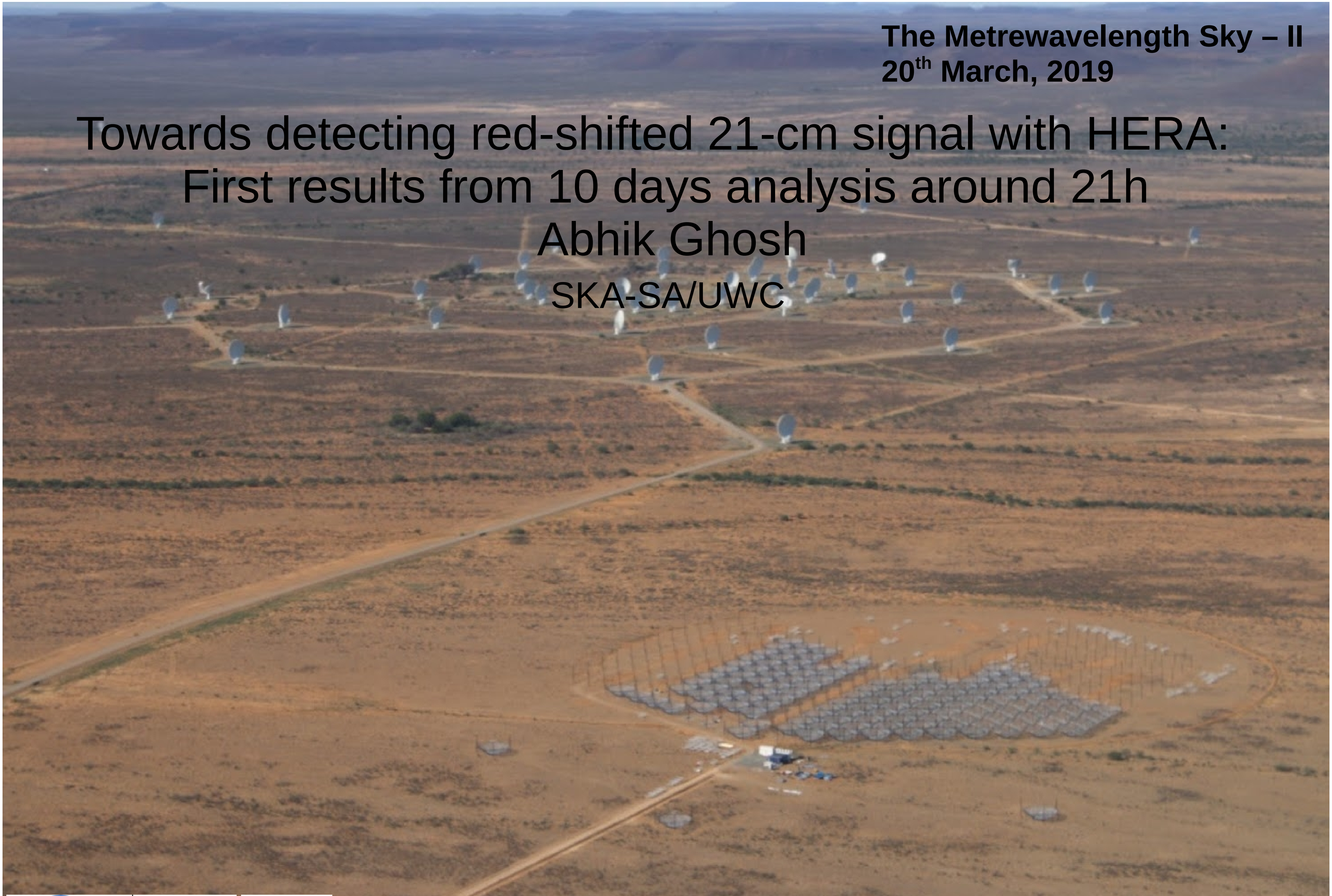


The Metrewavelength Sky – II
20th March, 2019

Towards detecting red-shifted 21-cm signal with HERA: First results from 10 days analysis around 21h

Abhik Ghosh

SKA-SA/UWC

NRF National Research Foundation | SARAO South African Radio Astronomy Observatory | UNIVERSITY of the WESTERN CAPE

In collaboration with Gianni Bernardi(INAF, Italy), Nicholas S. Kern(Berkeley), [Leon Koopmans, Florent Mertens](Kapteyn Astronomical Institute, Netherlands), [Oleg Smirnov, Mario Santos, Trienko Grobler](SKA-SA) + HERA collaboration members

Two scenarios of CMB photons interacting with a cloud of hydrogen gas placed at different distances along the line of sight :

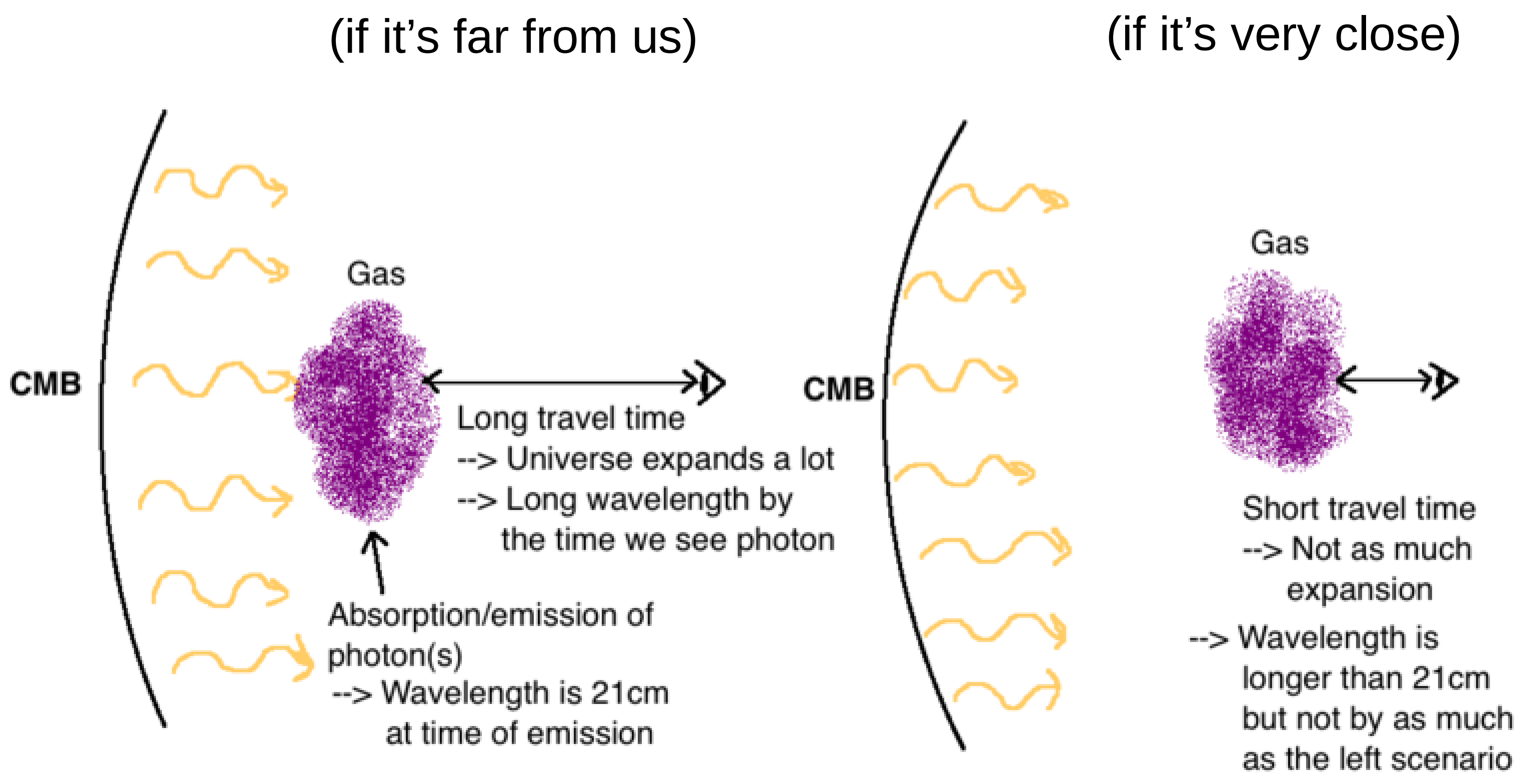


Figure credit: Adrian Liu

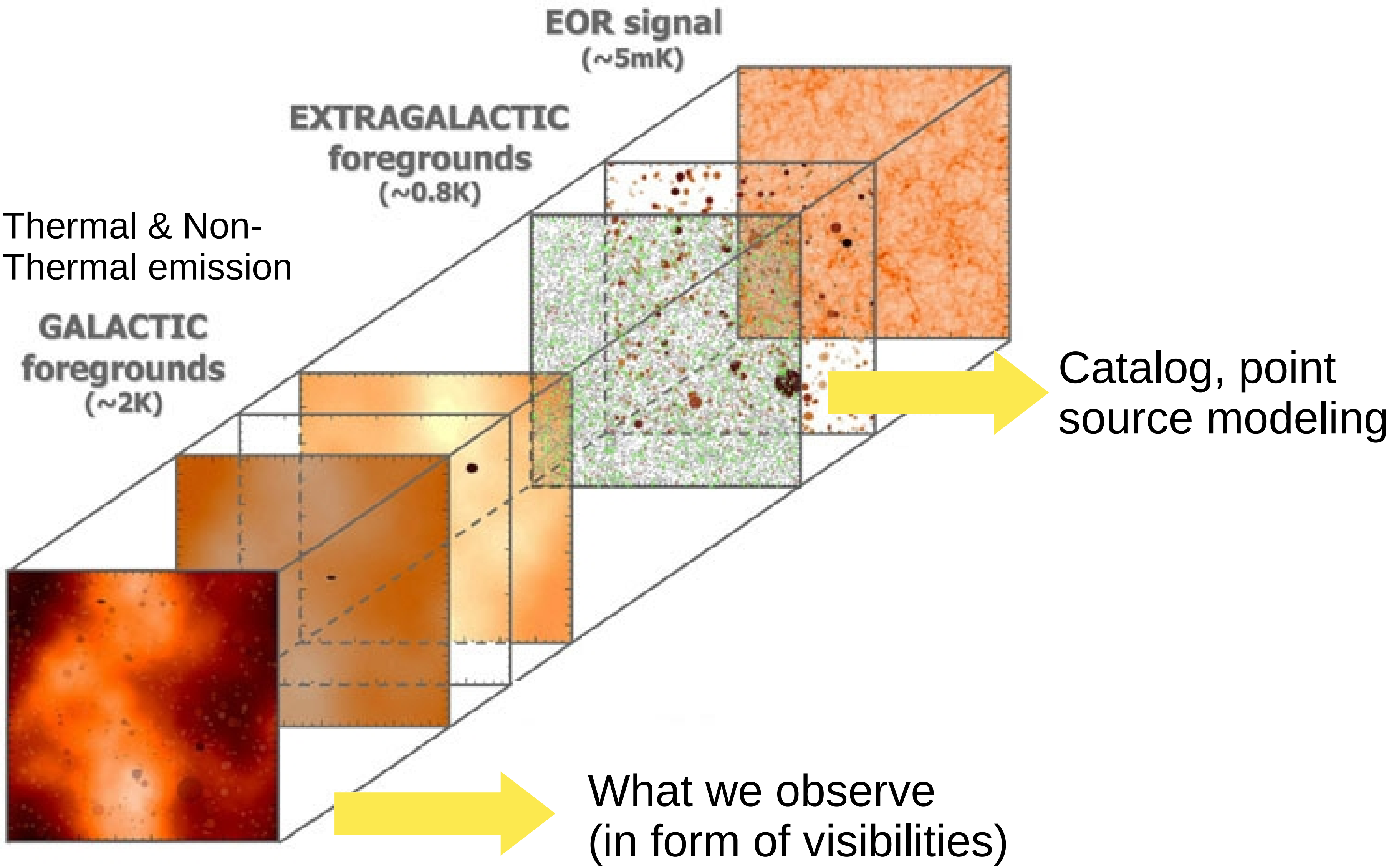
Quantity of intetrest for redshifted 21-cm observations:

$$\delta T_b(\mathbf{n}, z) = \frac{T(\tau) - T_\gamma}{1+z} \approx \frac{(T_s - T_\gamma)\tau}{1+z}$$

$$\delta T_b \approx 28\mathrm{mK} \underbrace{(1+\delta)}_{\text{Cosmology}} \underbrace{x_{HI} \frac{T_s - T_{CMB}}{T_s}}_{\text{Astrophysics}} \underbrace{\frac{\Omega_b h^2}{0.02} \left[\frac{0.24}{\Omega_m} \left(\frac{1+z}{10} \right) \right]^{\frac{1}{2}}}_{\text{Cosmology}}$$

* the product of $x_{HI}\delta$ enters the expression for δT_b, the 21 cm line is clearly sensitive to the correlations between density and ionization

Detecting 21-cm signal



Credit: Vibor
2010iska.meetE..28J

Detecting the 21-cm signal:

- Observed signal is composed of:
  * Galactic and extragalactic continuum smooth foregrounds.
  * Additioanl frequency structure may come from (instrument chromaticity, imperfect calibration, ionosphere…).
  * 21-cm signal.
  * Noise.

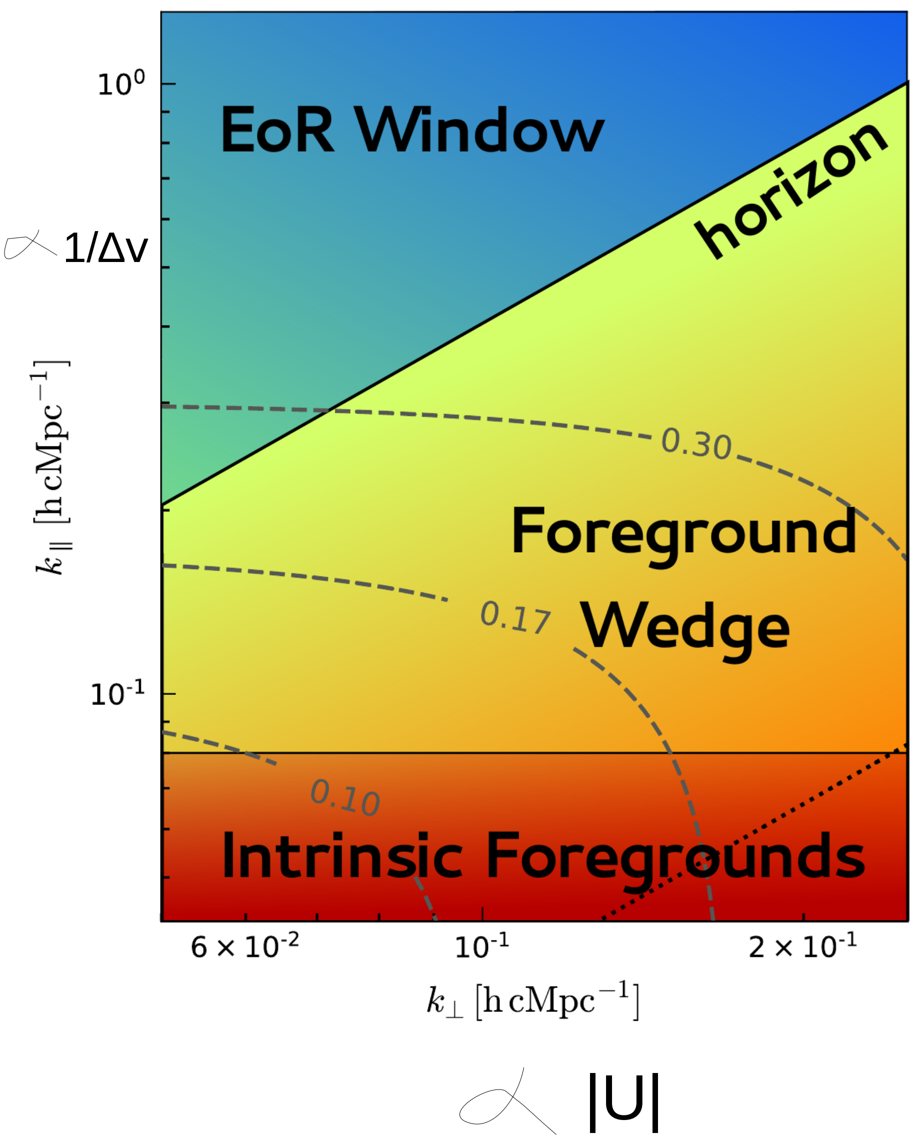
- HERA's SNR is significantly higher for $k < 0.15\ h\ Mpc^{-1}$, than for $k > 0.2\ h\ Mpc^{-1}$. The 21-cm signal also peaks at short baselines as do the Foregrounds !

- To model the foreground accurately, one needs to account for the different components of the observed data.

- This can be done using Gaussian Process Regression.

FG avoidance:

FG wedge & EoR Window:

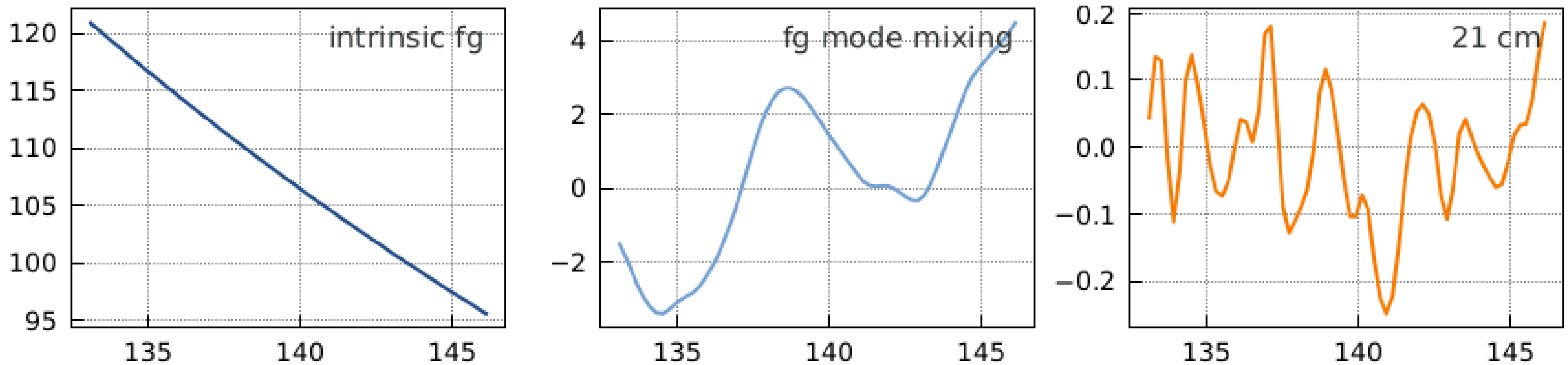


➢ FG removal with GPR:

- A Gaussian Process (GP) constitutes the generalization of the Gaussian distribution of random variables into the space of functions.
- Defined by its mean and covariance function.
- Posterior distribution gives fit to the data and confidence interval.
- Covariance function incorporate our knowledge on the data: smoothness, noise level...
- **d** = **f** + **n**, **f** are the true functional values where data is taken. In GPR the values of **f** follows a covariance structure, set by simple functional form such as Matern Kernel etc. The noise covariance $\boldsymbol{\Sigma}_n$ is used in the likelihood and $\boldsymbol{\Sigma}_f$ as a prior in the usual Bayessian sense.

GPR modeling for 21-cm experiments

Observed data can be decomposed in three main components:



Residual astrophysical sources: Smooth in frequency

Mode mixing: Less frequency smooth

21-cm signal: Uncorrelated ~ MHz

GPR: uses Gaussian Process (GP) as prior information $$\mathbf{f} \sim \mathcal{N}(0, K)$$

$$E(\mathbf{f}_{\rm fg}) = K_{\rm fg}\left[K_{\rm fg} + K_{21} + \sigma_n^2 I\right]^{-1}\mathbf{d}$$

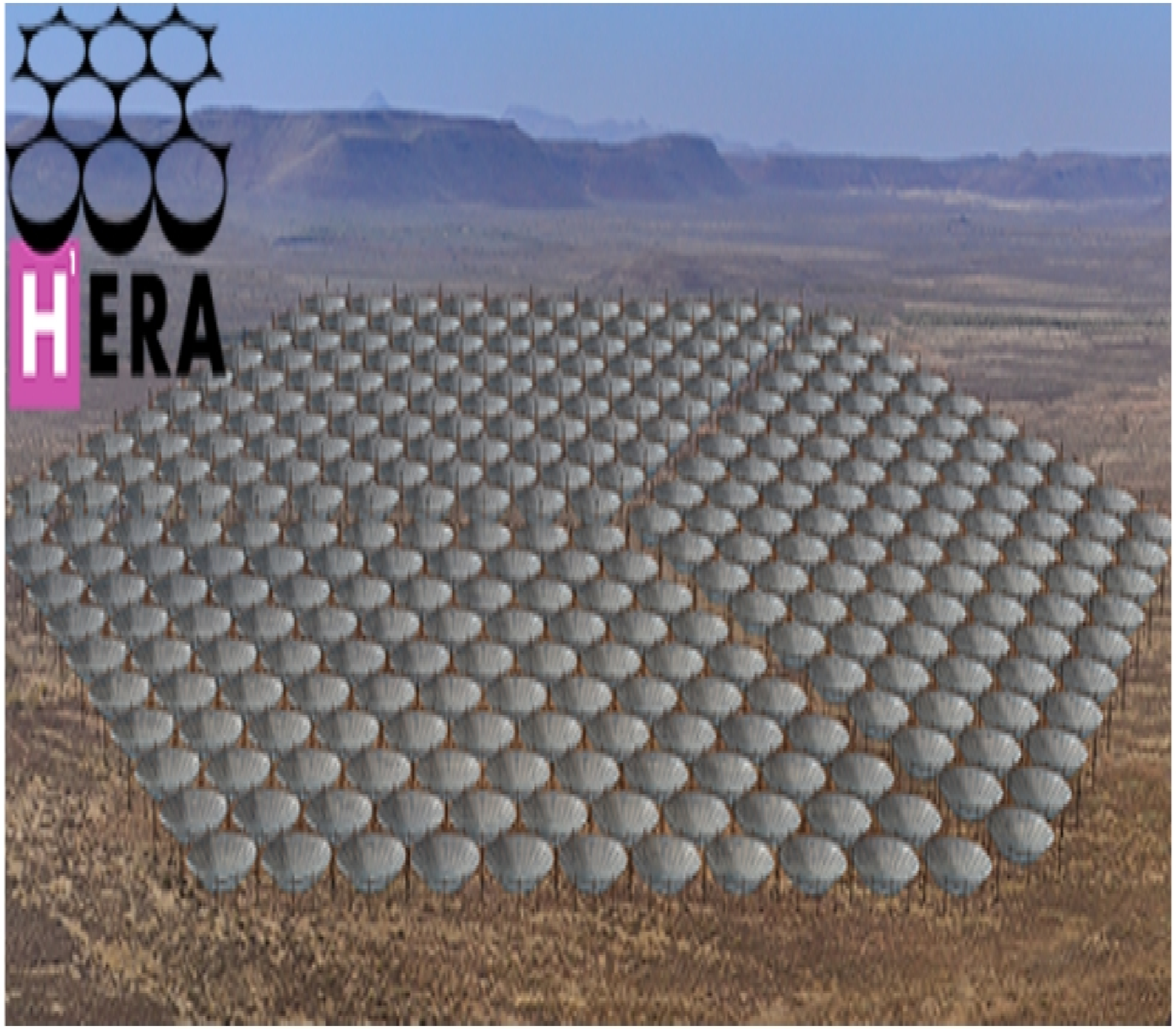
$$\mathrm{cov}(\mathbf{f}_{\rm fg}) = K_{\rm fg} - K_{\rm fg}\left[K_{\rm fg} + K_{21} + \sigma_n^2 I\right]^{-1} K_{\rm fg}$$

$$K_{\rm fg} = K_{\rm fg\_int} + K_{\rm fg\_mix}$$

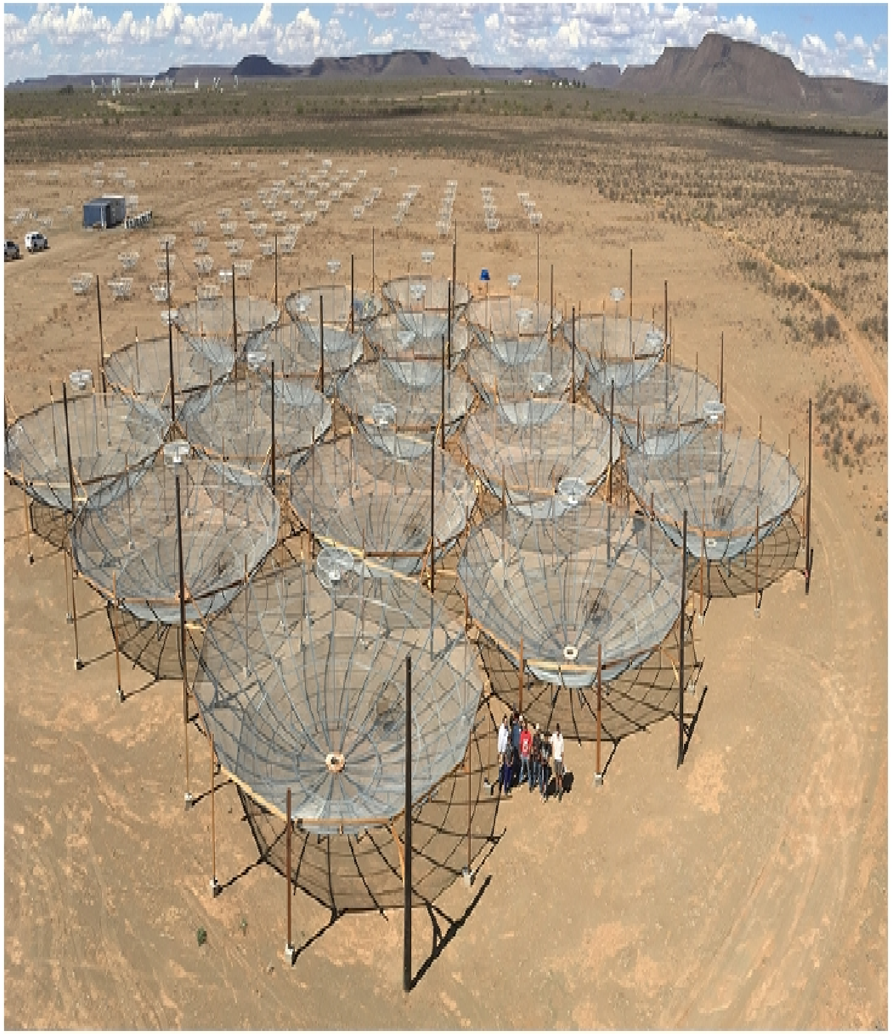
- Currently we only model in the frequency direction.
- Two parameters per co-variance functions: variance and frequency coherence-scale.
- Optimized by maximizing the marginal likelihood.

2018MNRAS.478.3640M
Mertens, F. G.; Ghosh, A.; Koopmans, L. V. E.

Full HERA



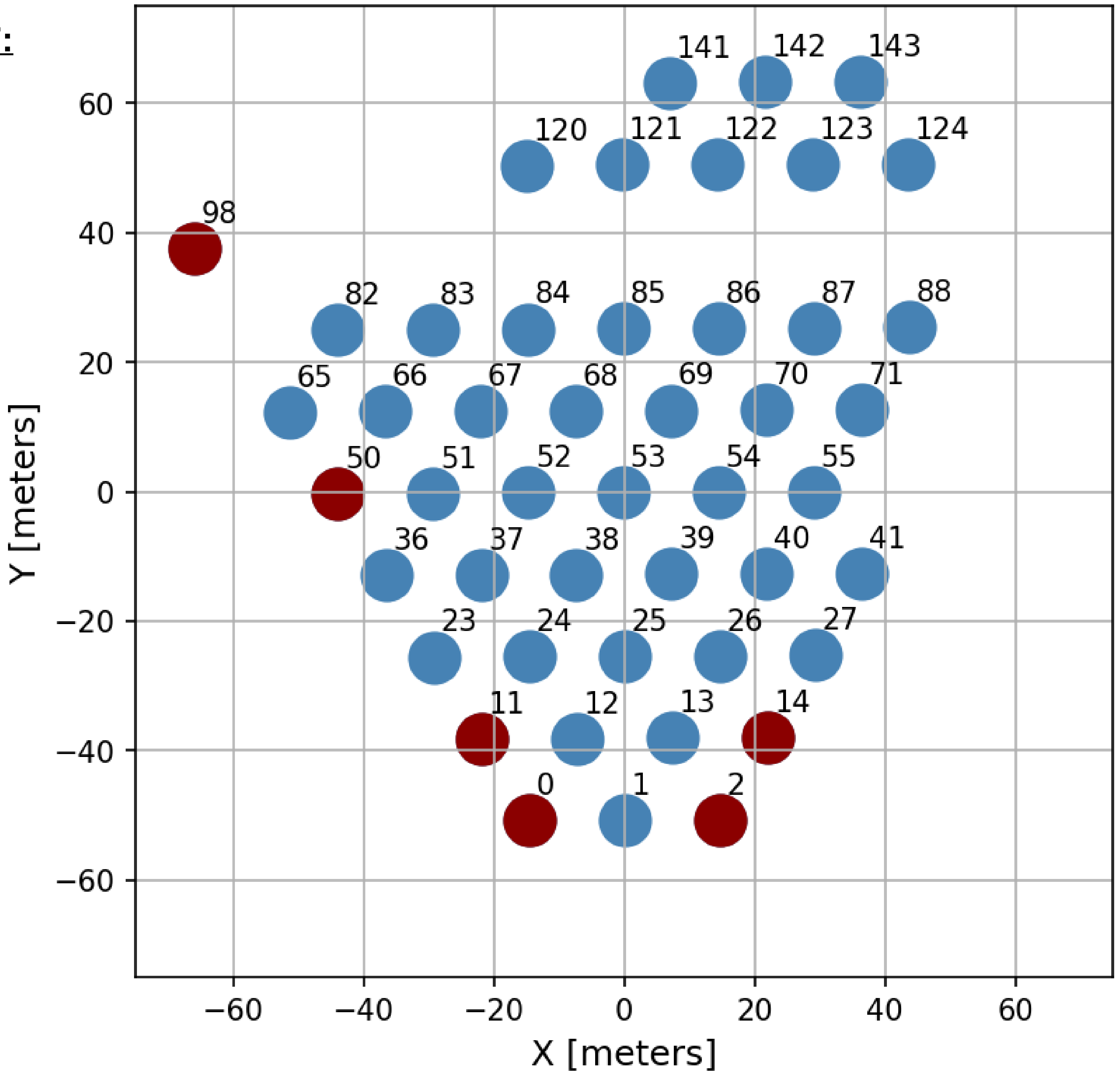
HERA-19



Goal : Statistical (direct ?) detection of Faint EoR Signal

- Redundancy increases SNR
- Measurements of same uv mode helps to build up a sky/model independent calibration
- Can in principle avoid incompleteness in sky model

HERA-47:



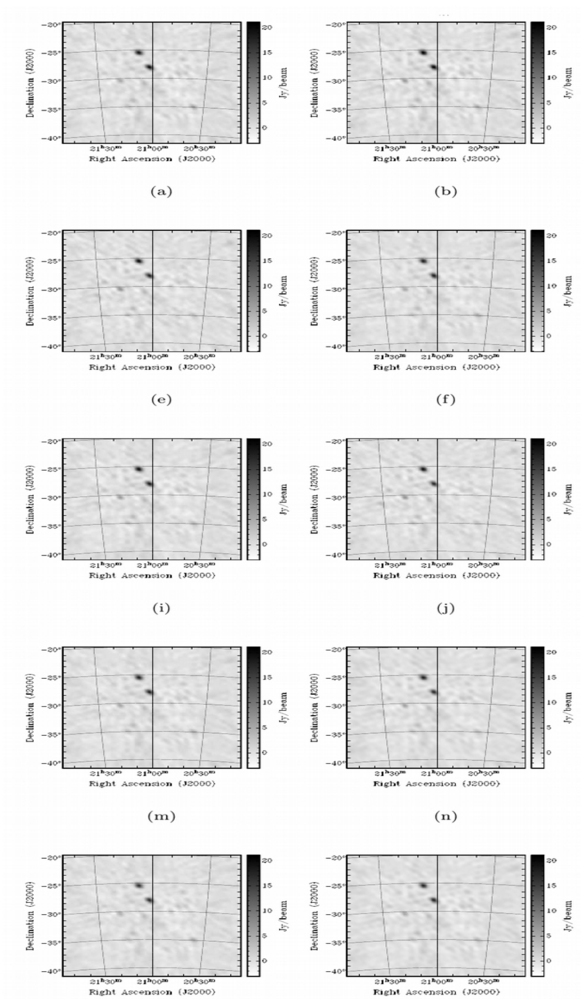
141
142
143
120
121
122
123
124
98
82
83
84
85
86
87
88
65
66
67
68
69
70
71
50
51
52
53
54
55
36
37
38
39
40
41
23
24
25
26
27
11
12
13
14
0
1
2
60
40
20
0
−20
−40
−60
Y [meters]
−60
−40
−20
0
20
40
60
X [meters]

GPR application on HERA-47:

Observations and data redunation:

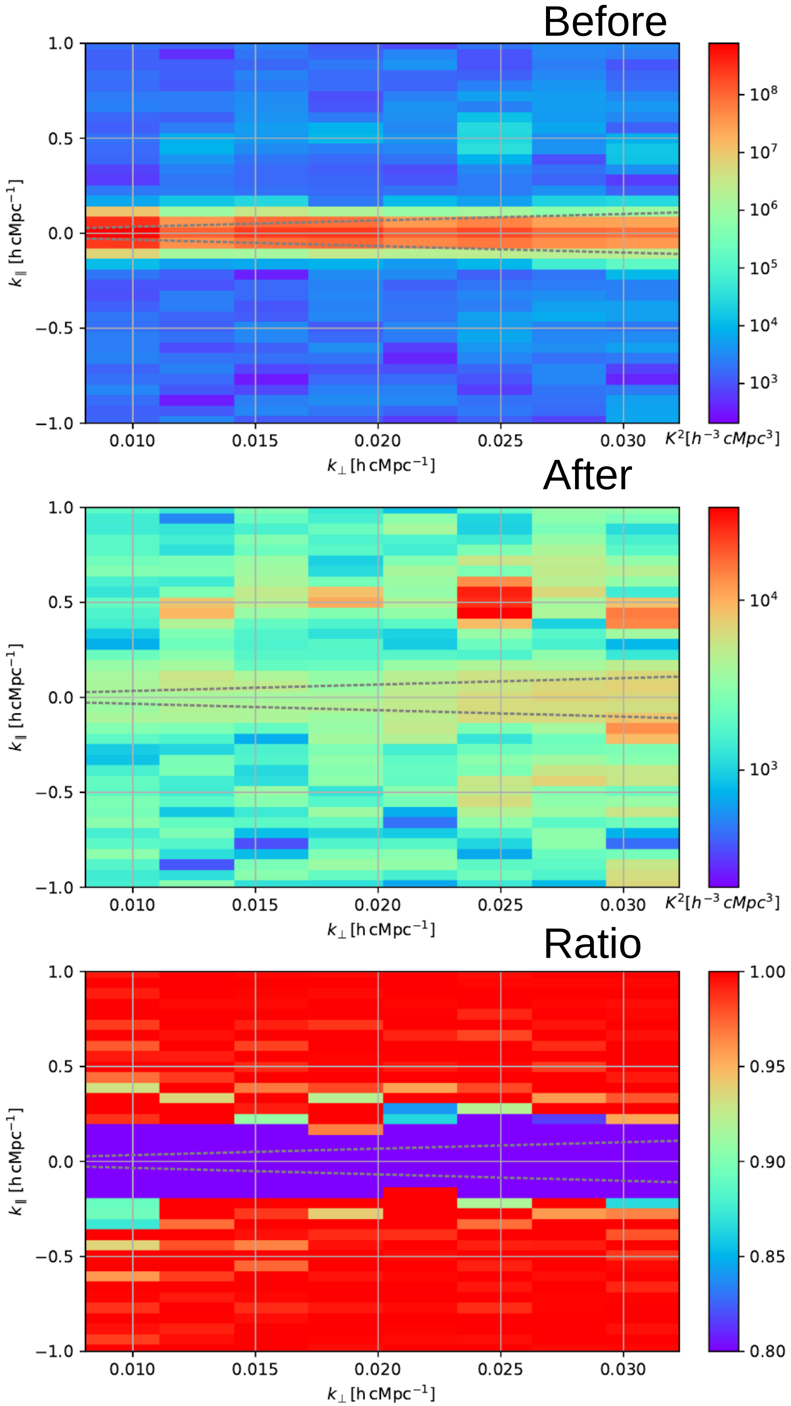
- We selected snapshots of 10 minutes of data close to 21 h.
- Calibration was performed using the brightest point sources within HERA’s FoV (N. Kern, Gianni, Chris Abscal memo -> PB corrected model).
- We used the sky model to solve for antenna based delay (K), complex gain for all the channels and whole 10 min interval (G) and a complex bandpass calibration (B).
- Calibration solutions from 1 night were applied to the rest of the 9 nights.

10 snapshot images from HERA-47 at 150 Mhz:



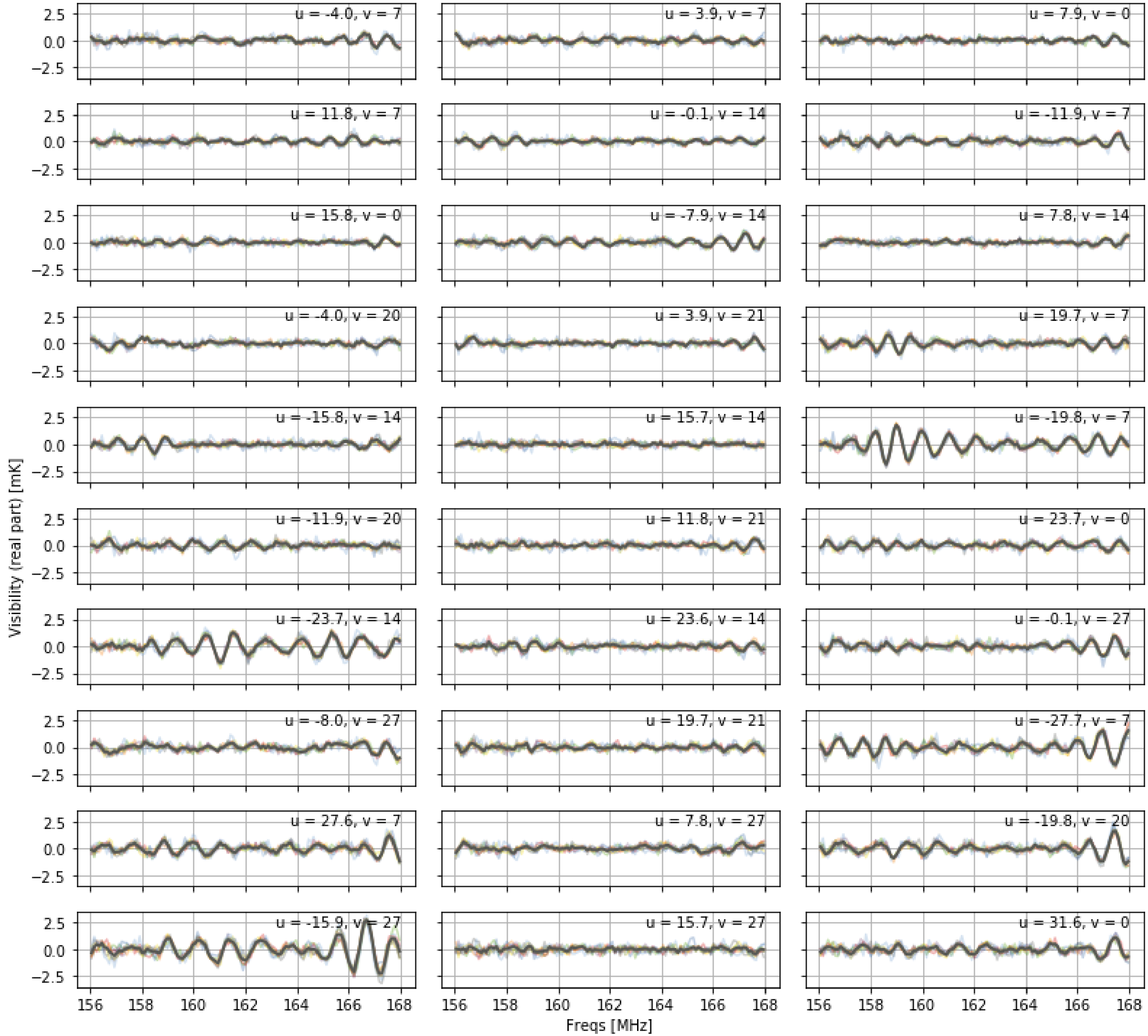
- 21^{o} X 21^{o} Images, with uniform weights & synthesized Beam of 43” X 35”.
- Off-source rms noise varies between 0.35 – 0.45 Jy/Beam

Cylindrical Power spectra (Before & After GPR):

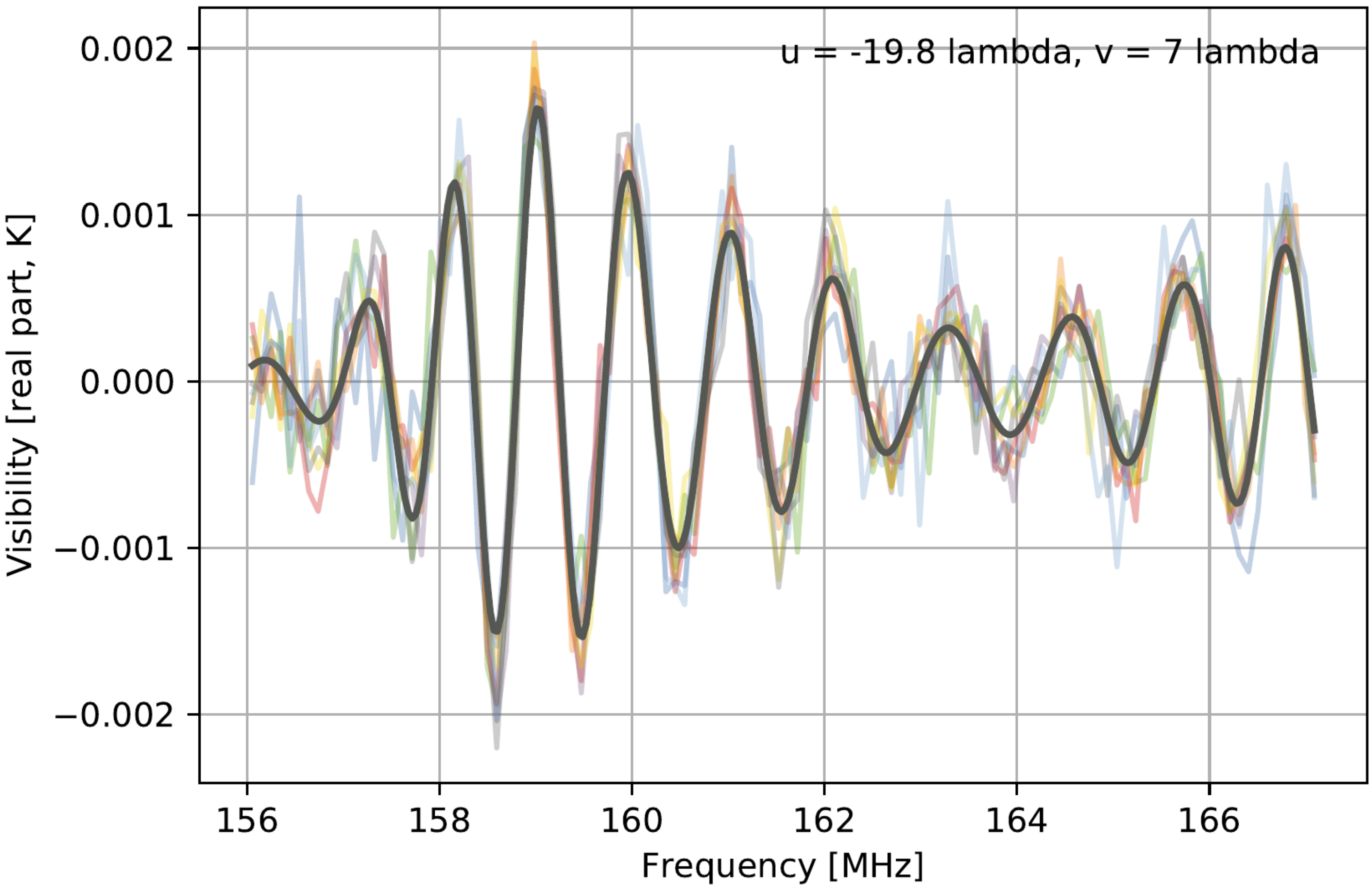


- We noticed 4-5 order of FG supression within the horizon line.
- The wedge like structure in the residuals are mainly due to higher noise levels at longer baselines (fewer redundant base lines).
  Also FG covariance estimation is altered by the periodic signal which is not modeled here. When it is included in the covariance model, the residual lowered in the wedge.

Baseline dependent periodic signal in residuals:



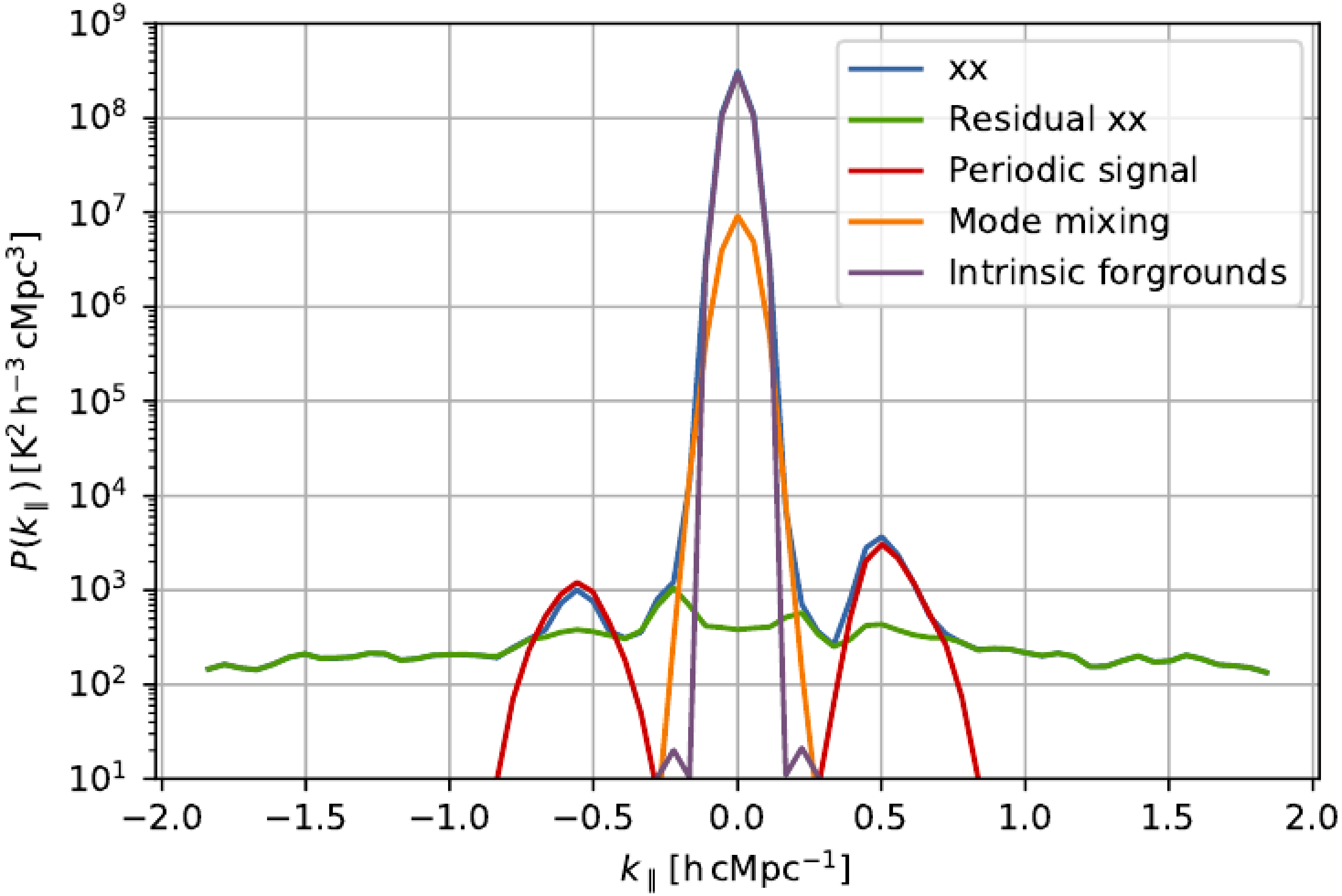
Nature of the periodic signal for a 38.6 m baseline:



- Periodic signal in frequency with a varying amplitude and period.
- We used an additional kernel in the FG covariance model, a combination of RBF and cosine kernel.

$$K_{\mathrm{fg}} = K_{\mathrm{sky}} + K_{\mathrm{mix}} + K_{\mathrm{per}} \qquad \kappa_{\mathrm{per}}(\nu_p, \nu_q) = \sigma^2_{\mathrm{per}} \exp(-\frac{r^2}{2l^2_{\mathrm{per}}}) \cos\left(\frac{2\pi r}{p_{per}}\right)$$

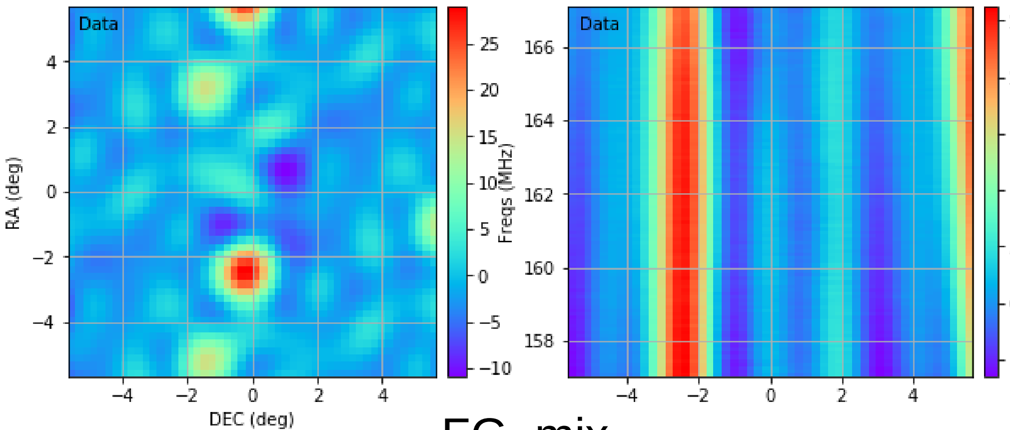
Filtering the periodic signal with GPR:



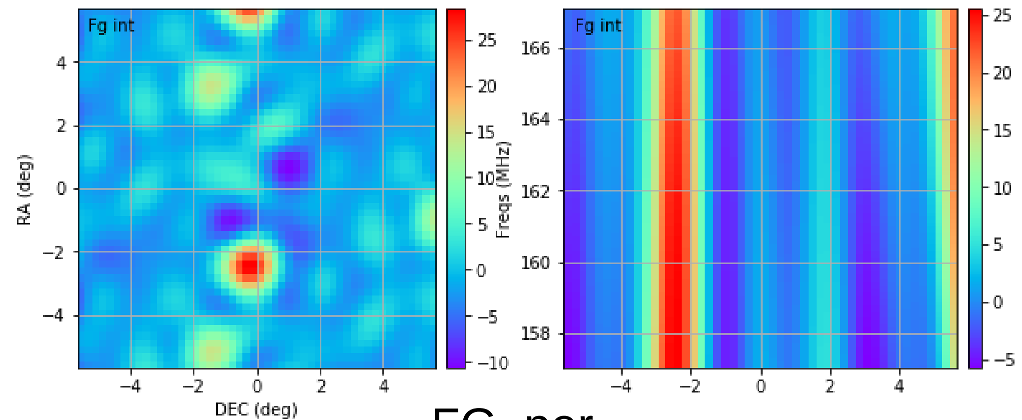
- Two remaining peaks at $k_{||}$ ~ 0.25 h Mpc^{-1} is mainly present for the 14.6 m baselinese, otherwise the residual is noise like.

ML Images of FG components:

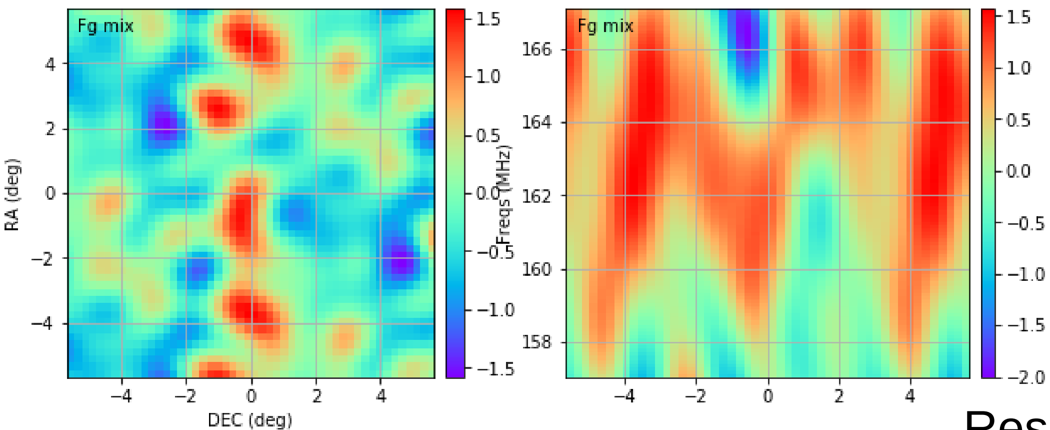
Data



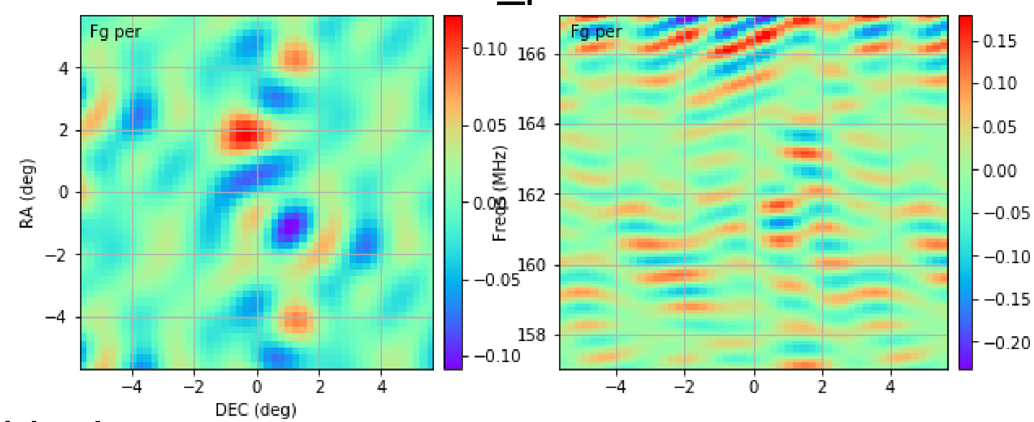
FG_Int



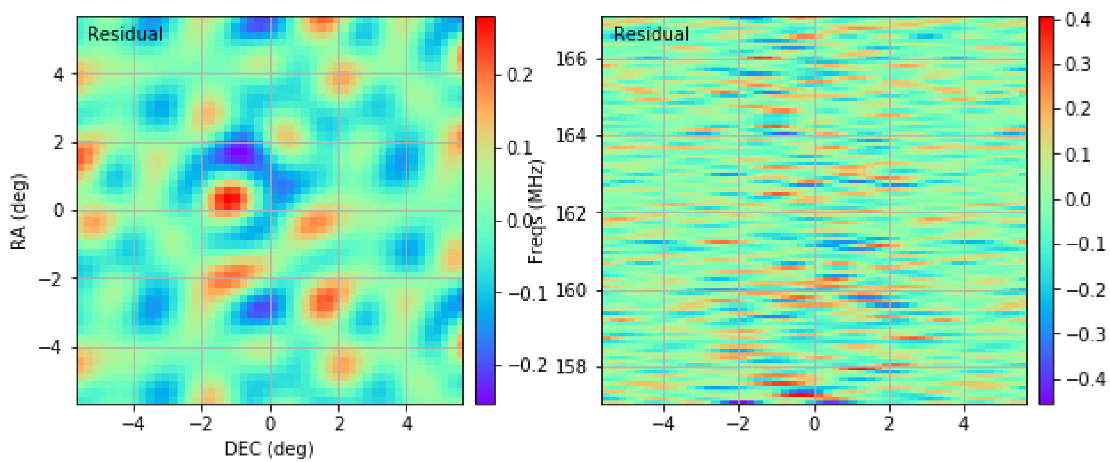
FG_mix



FG_per



Residual



FG model hyper-parameter uncertainty:

Posterior probability distribution of GP model hyperparameters

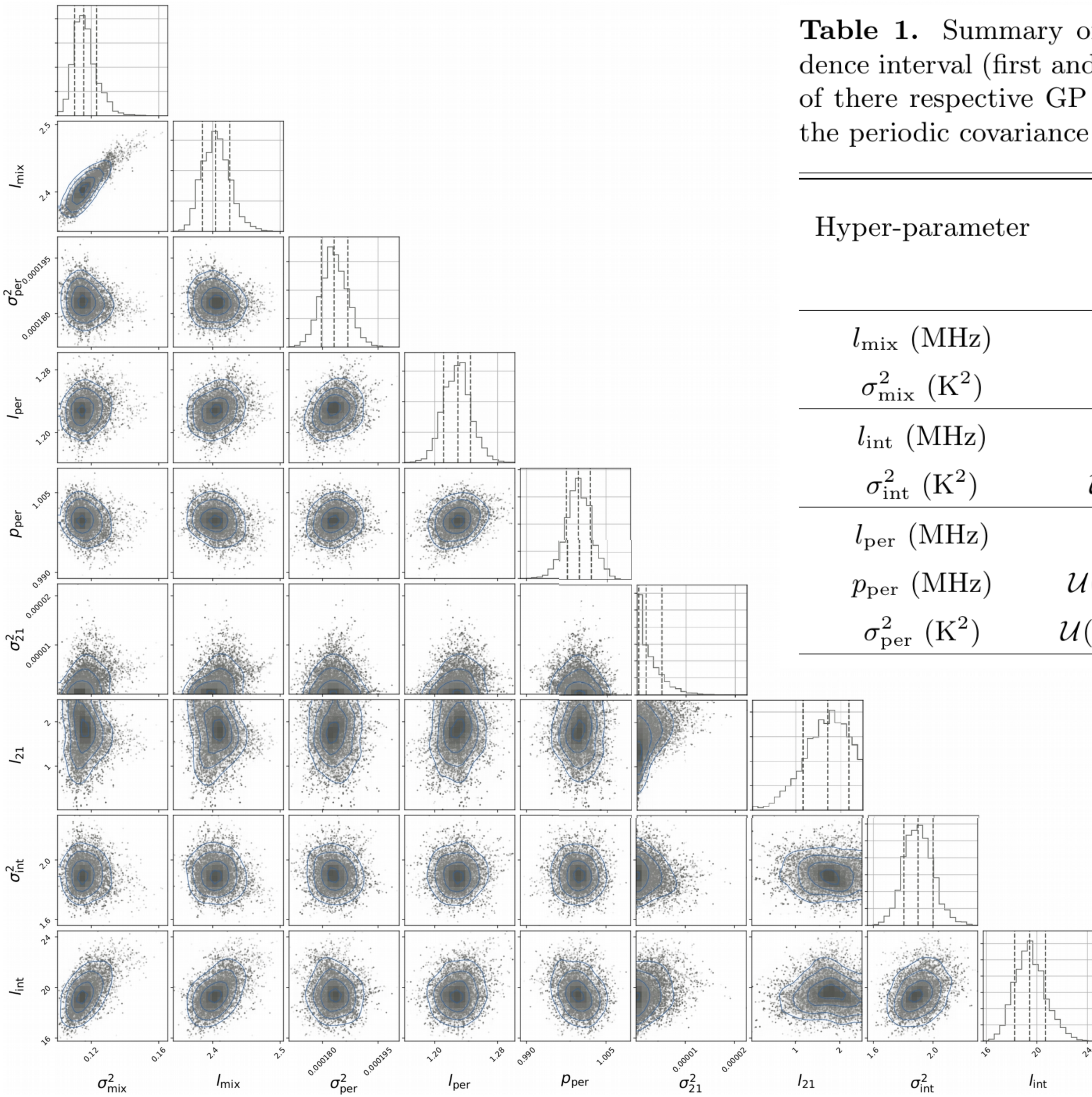


Table 1. Summary of the estimated median and confidence interval (first and third quantile levels (Q1 and Q3)) of there respective GP model hyper-parameters including the periodic covariance kernel

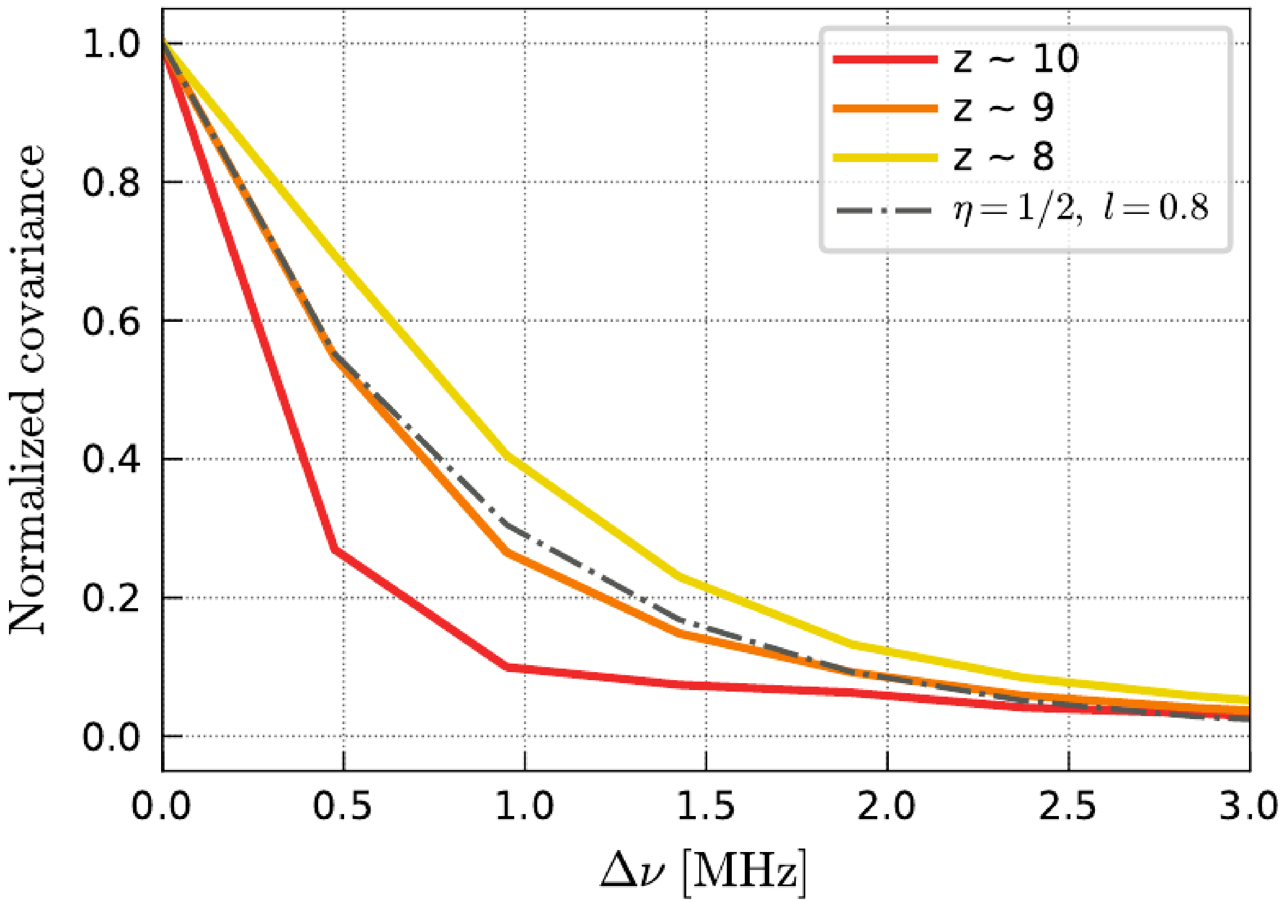
Hyper-parameter	Prior	Estimate
$l_{\rm mix}$ (MHz)	$\mathcal{U}(2, 20)$	$2.40^{+0.02}_{-0.01}$
$\sigma^2_{\rm mix}$ ($\rm K^2$)	$\mathcal{U}(0.1, 0.9)$	$0.115^{+0.007}_{-0.005}$
$l_{\rm int}$ (MHz)	$\mathcal{U}(10, 200)$	$19.42^{+1.25}_{-1.18}$
$\sigma^2_{\rm int}$ ($\rm K^2$)	$\mathcal{U}(0.02, 2.5)$	$1.89^{+0.10}_{-0.09}$
$l_{\rm per}$ (MHz)	$\mathcal{U}(1, 5)$	$1.23^{+0.01}_{-0.01}$
$p_{\rm per}$ (MHz)	$\mathcal{U}(0.628, 1.256)$	$0.999^{+0.002}_{-0.002}$
$\sigma^2_{\rm per}$ ($\rm K^2$)	$\mathcal{U}(0.00001, 0.01)$	$0.000183^{+0.000004}_{-0.000003}$

21-cm signal injection tests:

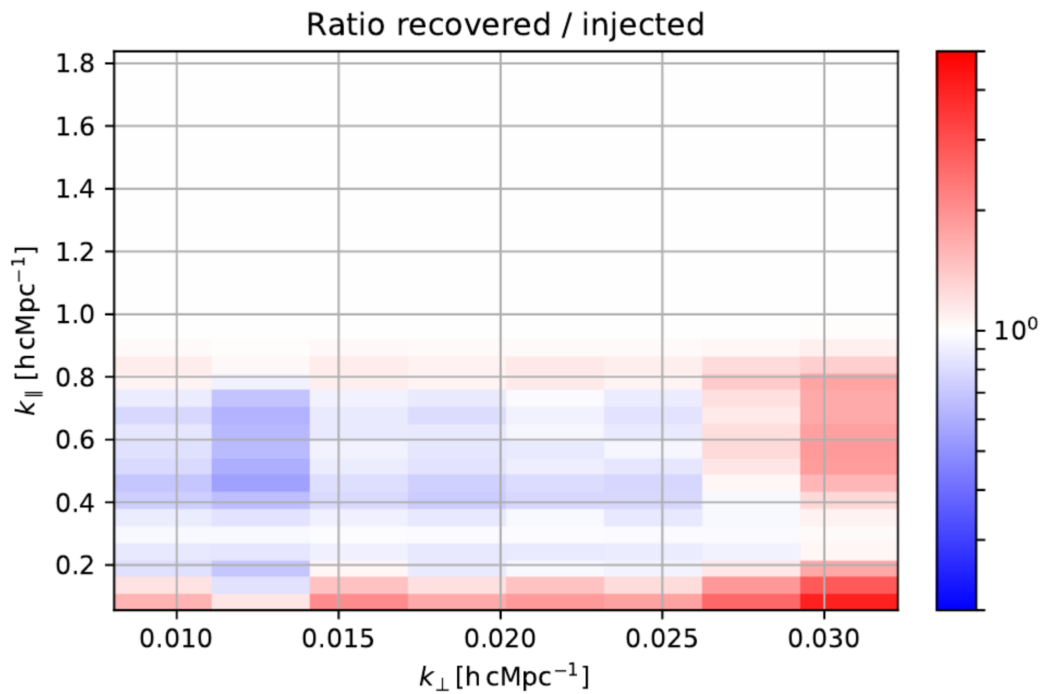
$$V_{\mathrm{rec}} = \mathcal{F}_{\mathrm{GPR}}(V_{\mathrm{data}} + V_{\mathrm{inj}}) - \mathcal{F}_{\mathrm{GPR}}(V_{\mathrm{data}}).$$

21cm signal is simulated with a GP Exponential covariance function with a frequency covariance scale of 0.8 Mhz and variance $\sigma^2_{\mathrm{inj}} = 0.01\sigma^2_{\mathrm{noise}}$

GP Exponential covariance Vs Simulated EoR signal:



Injected Vs Recovered signal ps:



- We found 47% under-subtraction in FG dominated part, $k < 0.25$ hMpc^{-1}
- 11% over-subtraction in periodic signal dominated regions, $0.35 < k < 0.7$ hMpc^{-1}

Summary and Outlook:

- GPR offers a complete statistical description of all components contributing to the observed signal: To-Do: improve the accuracy of foregrounds modeling.

- The optimized 'coherence-scale' of the 'sky' and 'mix' covariance kernel are about 20 Mhz and 2.4 Mhz. For the periodic kernel, GPR estimate of the coherence scale l_{per} ~ 1.2 Mhz and the period p_{per} ~ 1.0 Mhz.

- Through foreground modeling and subtraction, we achieved more than four order of foreground suppression in the `foreground window' and in the `EoR window' (corresponding to $|k_{||}|$ > horizon line + 0.1 h $cMpc^{-1}$) the foreground contamination decreases by factor of ~ 1.9 compared to a foreground avoidance scheme.

- Improve FG covariance model by integrating: The instrument chromaticity, Calibration error, Ionospheric disturbances. Tests with HERA simulations.

- Improve 21-cm covariance model using wide range of simulations. Tests performance using full HERA simulation and with 21-cm signal injection test.

- Optimize GPR model for several redshifts at once. Include baseline dependence in the covariance kernel and accomodate this as part of optimization.