11 bright point sources
- **HyFoReS:** estimates gains by cross-correlating low-delay with high-delay data
- Reduced gain fluctuations by ~ a factor of 2

# Daily Foreground filter pipeline : Change the pipeline framework and write many new tasks



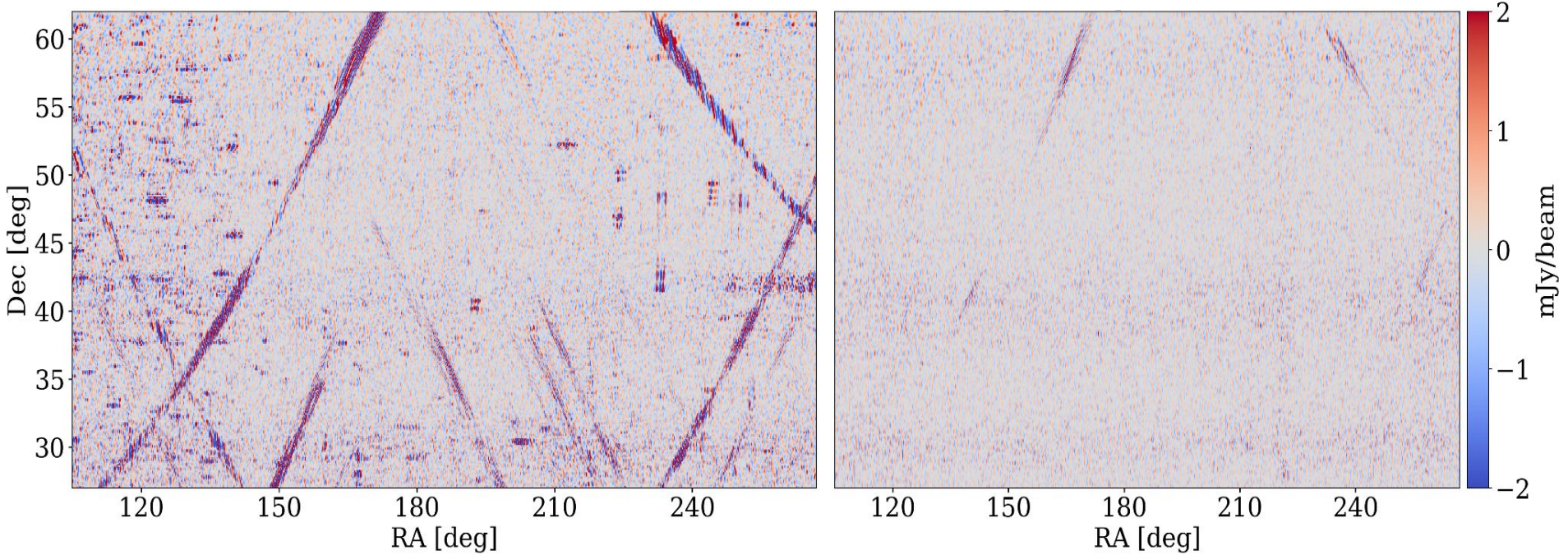
# Data available after flagging in 2019



# Progress on data analysis to clean the CHIME data to measure Auto-Power Spectrum

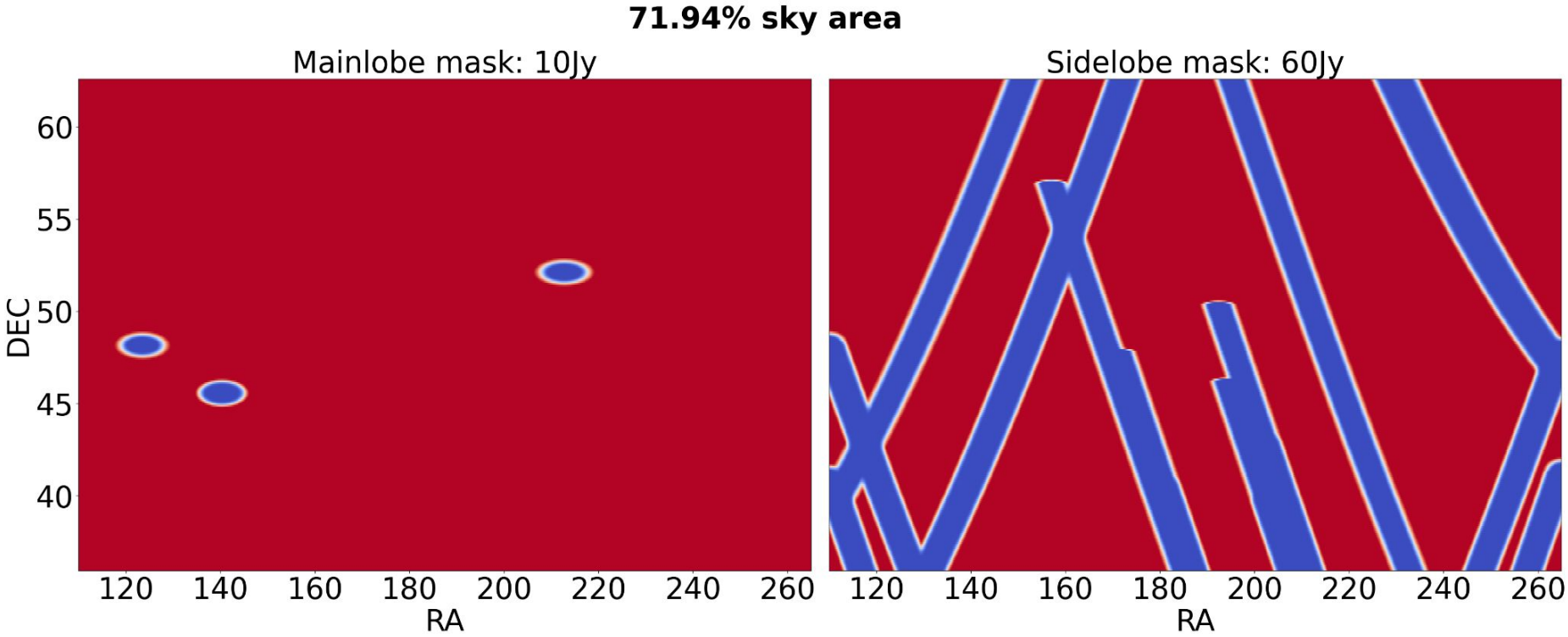
Data from cross-corr : **2022**

Data after new processing: **2025**

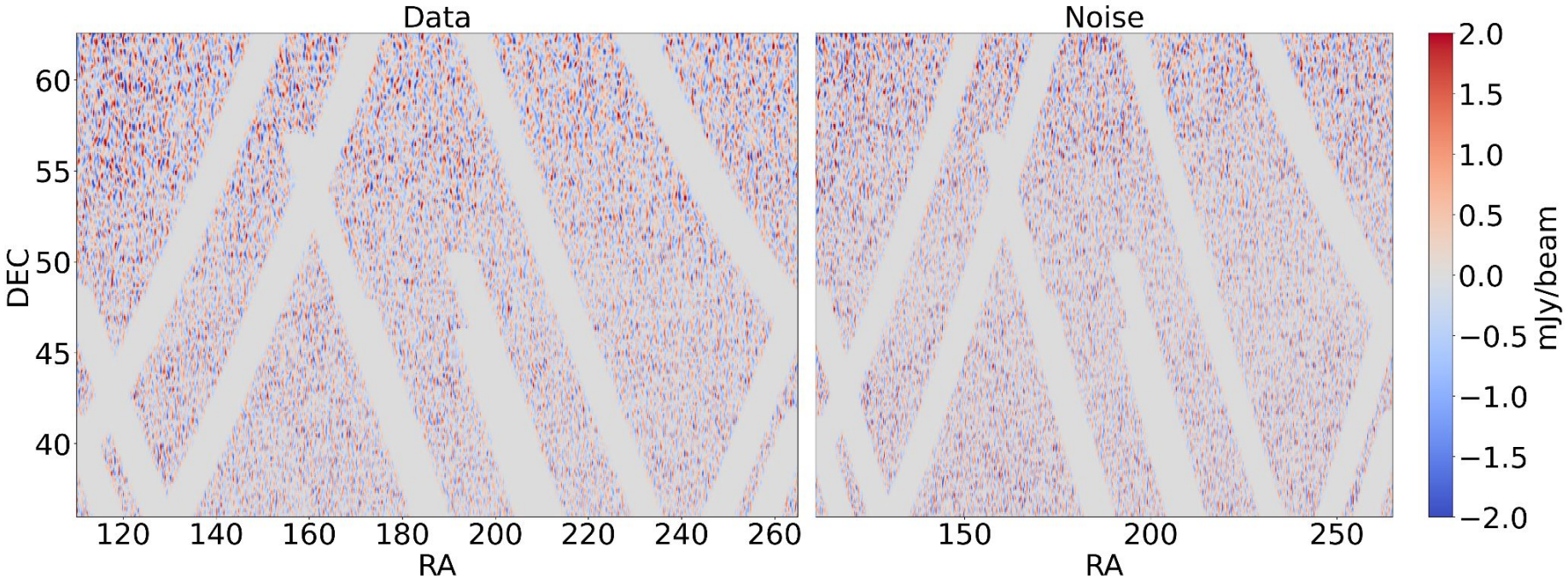



**After all these improvements we get this clean map, average of 94 nights from 2019**

# Applying a Spatial Mask for very bright point sources in the sky



# Comparison of data with noise after FG filtering and applying mask



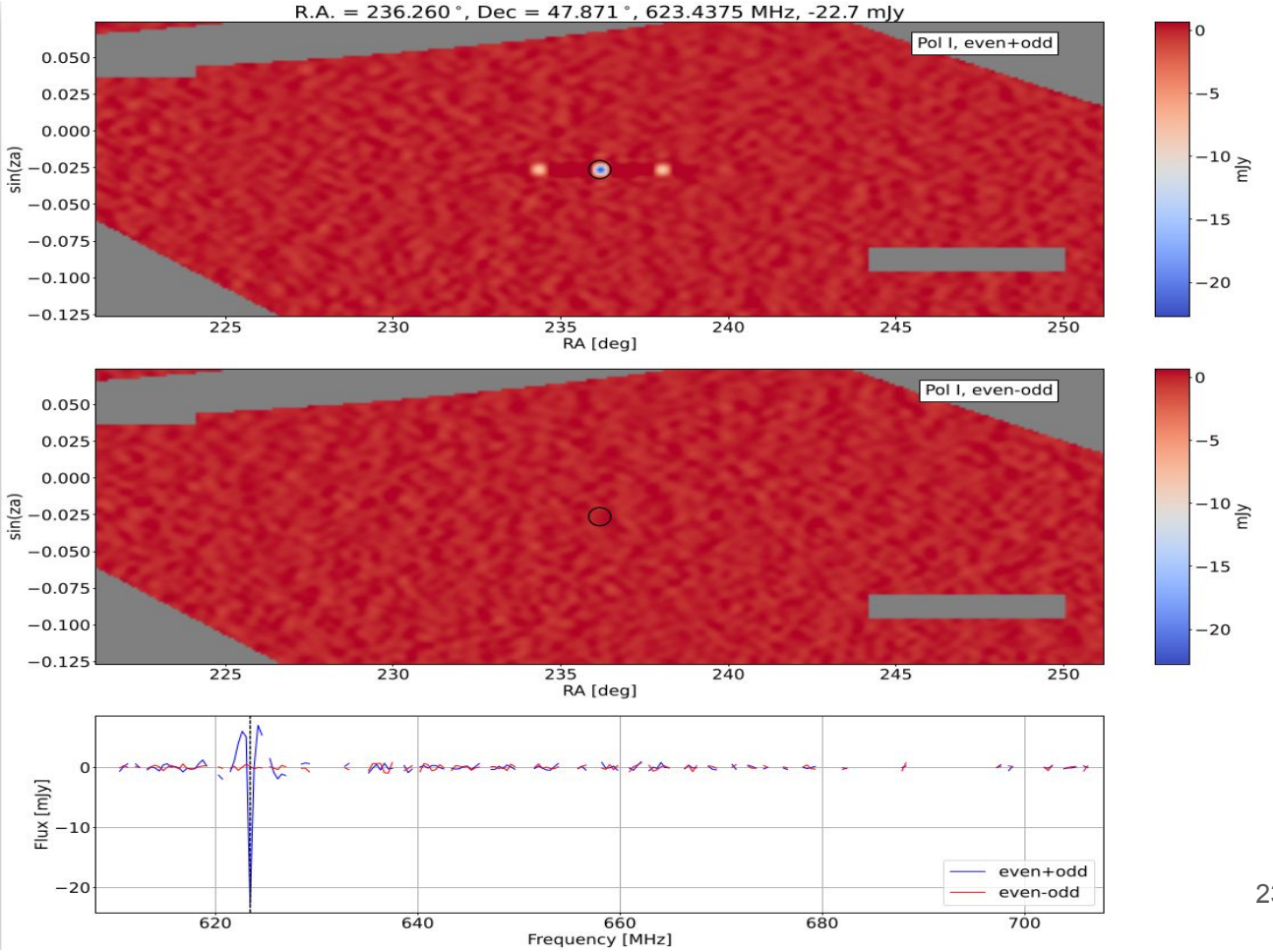

- We compare the data and thermal noise after applying FG filter and pixel mask. The thermal noise is estimated using fast-cadence method.

# Discovery of HI absorption system in the Cosmology data

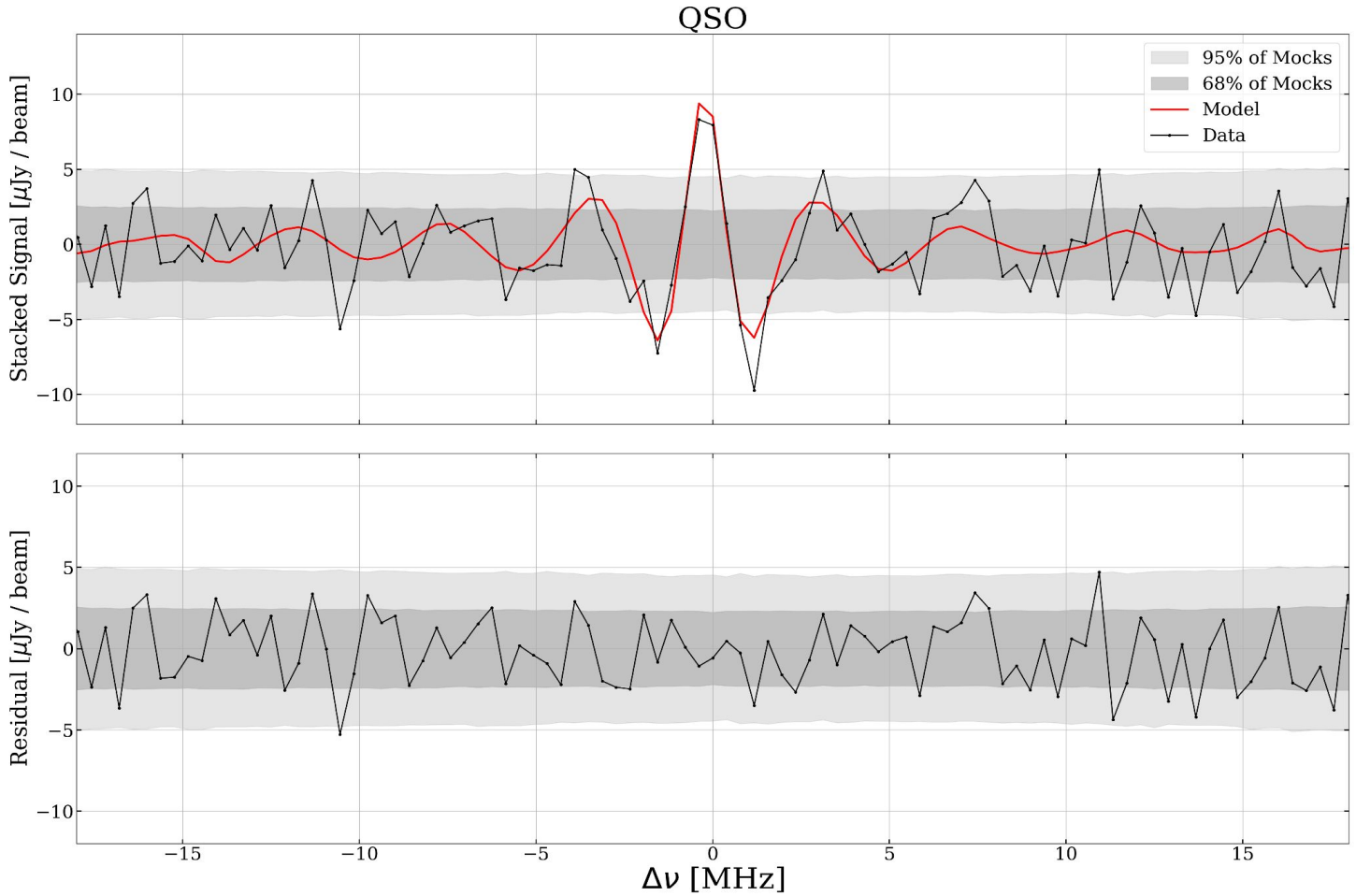
Credit: Seth Siegel

**Known absorber**

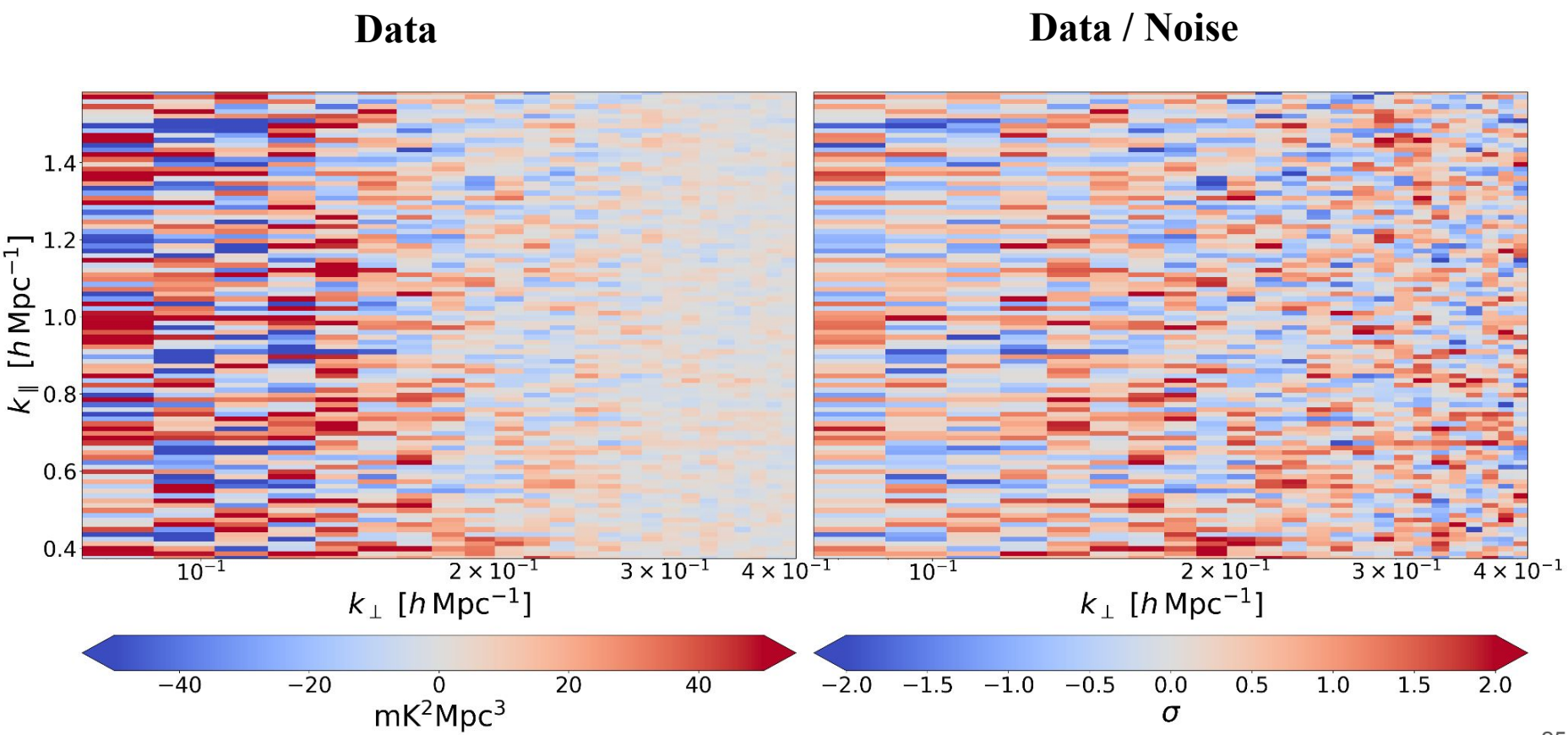
We have found many new absorbers in our band, which we are masking in this analysis and in parallel, we are also working on these new absorber systems.



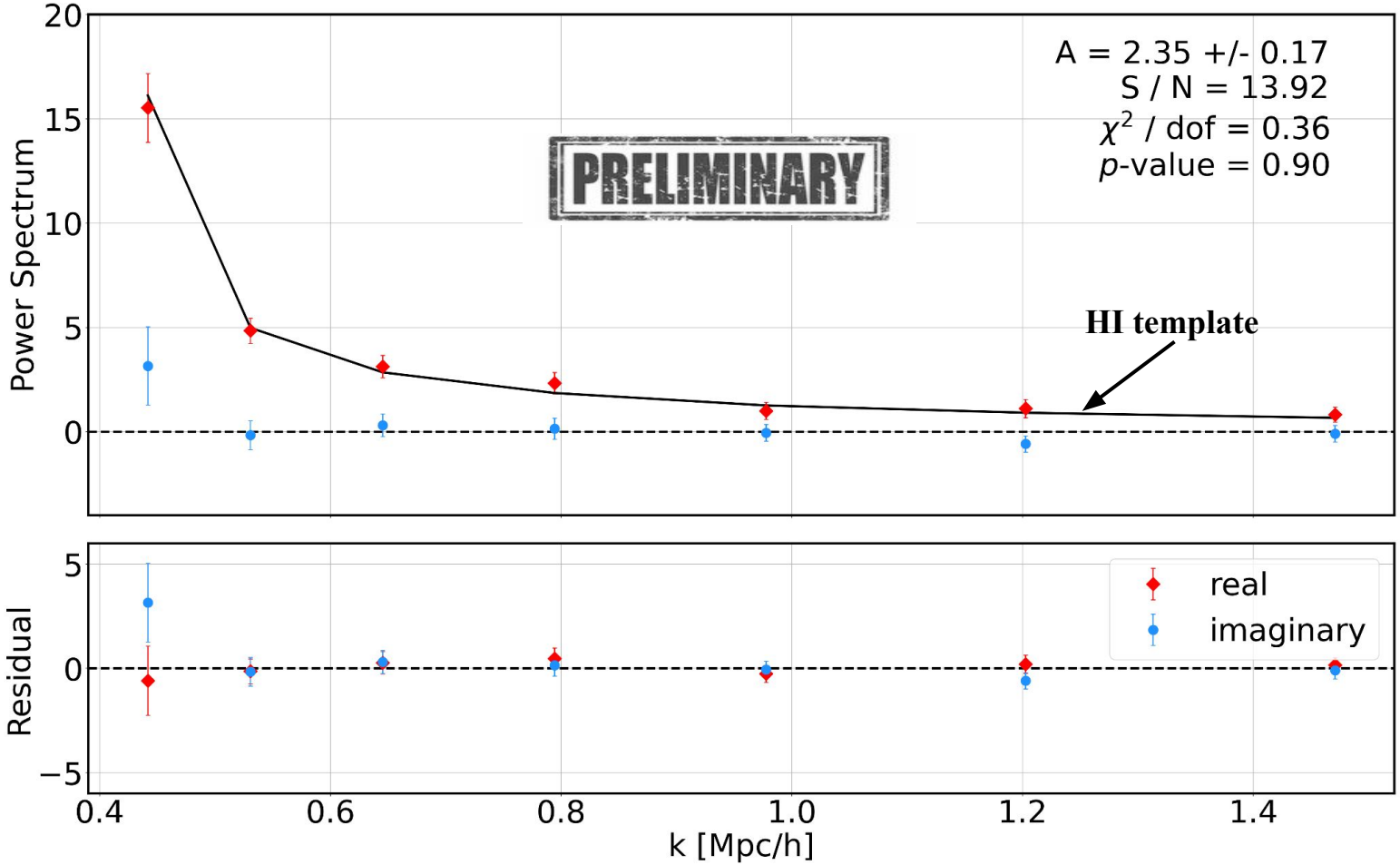
# Cross-Correlation : Stacking over 100 MHz sub-band



# 2D Cylindrically averaged power spectrum

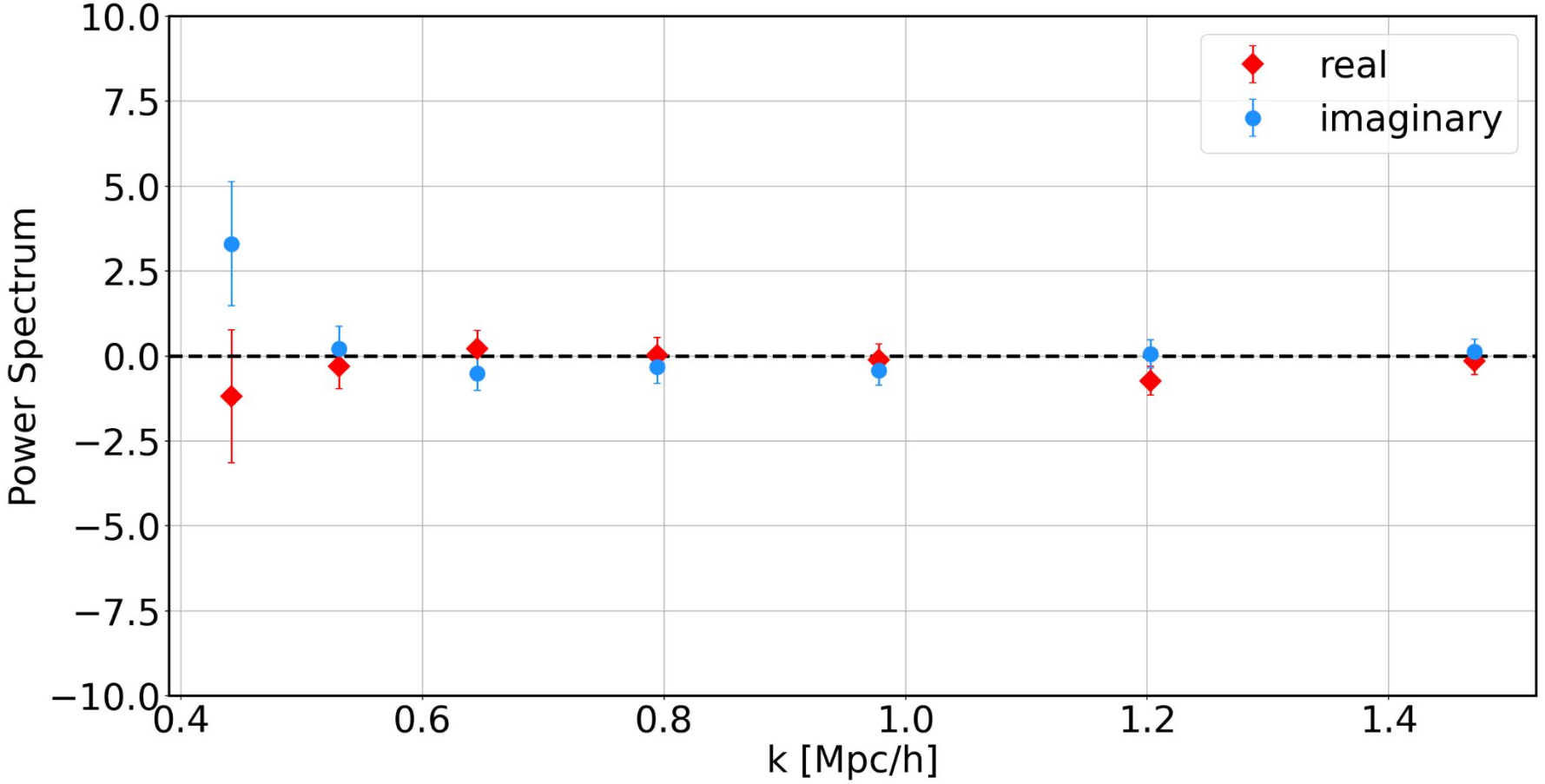


# Spherically averaged Power Spectrum at redshift ~ 1.16

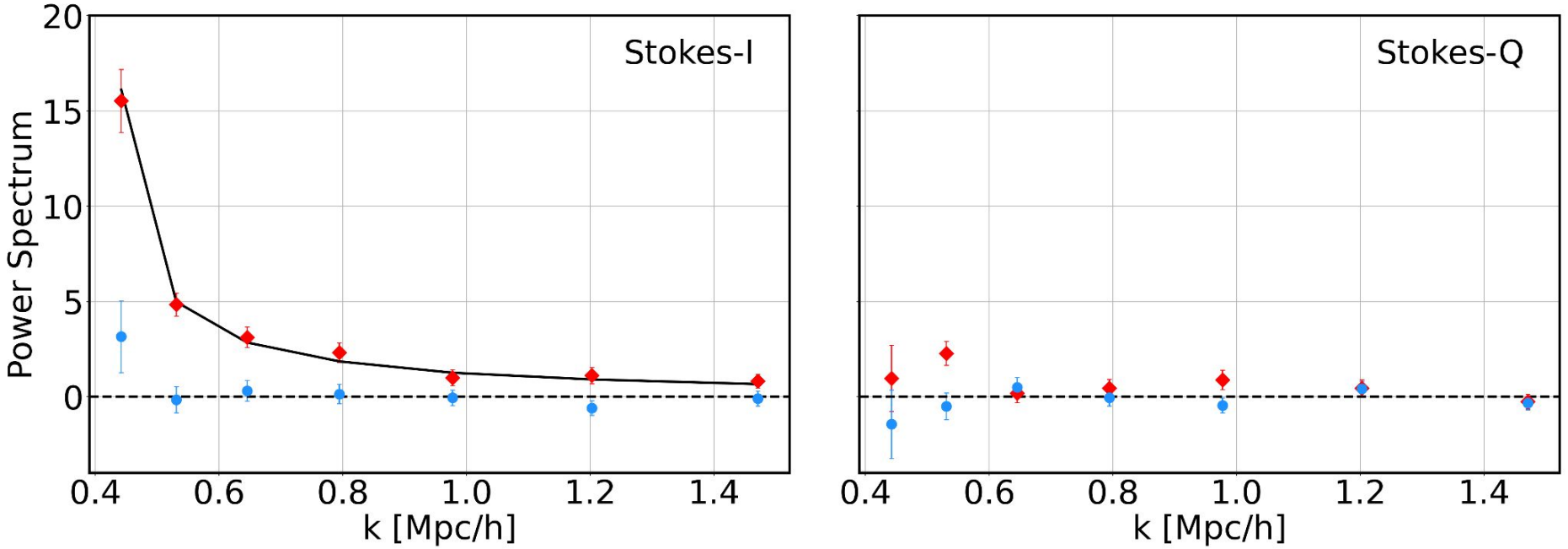


| Statistical Tests | Checks |
| --- | --- |
| Varying the fluxcut of sources to generate mask - mainlobe mask.<br>• If we see FG residual, then lowering the fluxcut will also lower the power spectrum. So we change a varying power as a function of fluxcut.<br>• If we do not see FG residual: then there should not be any varying power as a function of fluxcut | ✗<br>✓ |
| Varying the fluxcut of sources to generate mask - sidelobe mask.<br>• We may expect to see sidelobes of faint source (not seen in the map). So lowering the fluxcut of sidelobe mask will also lower the power spectrum.<br>• If we do not see any such variation in the power with fluxcut of sidelobes, then we can rule out the hypothesis of sidelobe FG leakage. | ✗<br>✓ |
| Divide the RA range into two separate RA bins<br>• If systematic is localized in LSTs, then that will show up in power spectrum. Also, FG sky is changing as a function of LST, but HI is same. So, if it is FG, then we will expect different power at different bins.<br>• If it is not systematic or FG residual, then we should expect consistent power spectrum in each bin | ✗<br>✓ |
| Power spectrum of even-odd partitions<br>• The transient RFI and day-to-day change will remain in this difference and will give excess variance in the power spectrum<br>• However, cosmological signal should drop out, and we should see power spectrum consistent with noise | ✗<br>✓ |

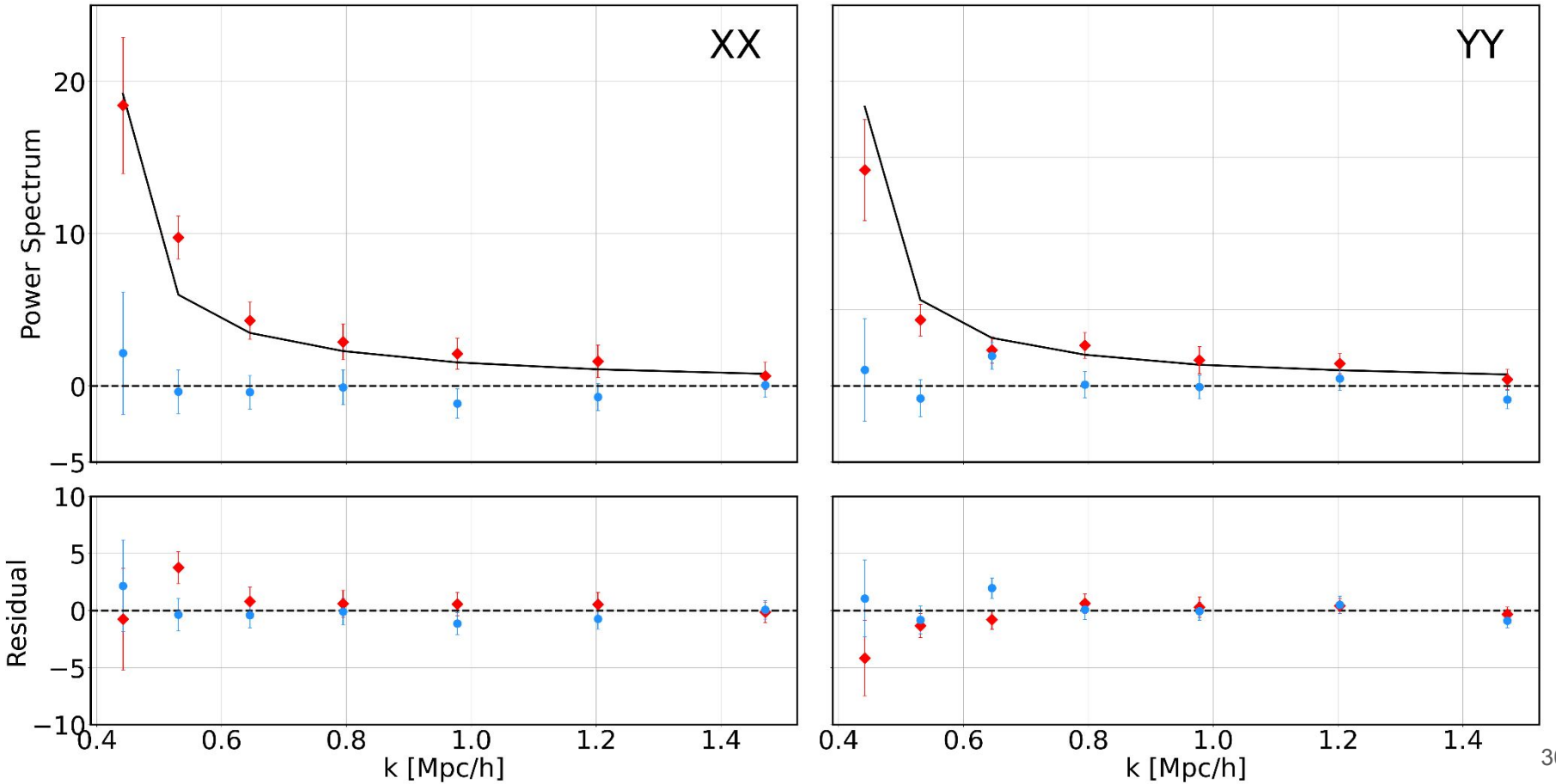
# Jackknives: Powerspectrum of (Even - Odd )



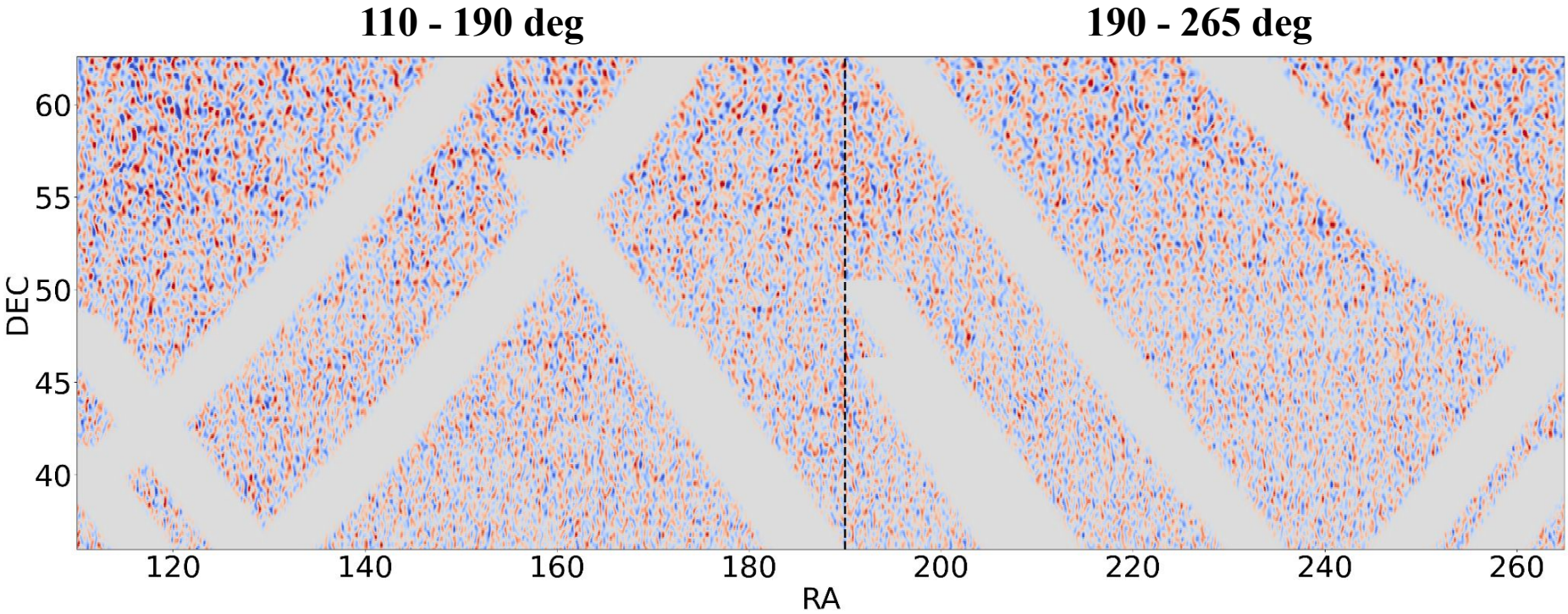
# Polarization consistency : Stokes-I and Stokes-Q



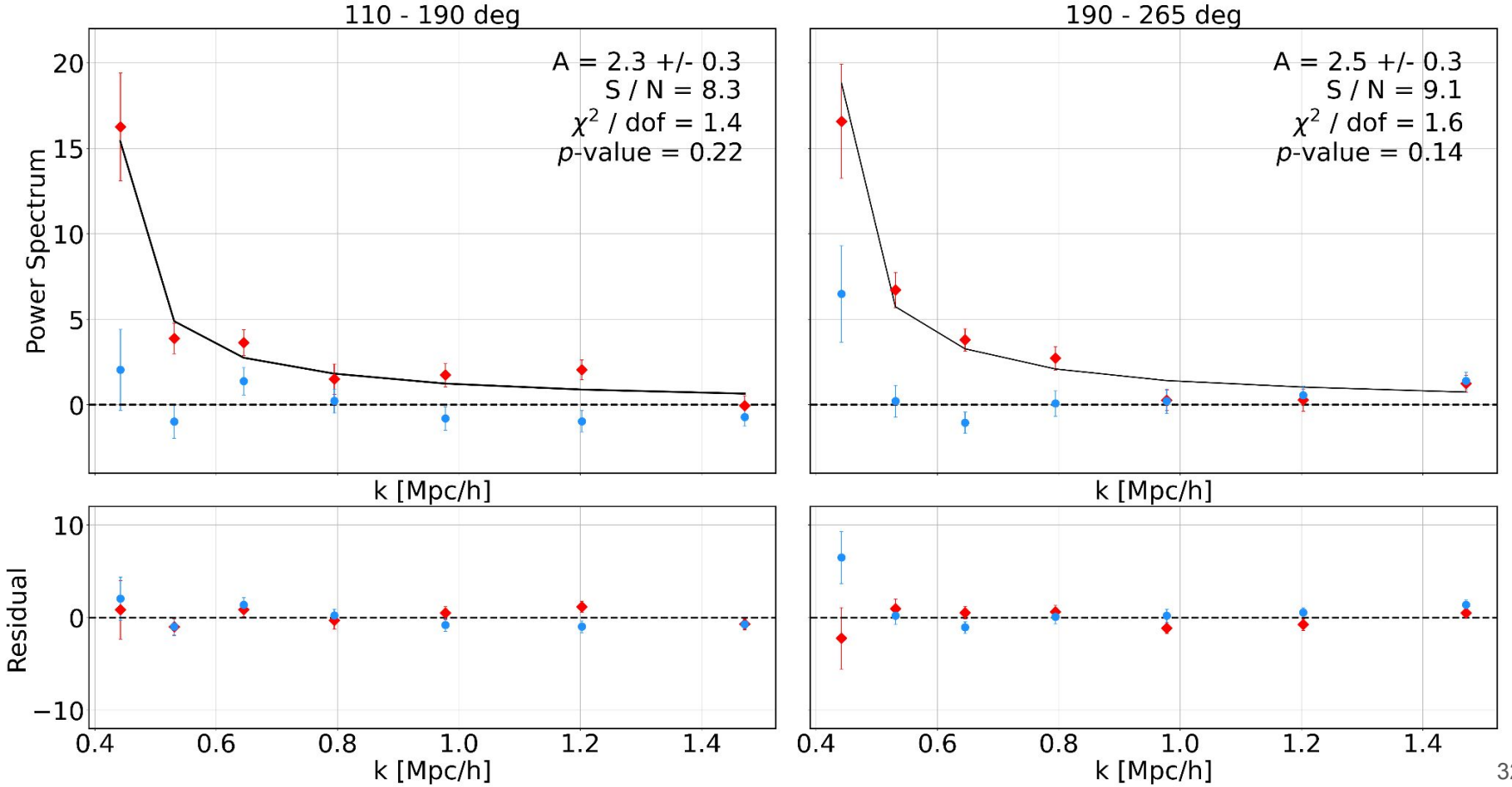
# Polarization consistency : XX and YY pol



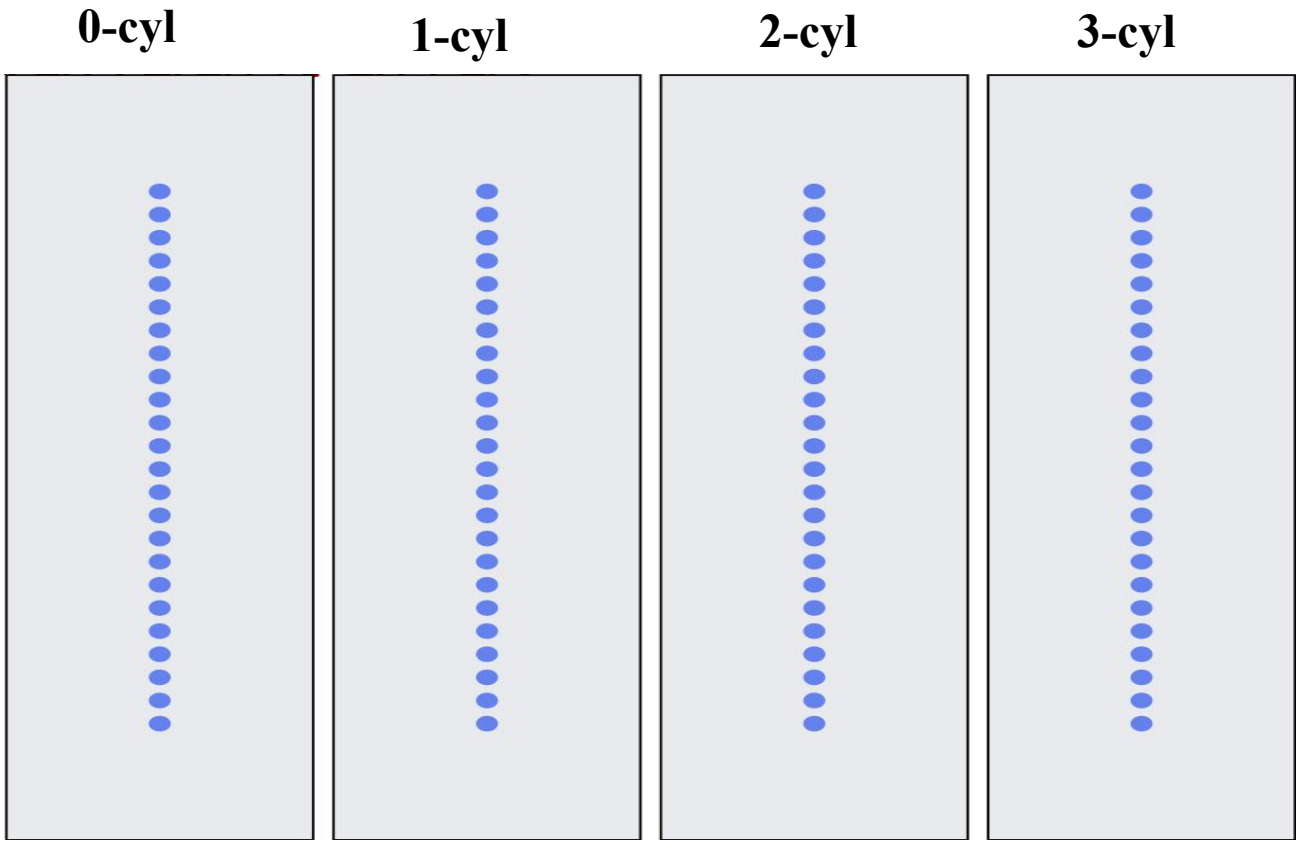
**RA divisions:** RA range of the NGC region is divided into two independent RA bins



# RA divisions: RA range of the NGC region is divided into two independent RA bins

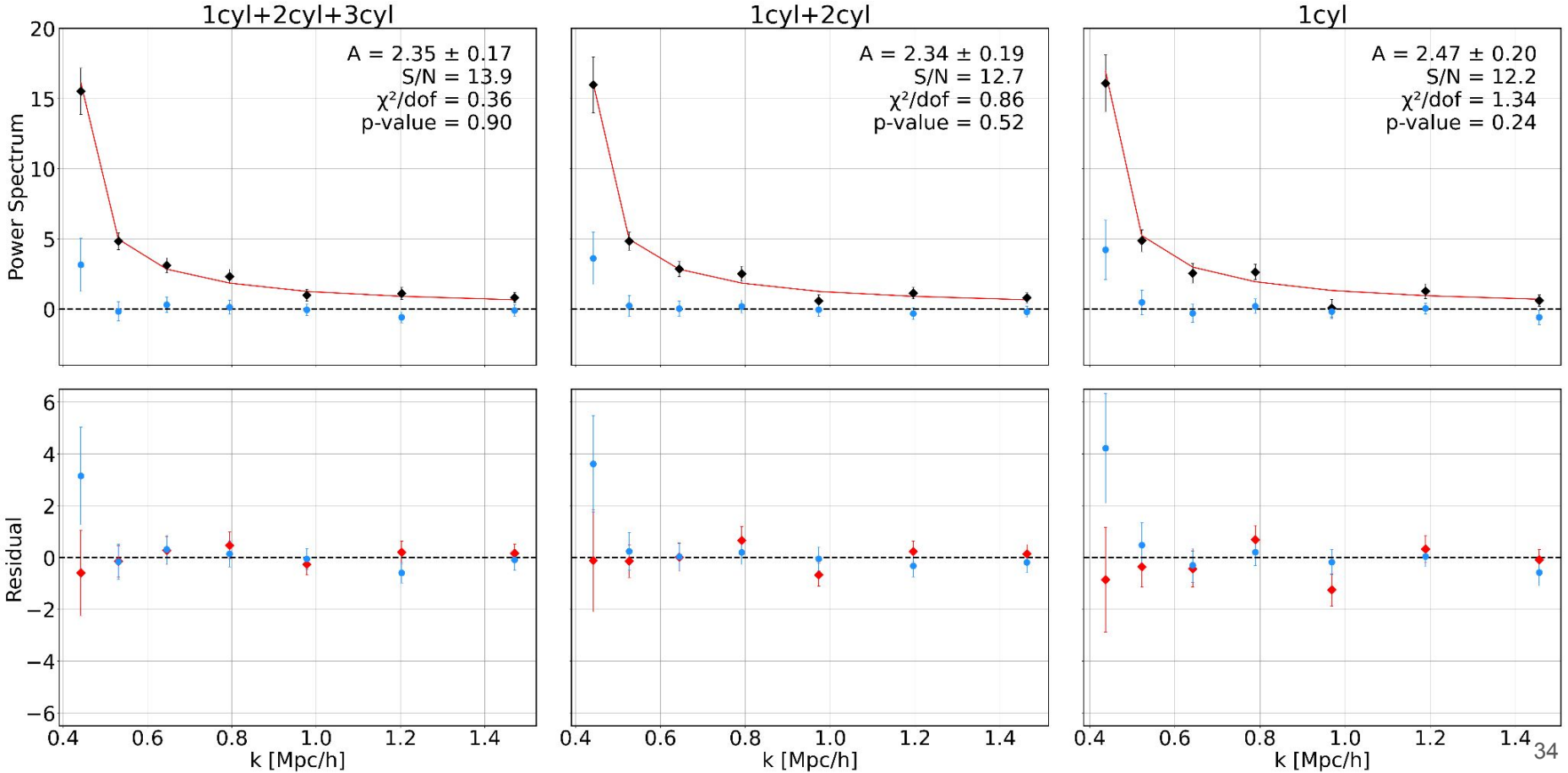


# Baselines cuts: including 3cyl baselines vs 2-cyl baselines vs 1-cyl baselines

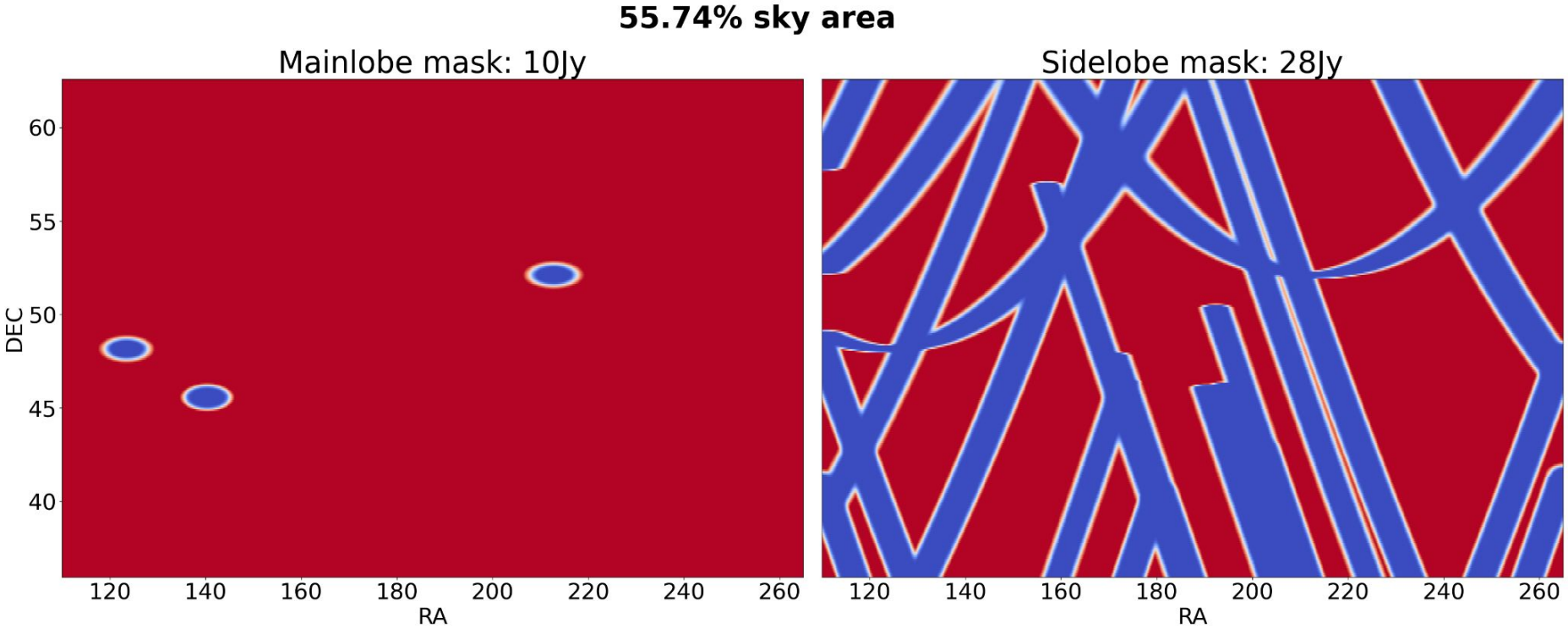



We do not include intra-cyl or 0-cyl separation baselines in this analysis

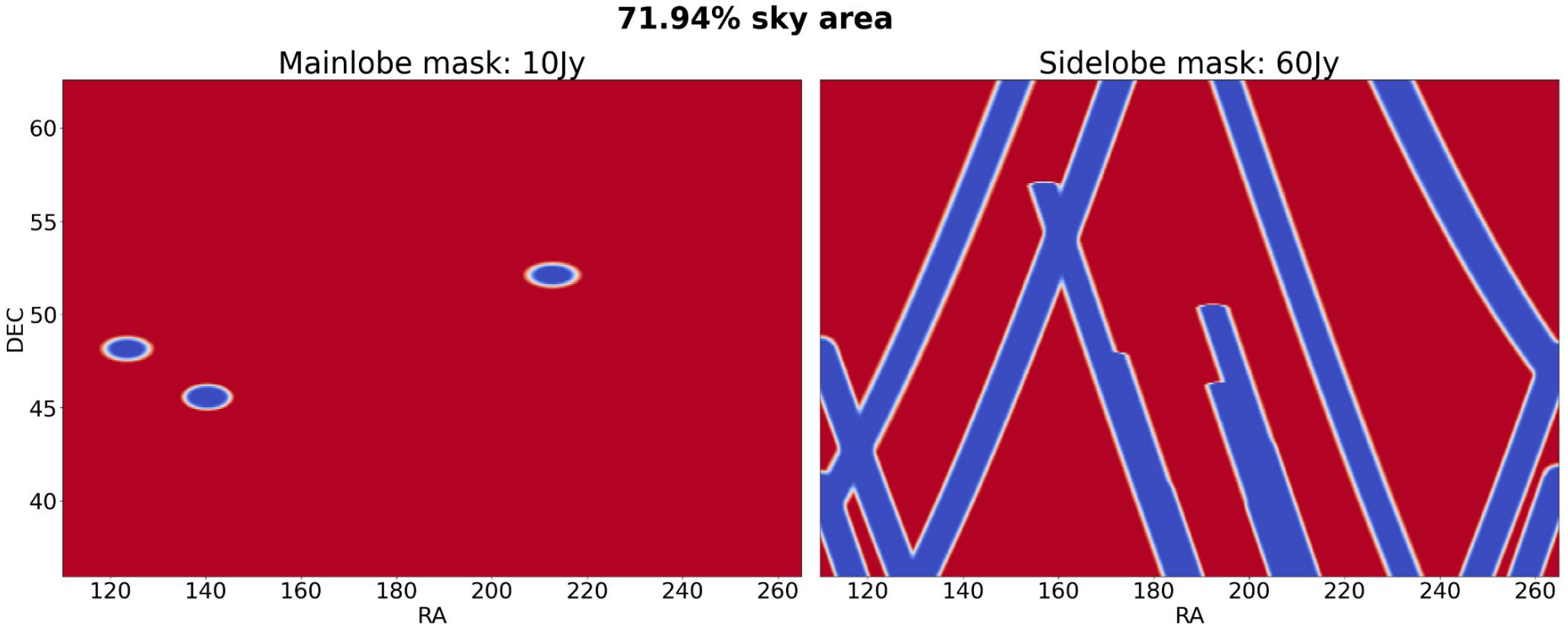
# Baselines cuts: including 3cyl baselines vs 2-cyl baselines vs 1-cyl baselines



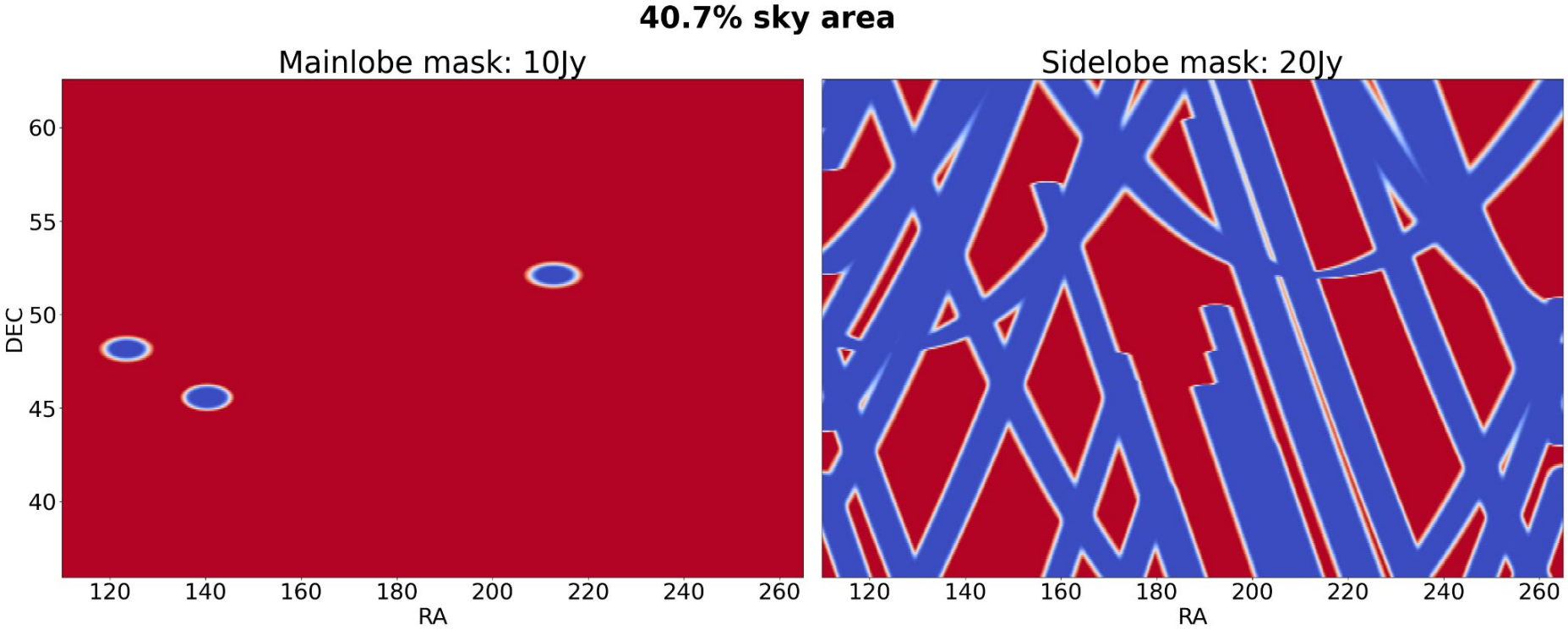
# Statistical tests: Varying Fluxcut of sources to generate a sidelobe mask



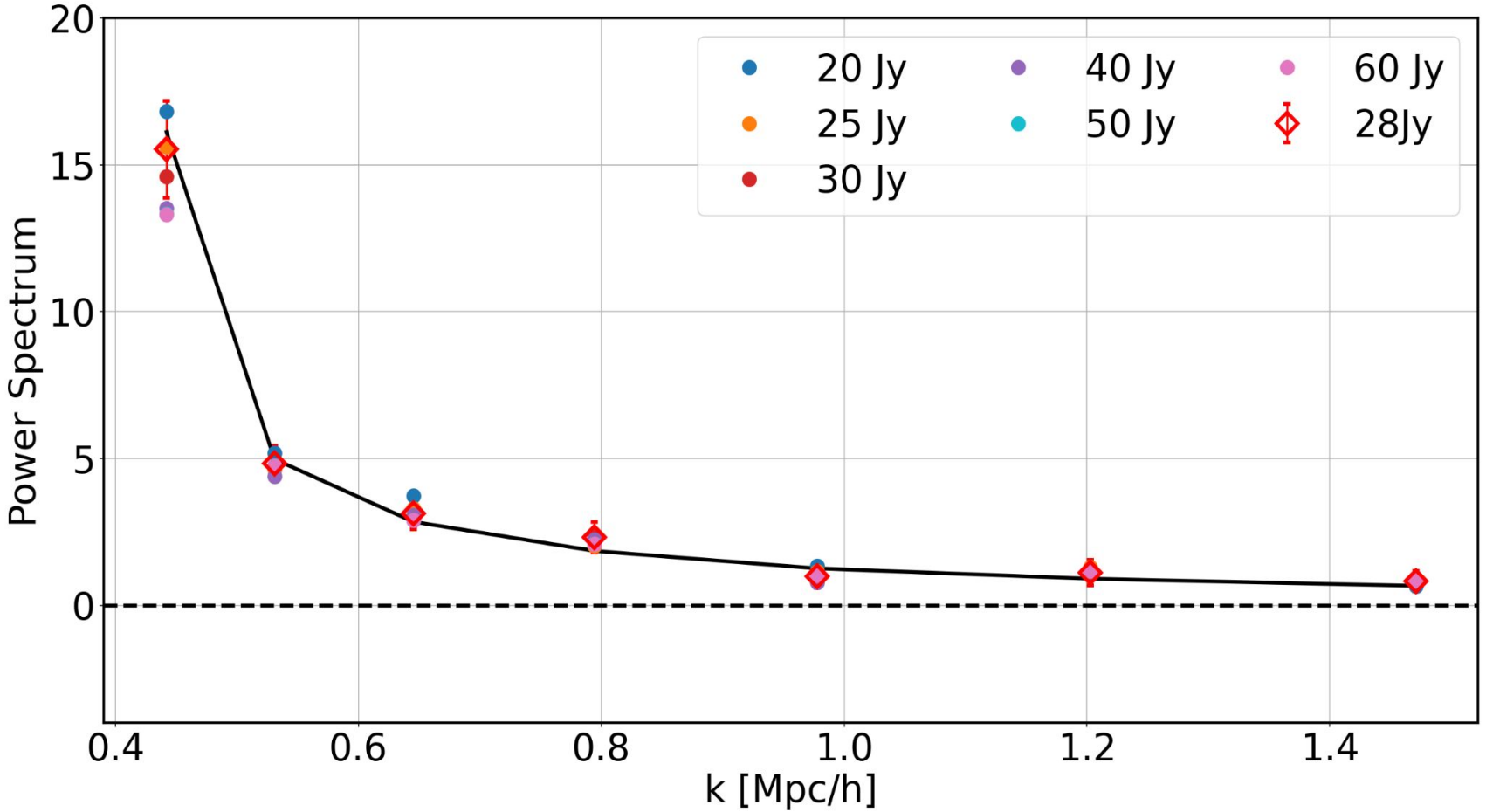
# Statistical tests: Varying Fluxcut of sources to generate a sidelobe mask



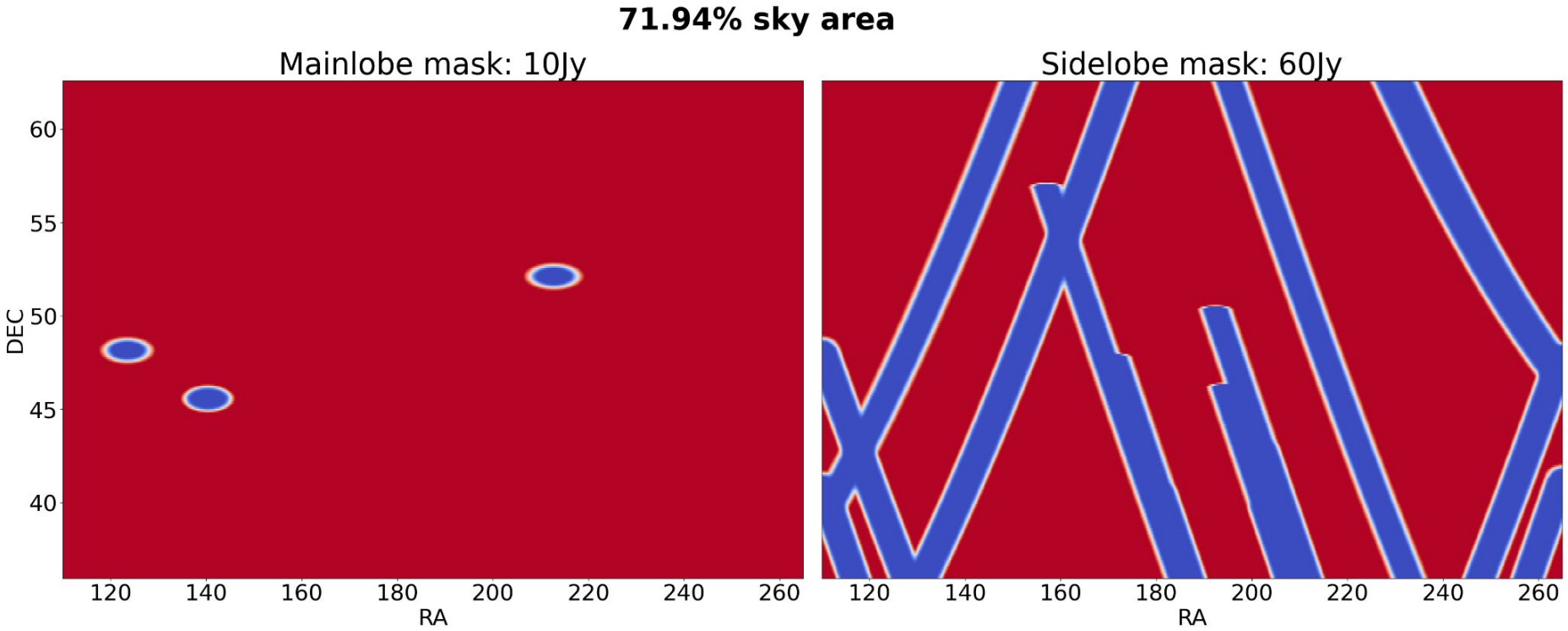
# Statistical tests: Varying Fluxcut of sources to generate a sidelobe mask



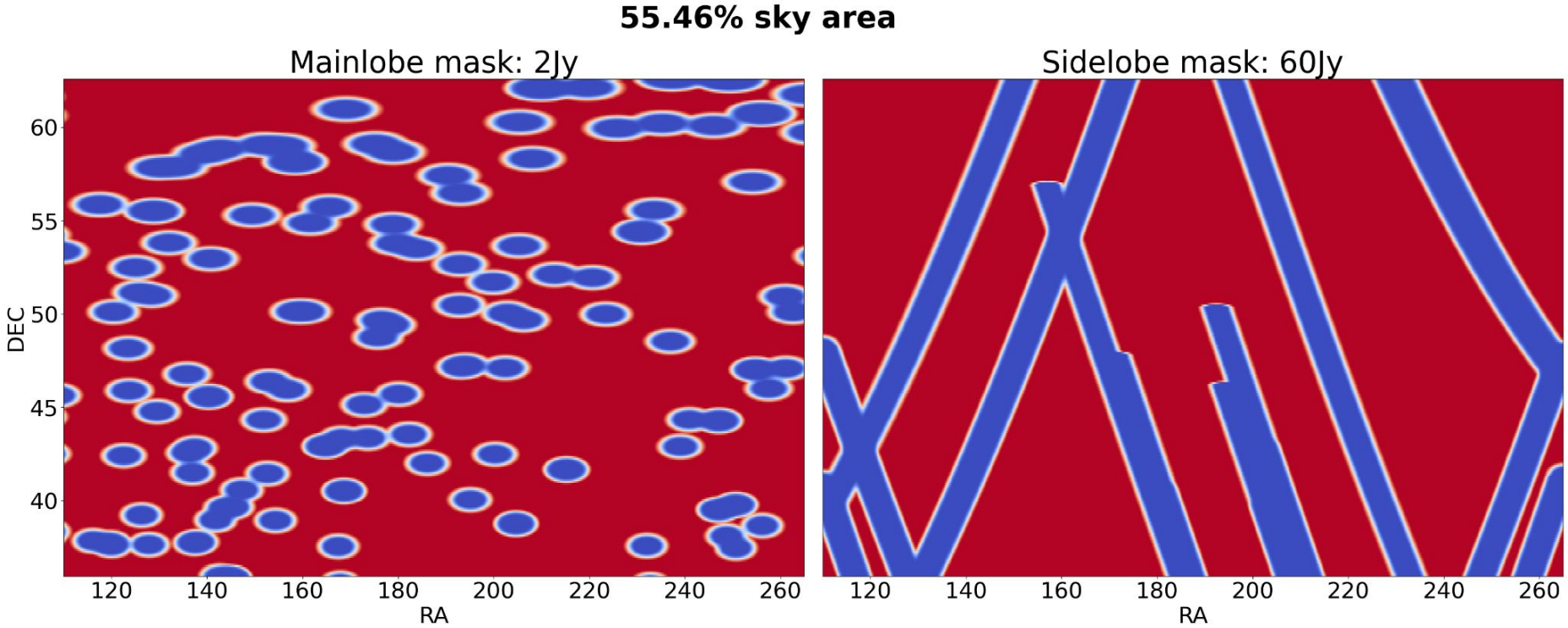
# Statistical tests: Varying Fluxcut of sources to generate a sidelobe mask



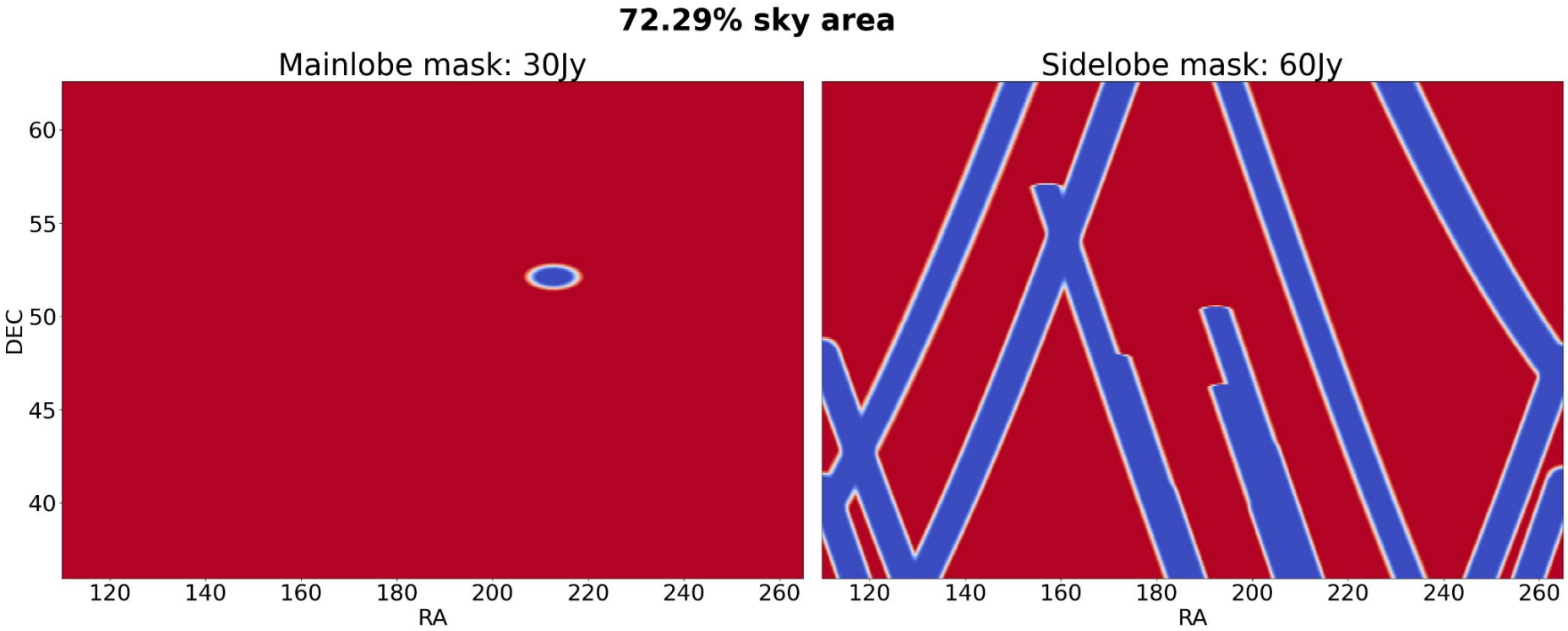
# Statistical tests: Varying Fluxcut of sources to generate a mainlobe mask



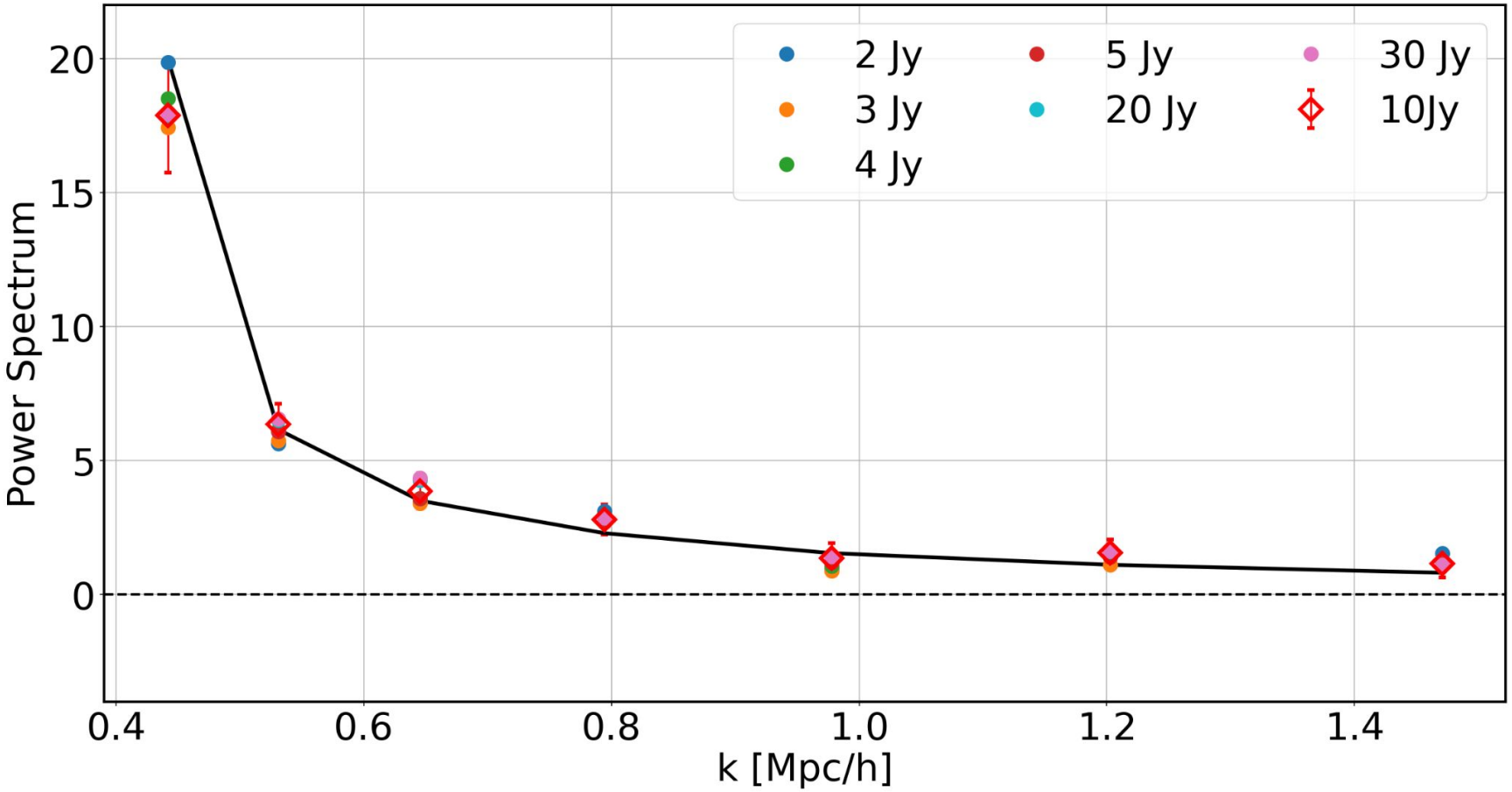
# Statistical tests: Varying Fluxcut of sources to generate a mainlobe mask



# Statistical tests: Varying Fluxcut of sources to generate a mainlobe mask



# Statistical tests: Varying Fluxcut of sources to generate a mainlobe mask



# Summary

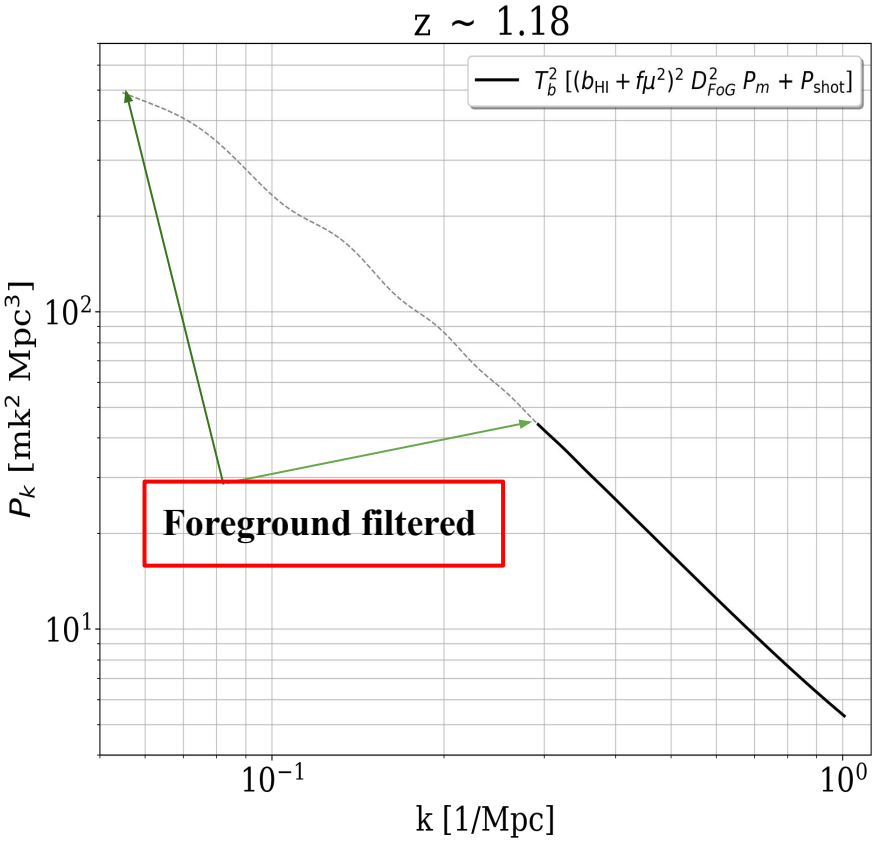
- CHIME is mapping large-scale structure from z = 0.8 to 2.5 using 21cm intensity mapping
- We have reported detection of cosmological 21-cm signal by cross-correlating CHIME map with eBOSS catalogs.
- We have identified and removed bulk of systematics from our data and made a cleanest map of the sky.
- We made progress in measurement of auto-power spectrum with CHIME. First result is coming soon……
- We made many statistical tests to rule out systematics contamination, currently work is going on in template fitting to the data.

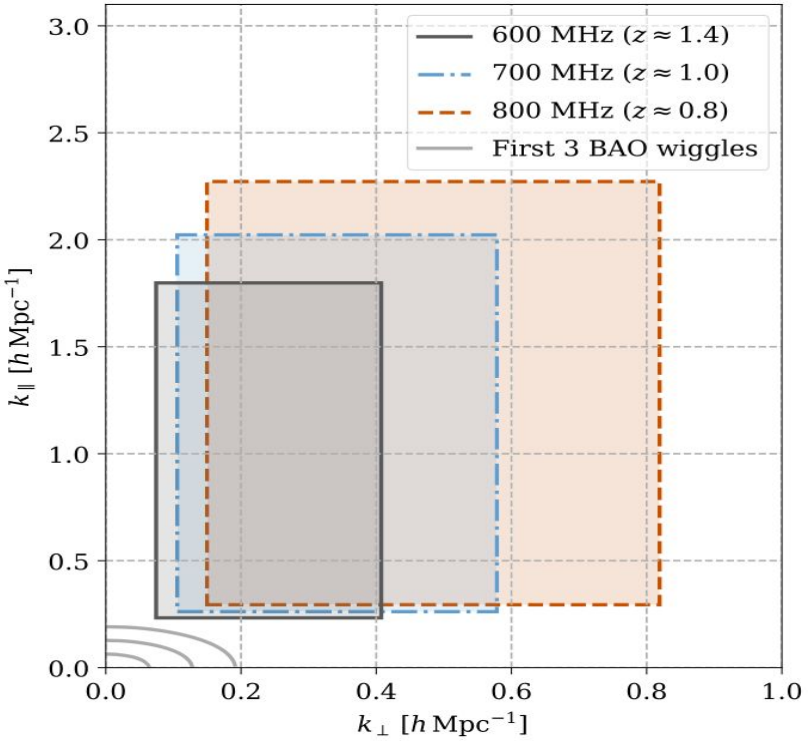
**Thank you!**

# Additional slides

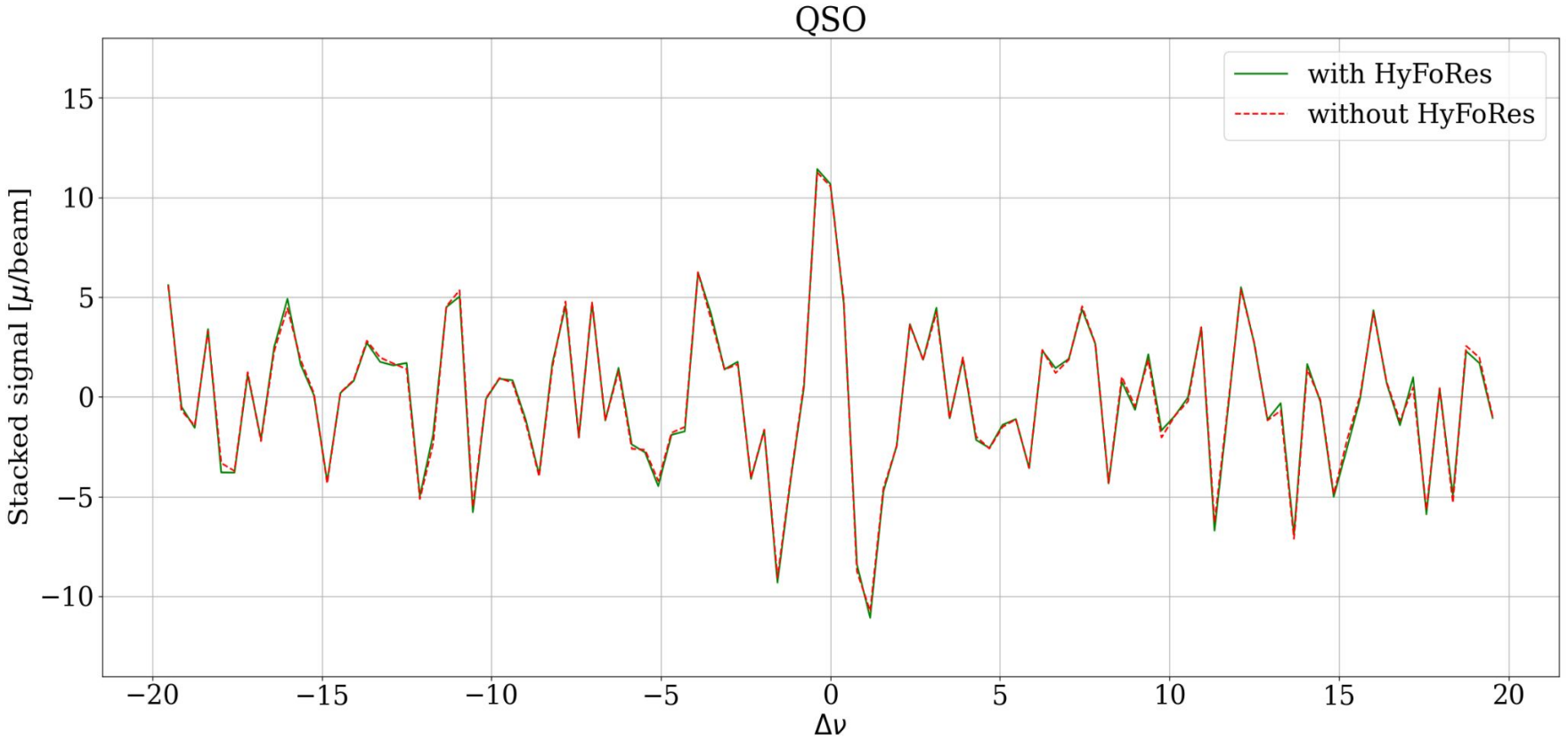
# Future directions towards getting BAO scale back

# Scales Probed in this analysis

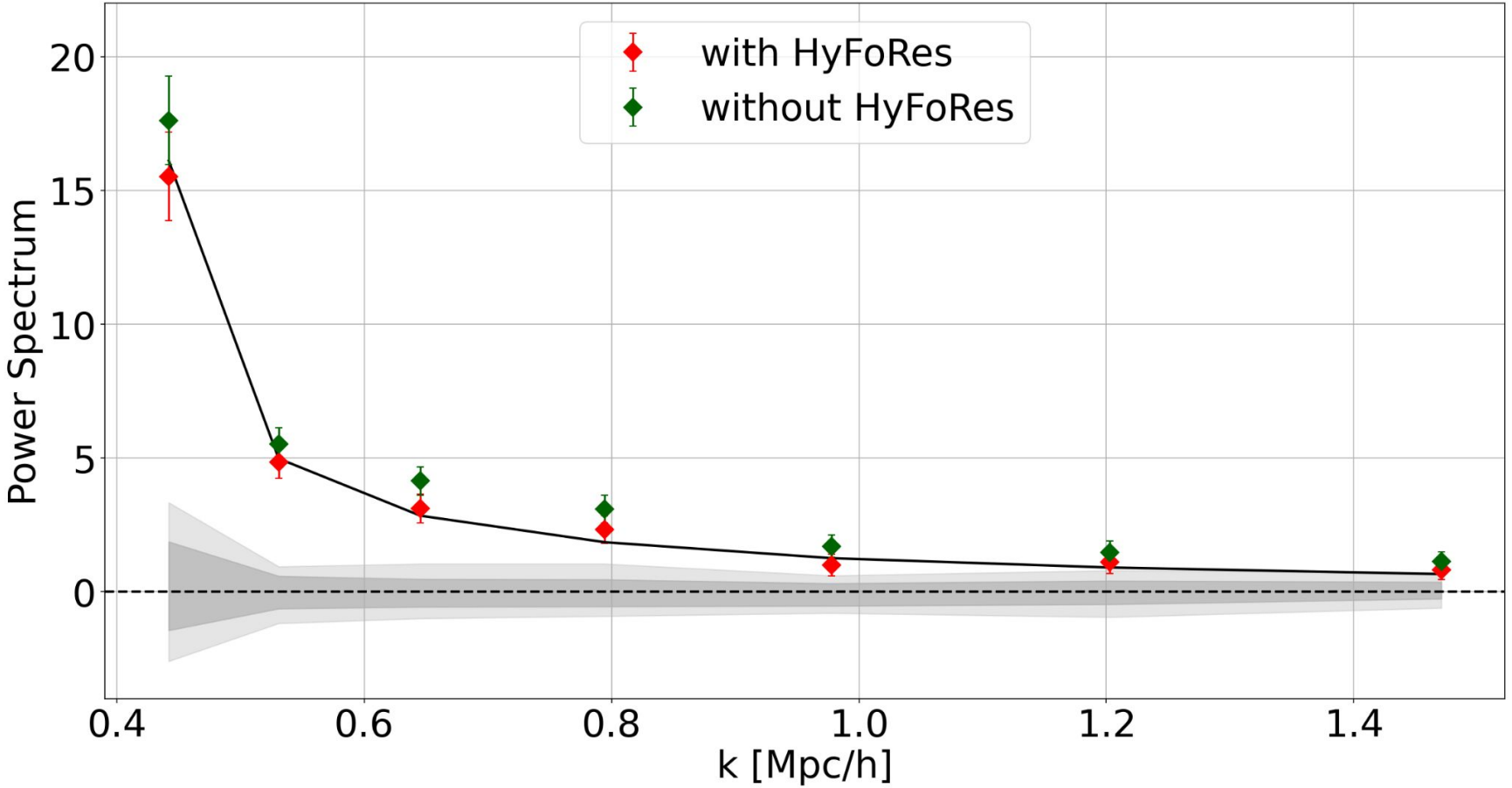


# HyFoReS signal loss estimate: HI signal stacked on eBOSS QSO

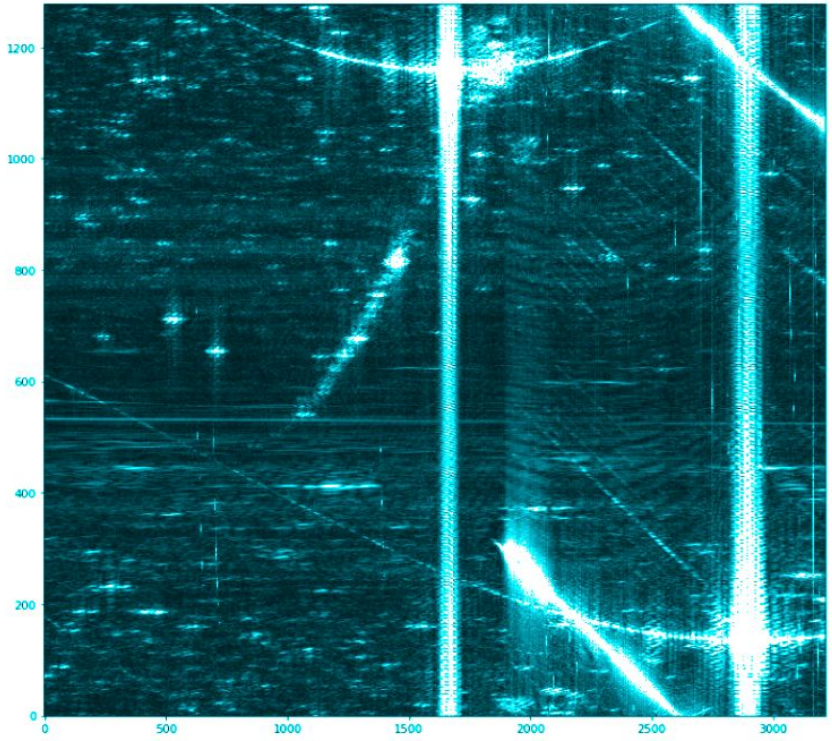


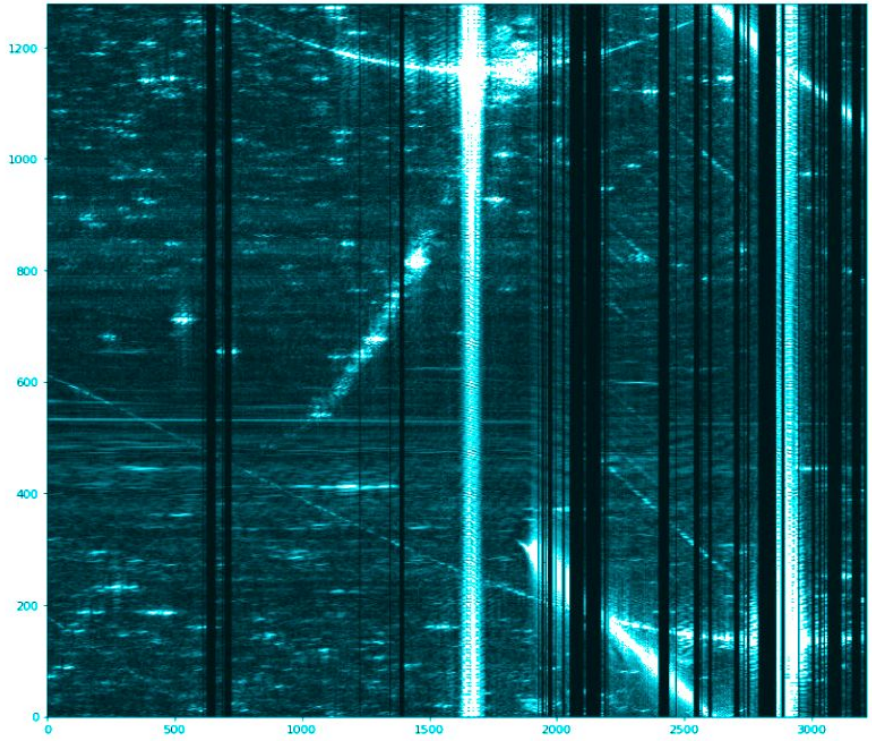
# Signal loss : with and without HyFoRes



# Improved RFI flagging, bad data identification and removal of systematics

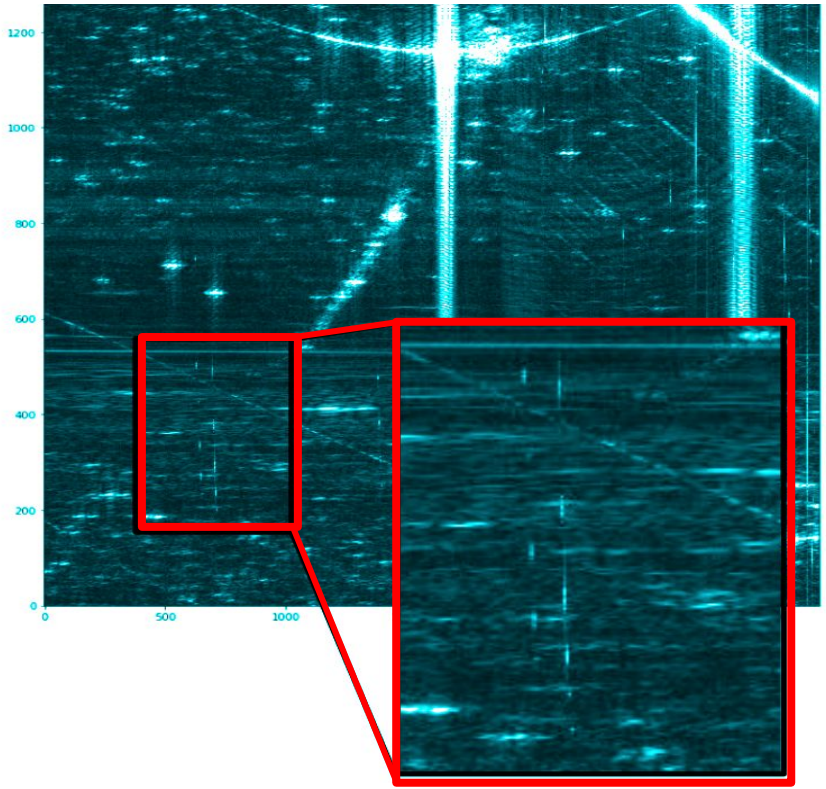
- "Stokes I" Masking:
  - Apply high-pass filter in time to remove background sky, beamform, identify outliers
  - Apply low-pass filter in time, average 2 and 3 cylinder separation visibilities, identify outliers

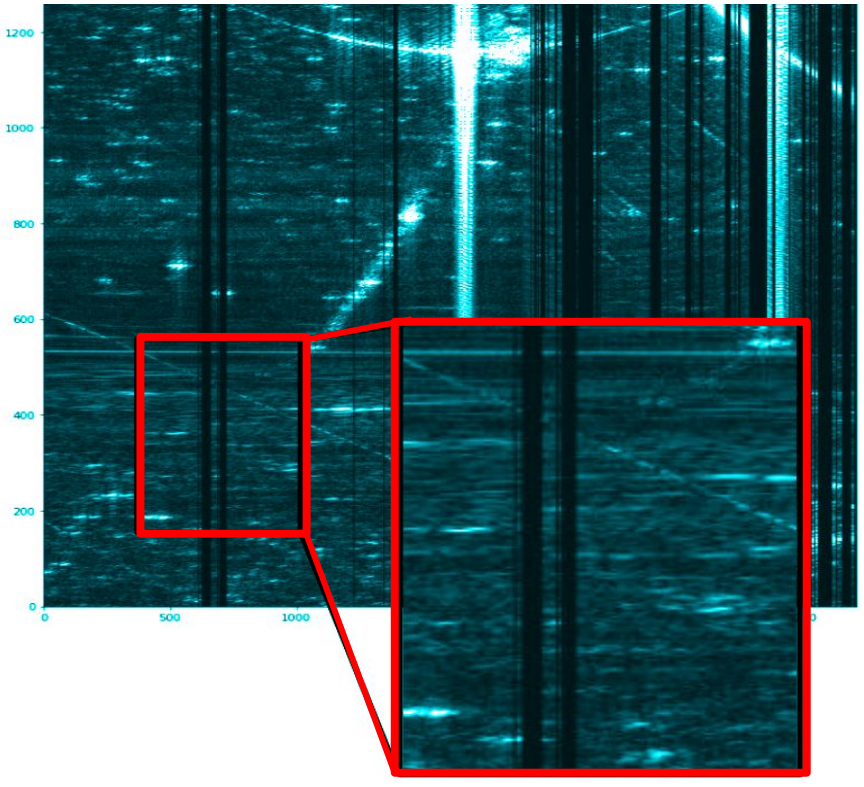


# Improved RFI flagging, bad data identification and removal of systematics

- “Stokes I” Masking:
  - Apply high-pass filter in time to remove background sky, beamform, identify outliers
  - Apply low-pass filter in time, average 2 and 3 cylinder separation visibilities, identify outliers



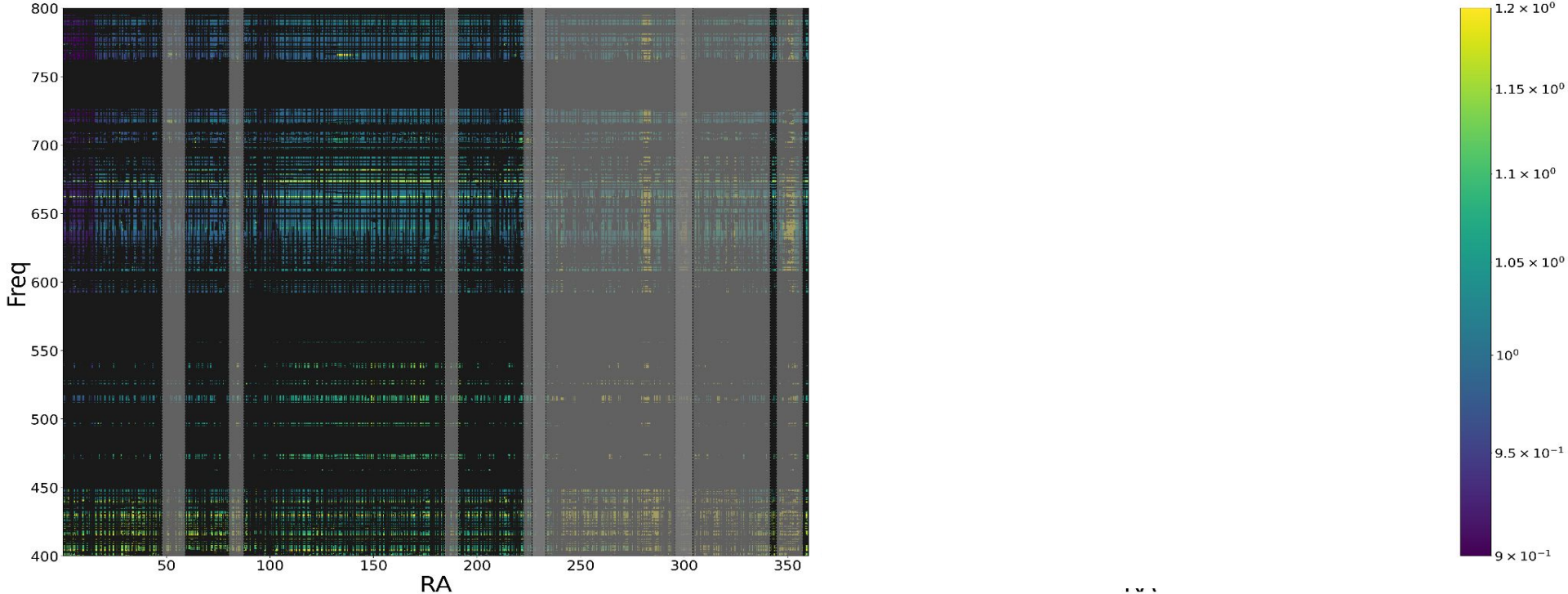
# RFI flagging : $\chi^2$ − test

- Identify outliers in aggressively-delay-filtered visibilities

$$\chi^2_{t\nu} = \sum_{ij} \frac{|\sum_{\nu'} H_{\nu\nu'} V_{ij,t\nu'}|^2}{\underset{t<5\text{min}}{\text{med}}\left[\text{Var}(V_{ij,t\nu})\right]}$$

Dayenu filtered visibilities with $\tau_{cut} = 400$ nsec

Sample variance from even-odd difference at 30 msec, smoothed with 5 min rolling median




**71.23% total flagging for this particular night**

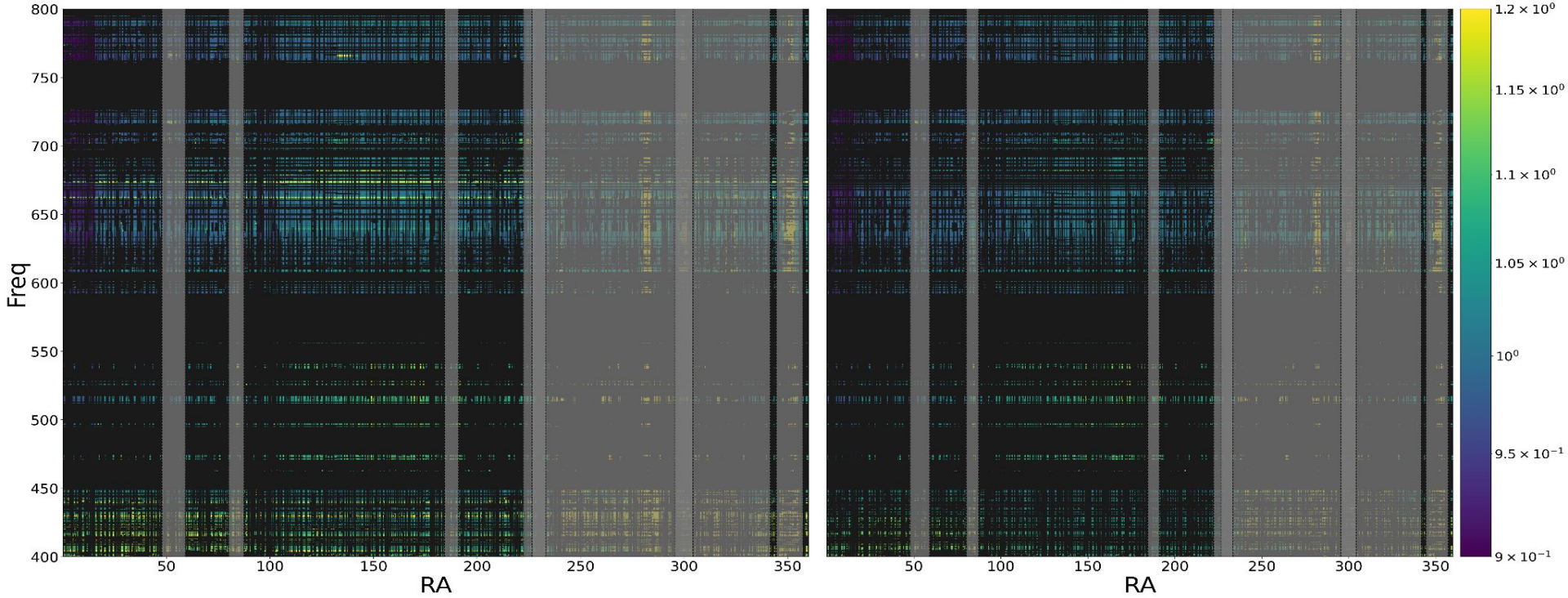
# RFI flagging : $\chi^2$ – test

- Identify outliers in aggressively-delay-filtered visibilities

$$\chi^2_{t\nu} = \sum_{ij} \frac{|\sum_{\nu'} H_{\nu\nu'} V_{ij,t\nu'}|^2}{\underset{t<5\text{min}}{\text{med}}\left[\text{Var}(V_{ij,t\nu})\right]}$$

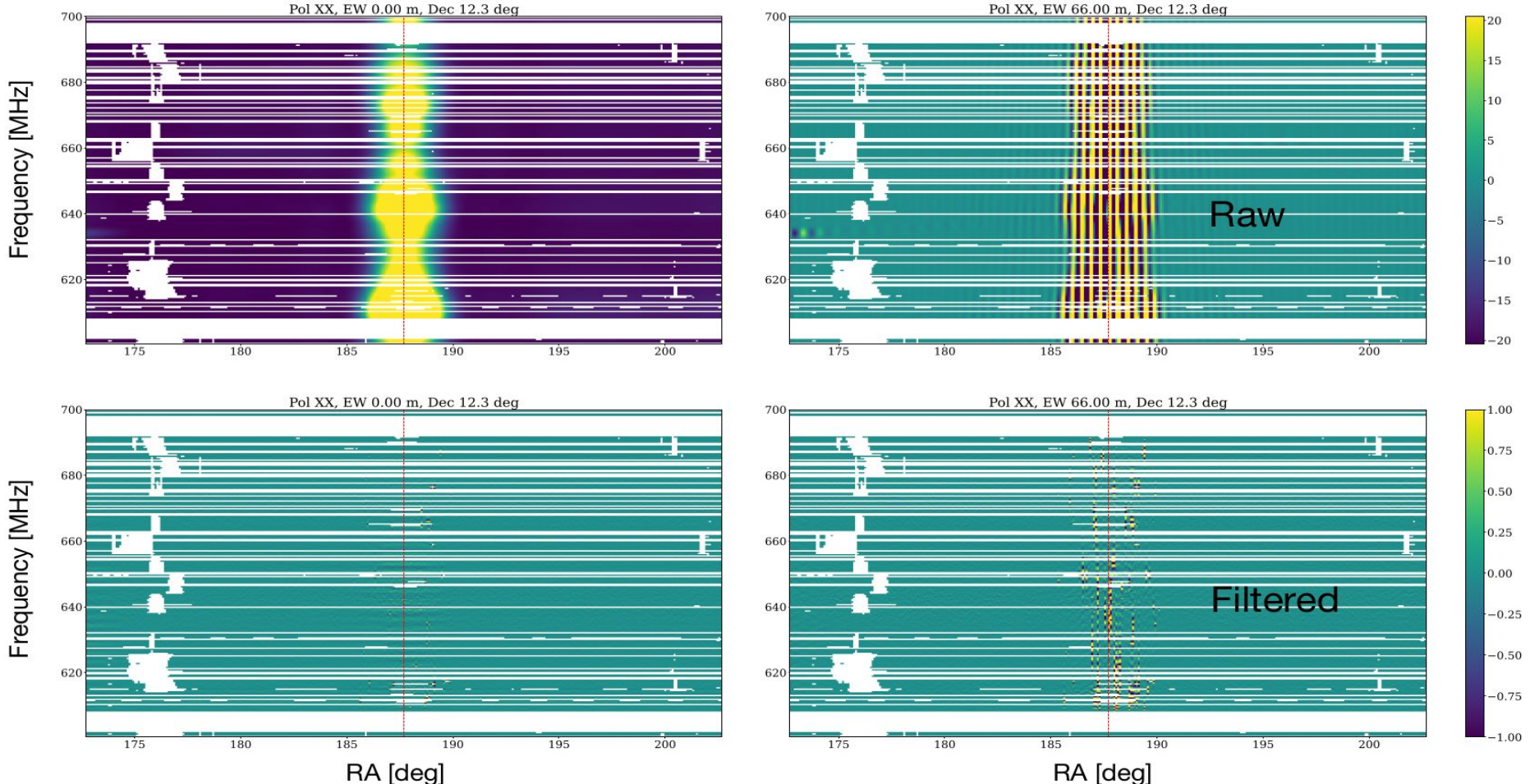
Dayenu filtered visibilities with $\tau_{cut} = 400$ nsec

Sample variance from even-odd difference at 30 msec, smoothed with 5 min rolling median

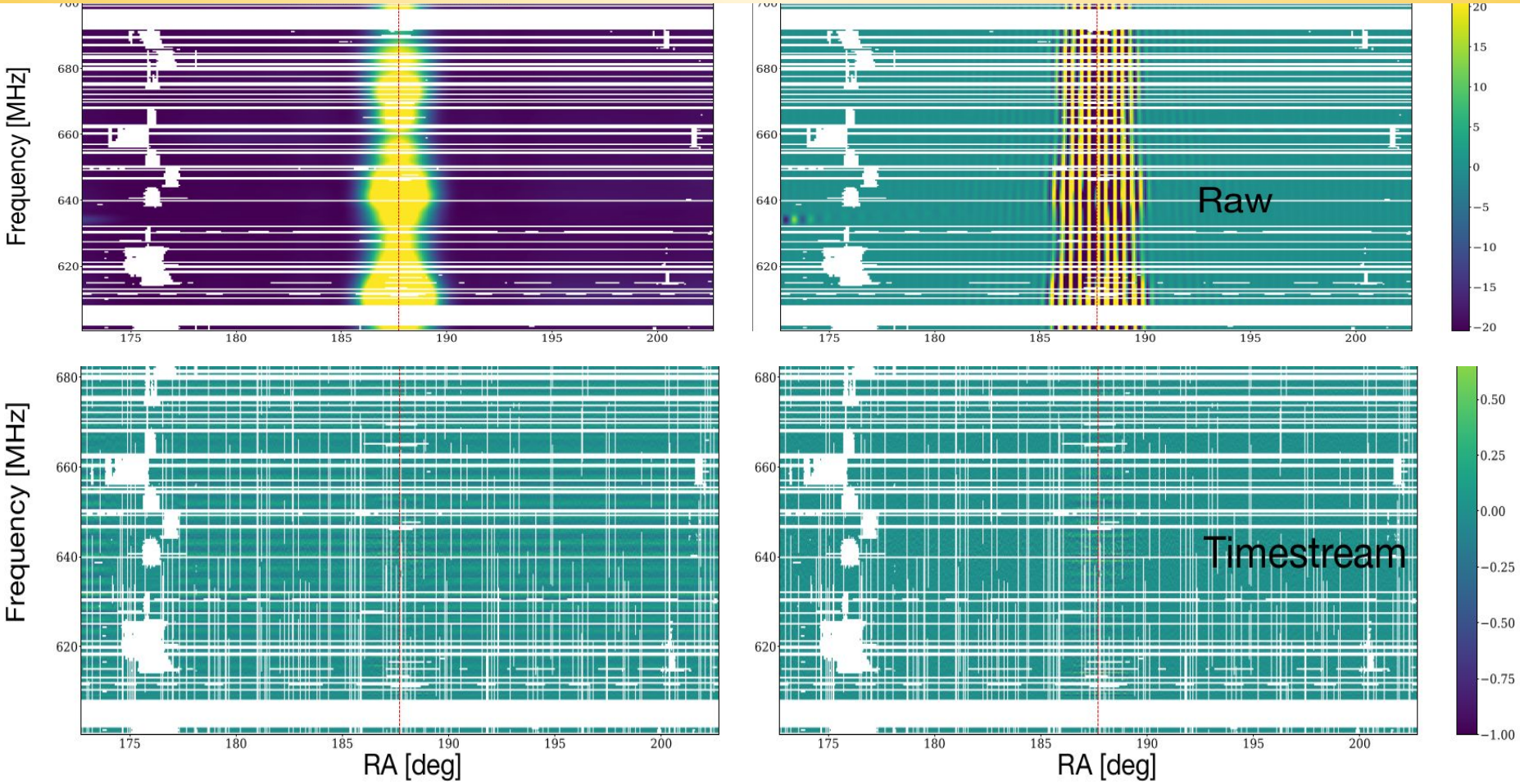



**71.23% total flagging for this particular night**

# Systematics : BeamFormed to VIR_A



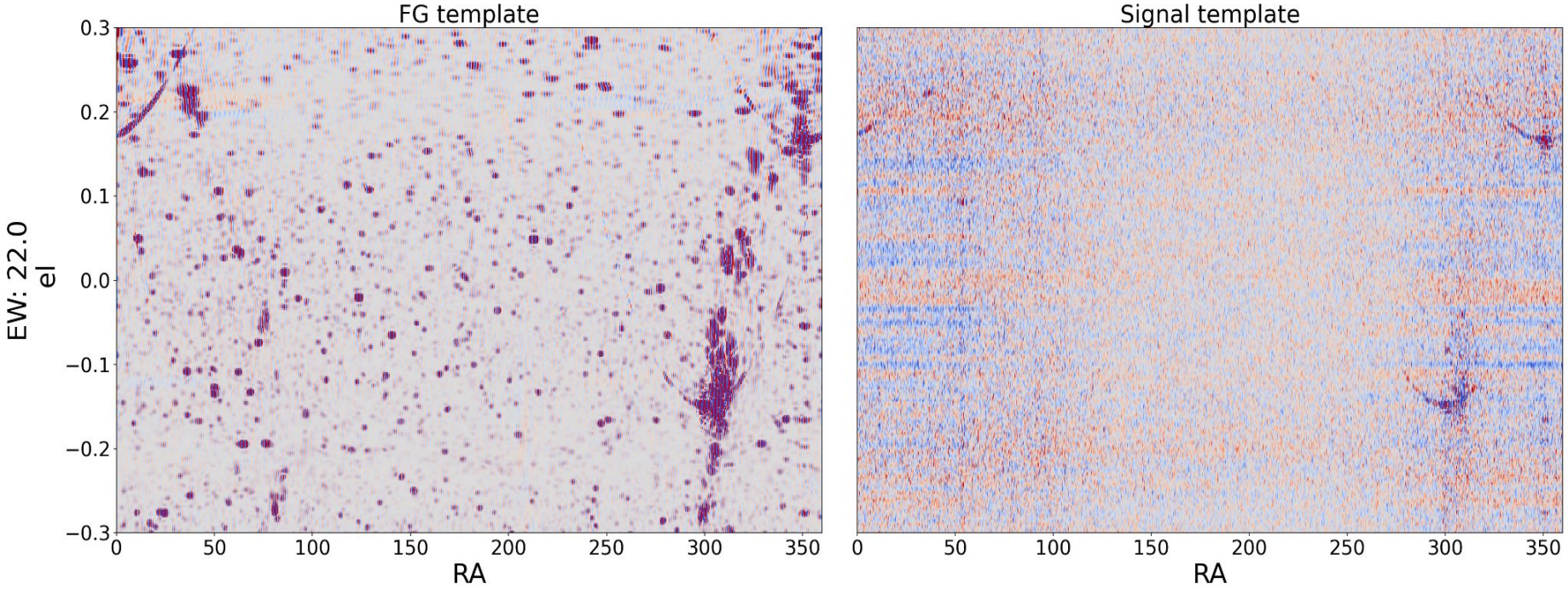
# Systematics : BeamFormed to VIR_A



# HyFoRes : To remove residual bandpass error

$$\hat{y}_{p,e,\nu}=\frac{\sum_i \hat{f}^{*}_{p,e,\nu,i}\hat{s}_{p,e,\nu,i}}{\sum_i \hat{f}^{*}_{p,e,\nu,i}\hat{f}_{p,e,\nu,i}}$$

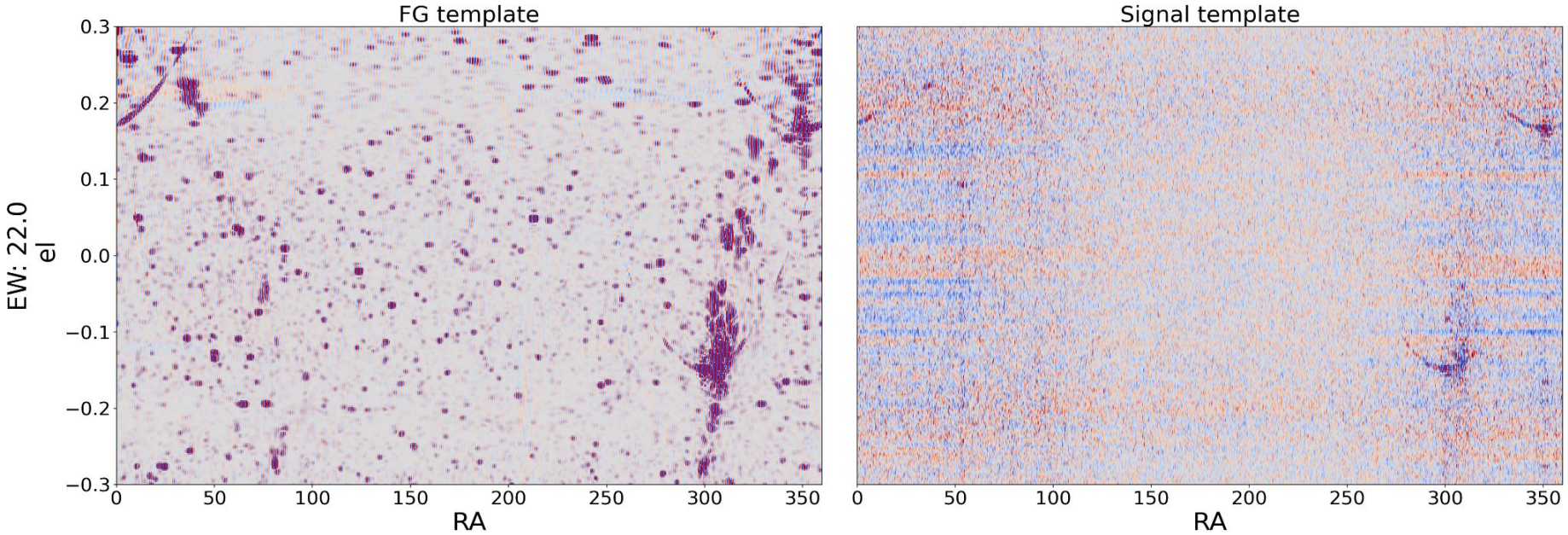
- **Assuming stationarity over RA and DEC**
- **Estimate gain per polarization and per basleines**



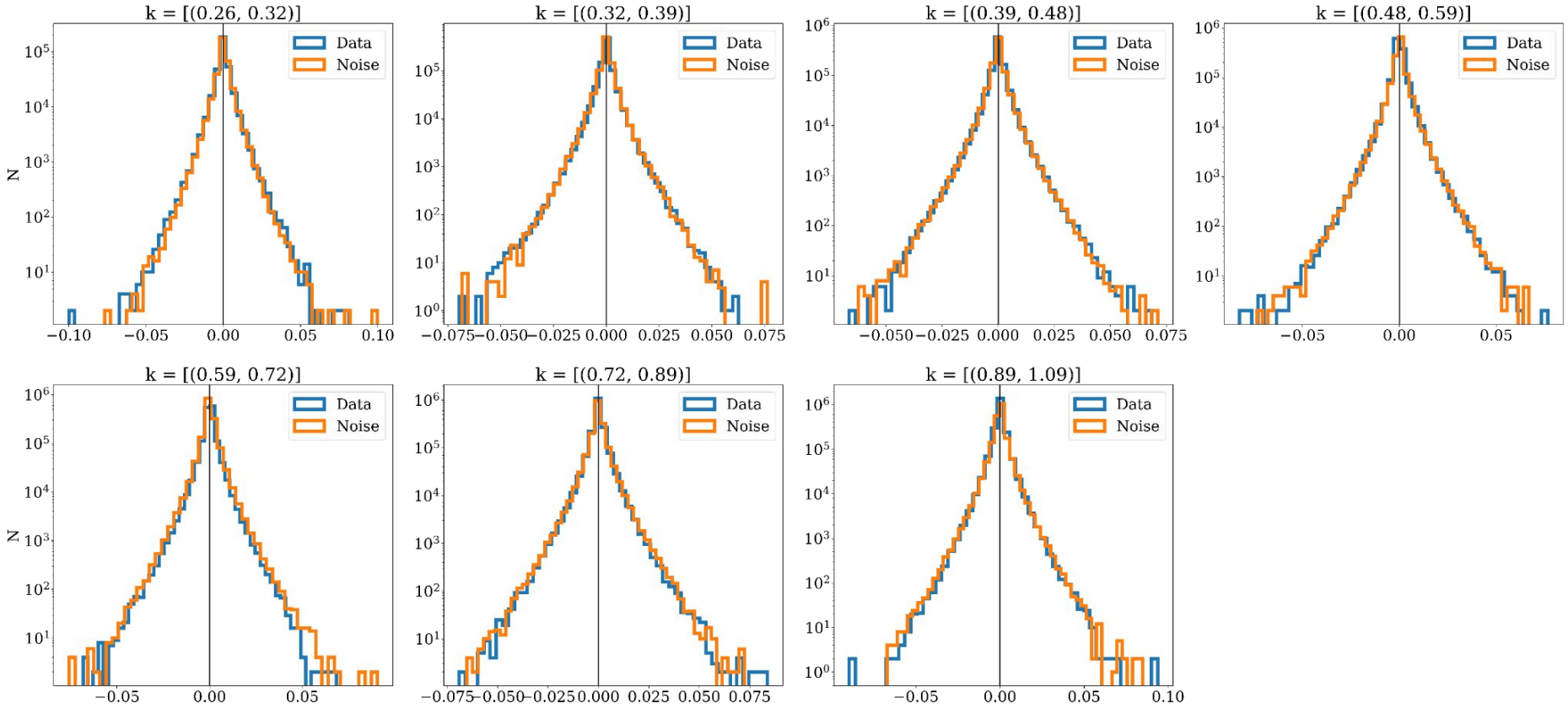
# HyFoRes : To remove residual bandpass error: Implementation

$$\hat{y}_{p,e,\nu} = \frac{\sum_i \hat{f}^{\cdot}_{p,e,\nu,i} \hat{s}_{p,e,\nu,i}}{\sum_i \hat{f}^{\cdot}_{p,e,\nu,i} \hat{f}_{p,e,\nu,i}}$$

- **We use a low-pass filter to the FG template before cross-correlation to remove noise bias at high delay**
- **We also subtract noise cross-talk of both FG and signal template before cross-correlation**

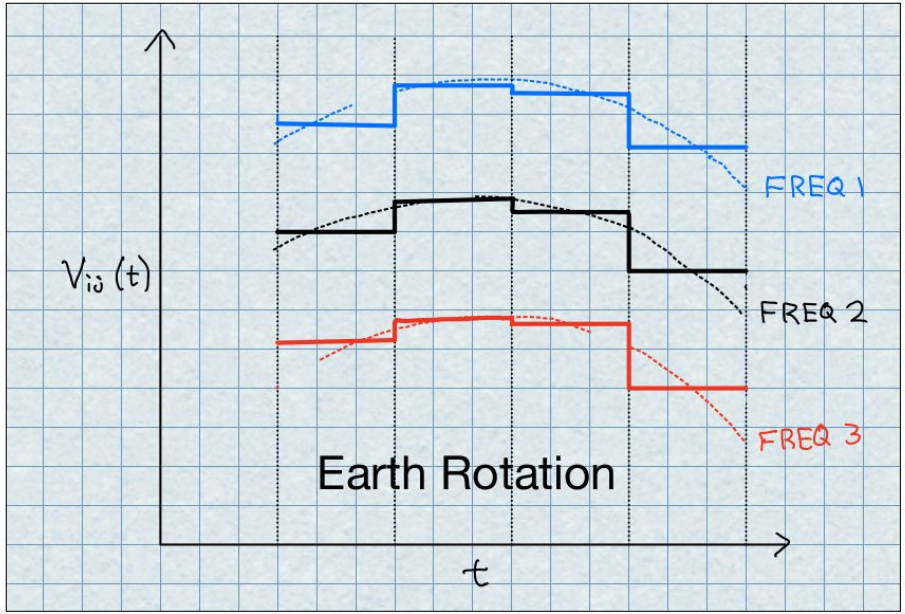


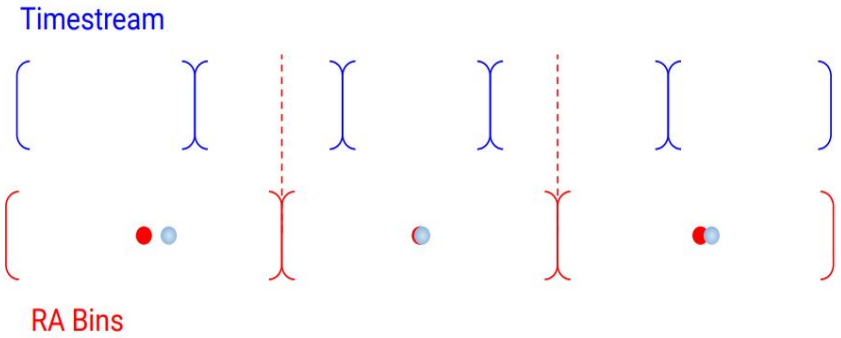
# Histogram of bin values in each 1D k-modes



# Systematics : Subtle issues with data averaging

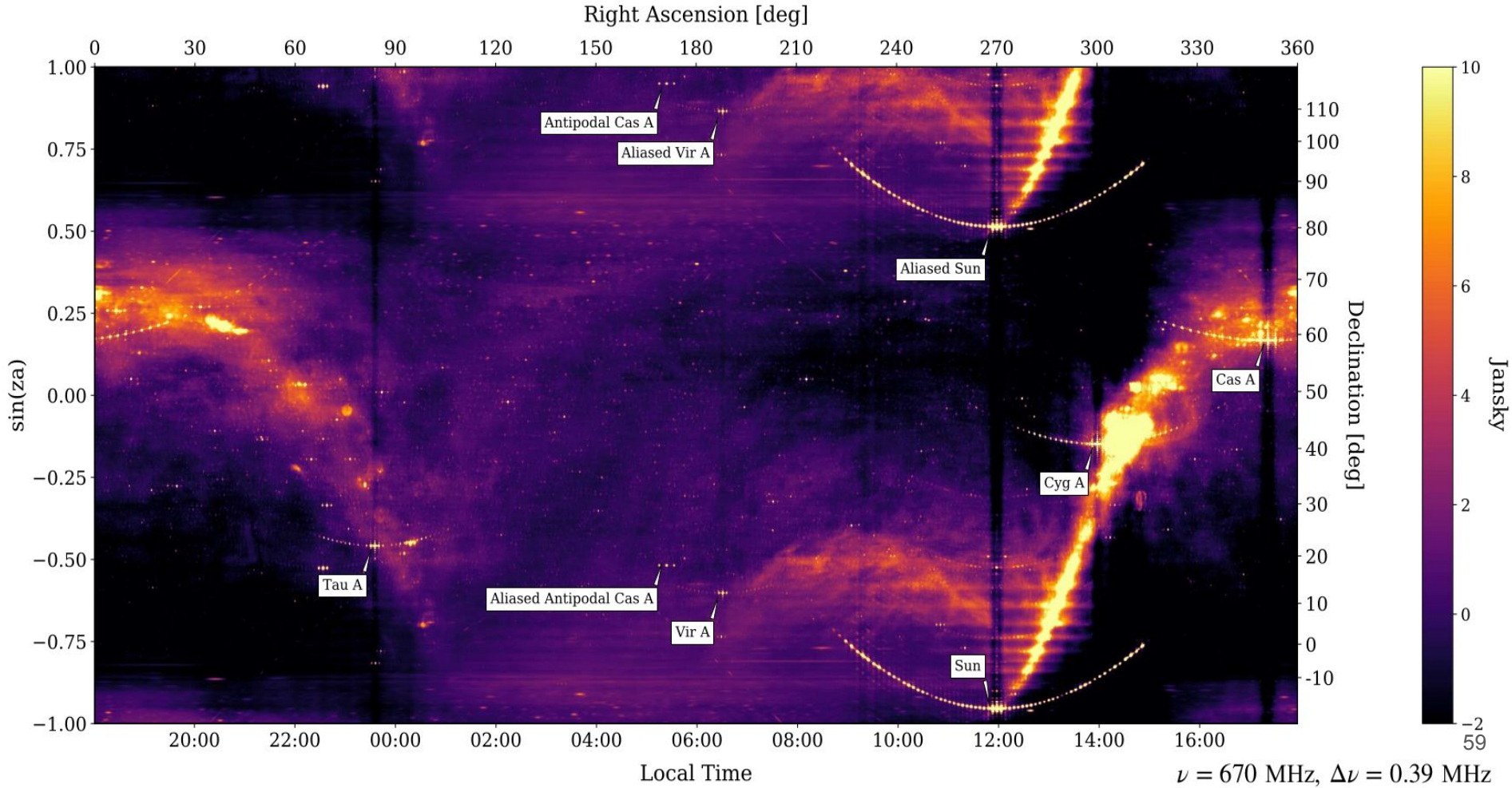
Integrating over visibilities that are changing as a function of time using frequency-dependent weighting leaks foreground power to high delays.



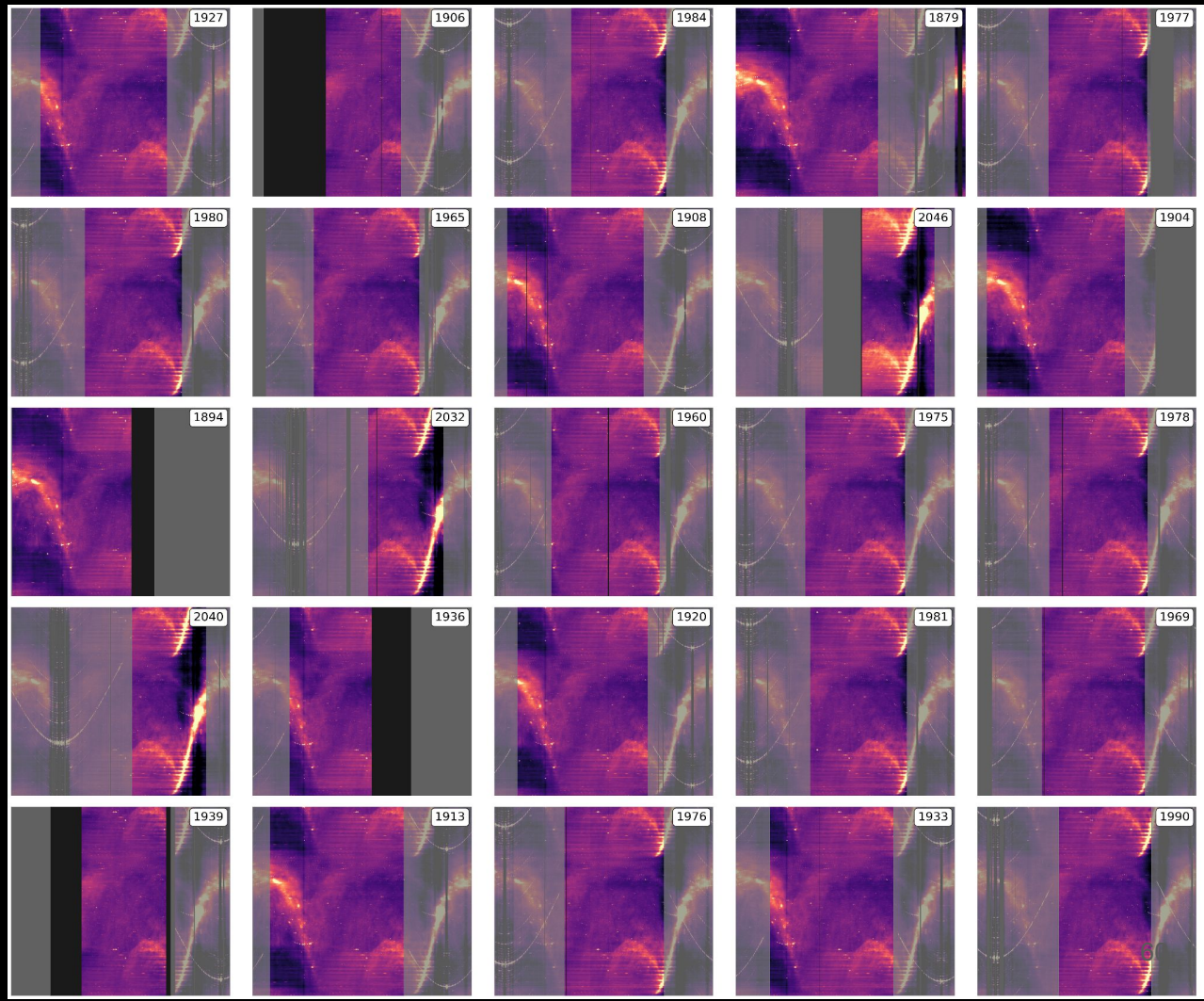

Solution : We take the **timestream data**, then beamformed along declination axis for each east-west baseline separation and use a constant FG filter cutoff for all declination and all baselines

# Radio Sky as seen by CHIME every night



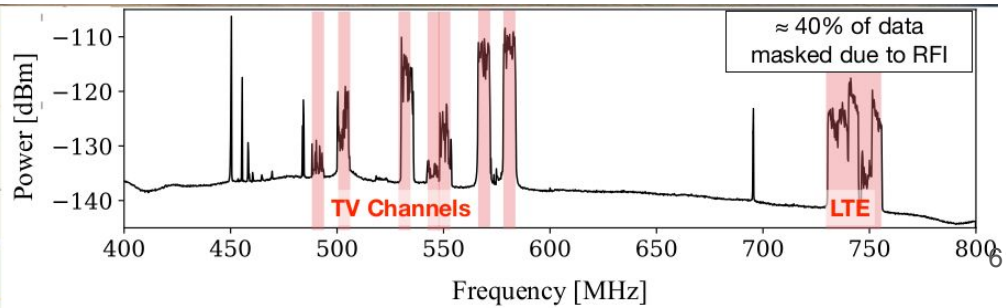
# CHIME data processing

- Data from 2019
- Roughly 100 nights after throwing away potential bad nights.
- Regrid the data and make a stack of 100 nights.

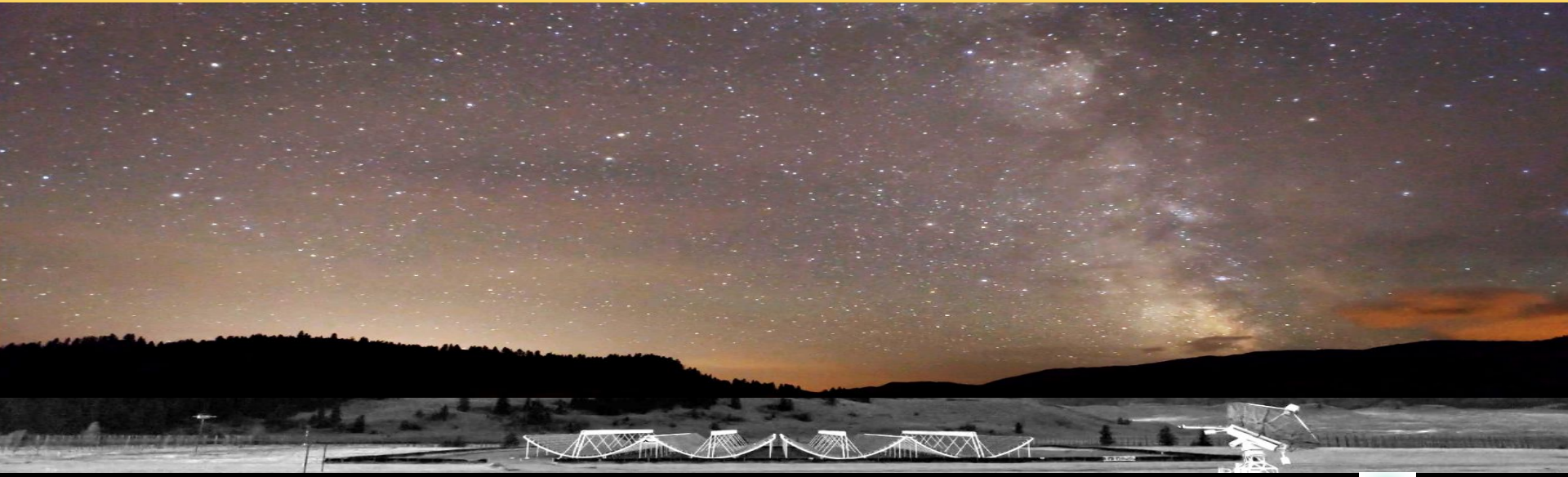


# Dominion Radio Astrophysical Observatory (DRAO)





# Canadian Hydrogen Intensity Mapping Experiment (CHIME)



- Transit radio interferometer (no moving parts).
- 4 cylindrical reflectors, 20m x 100m each.
- Observe between 400-800 MHz, corresponding to redshift of 2.5-0.8 at the 21 cm line.
- 1024 dual polarized antennas.
- Maps the northern radio sky every day.

Beam:
~120 x 2 deg