

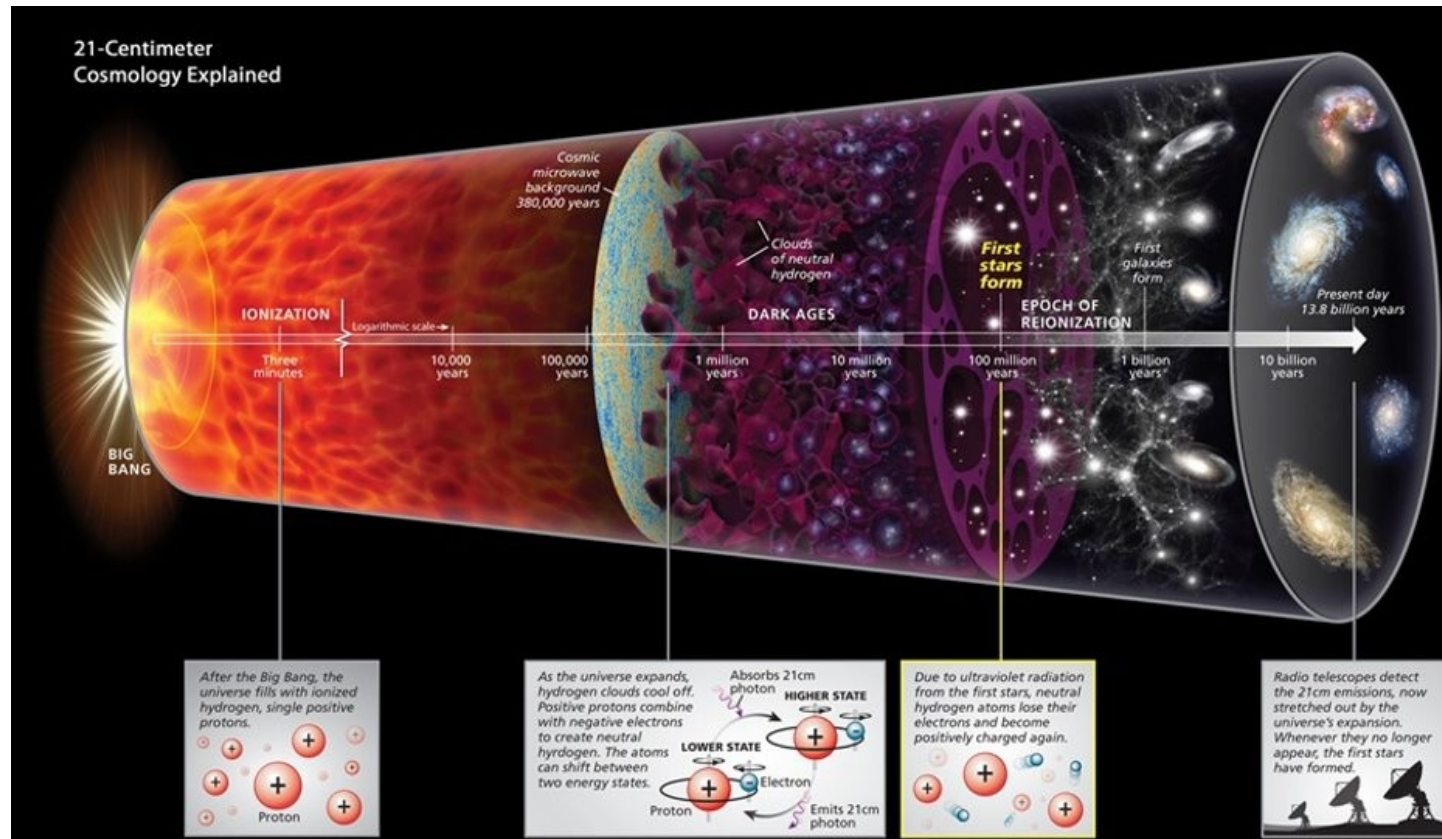
# Detecting the First Stars with REACH and Bayesian Statistics

Dominic Anstey | University of Cambridge

# Summary

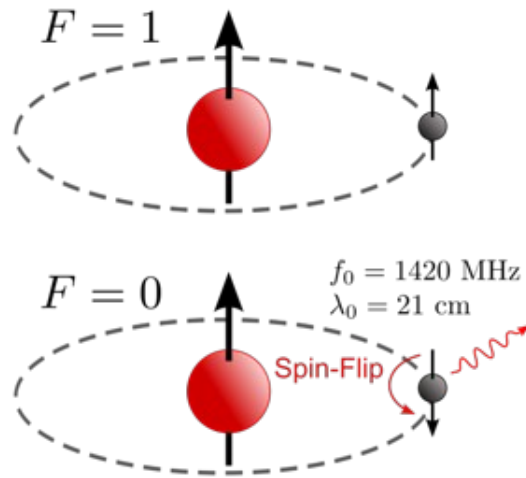
- Overview of 21-cm Cosmology
- Current State of the Field
- Challenges
- REACH
- Data Analysis Techniques

# Aims of 21-cm Cosmology



Roen Kelly, Discover Magazine

# The Hydrogen 21-cm Line



[https://en.wikipedia.org/wiki/Hydrogen\\_line](https://en.wikipedia.org/wiki/Hydrogen_line)

- Kinetic Gas Temperature  $T_K$

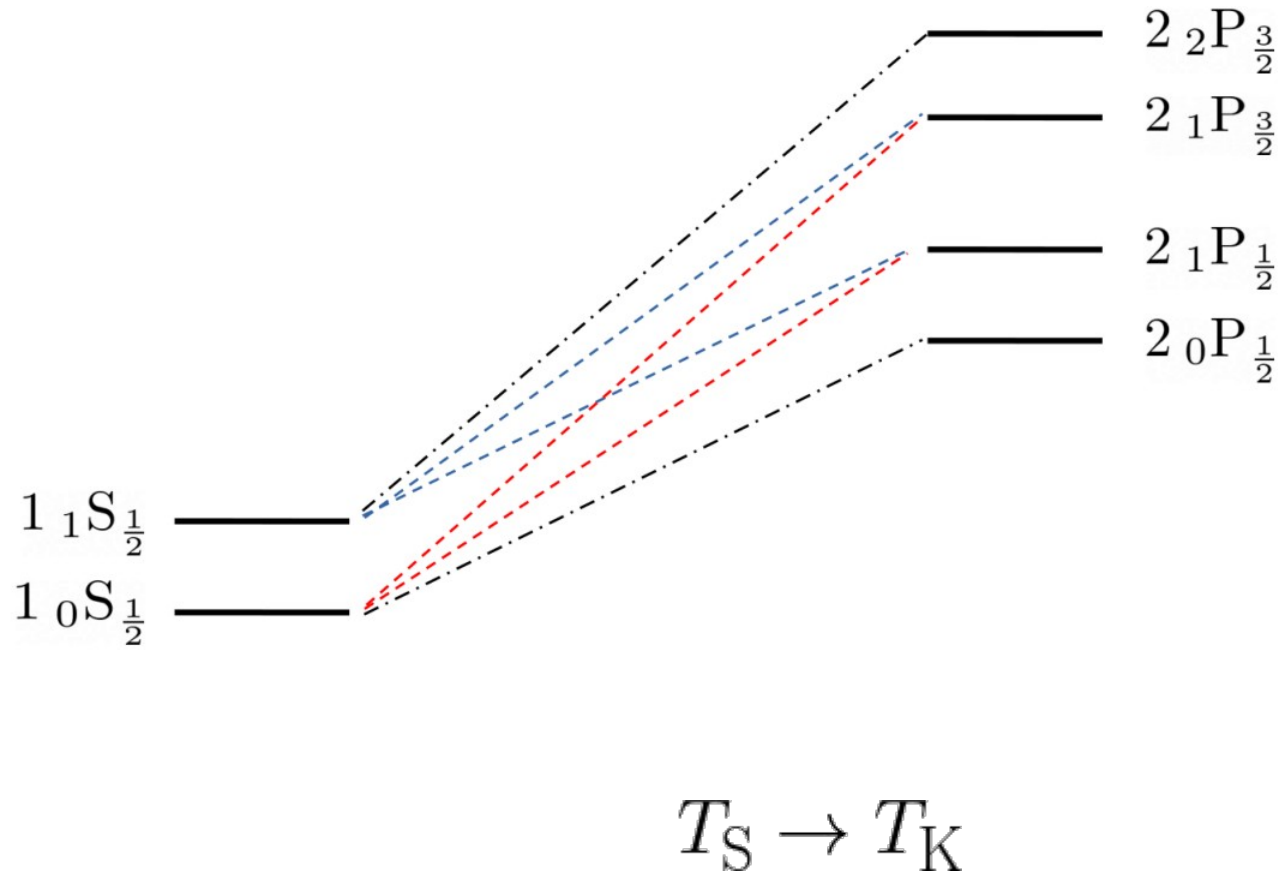
- Background Radiation Temperature  $T_{\text{CMB}}$

- Spin Temperature  $T_S$   $\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-\frac{T^*}{T_S}}$

In equilibrium  $T_S \rightarrow T_{\text{CMB}}$

# Wouthuysen-Field Effect

- Wouthuysen (1952)
- Field (1958)



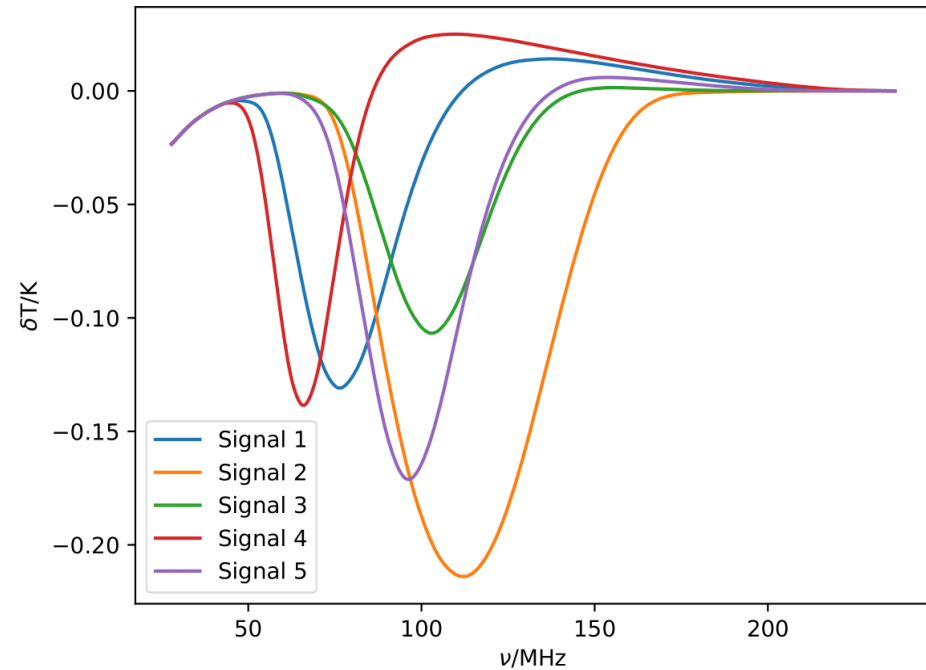
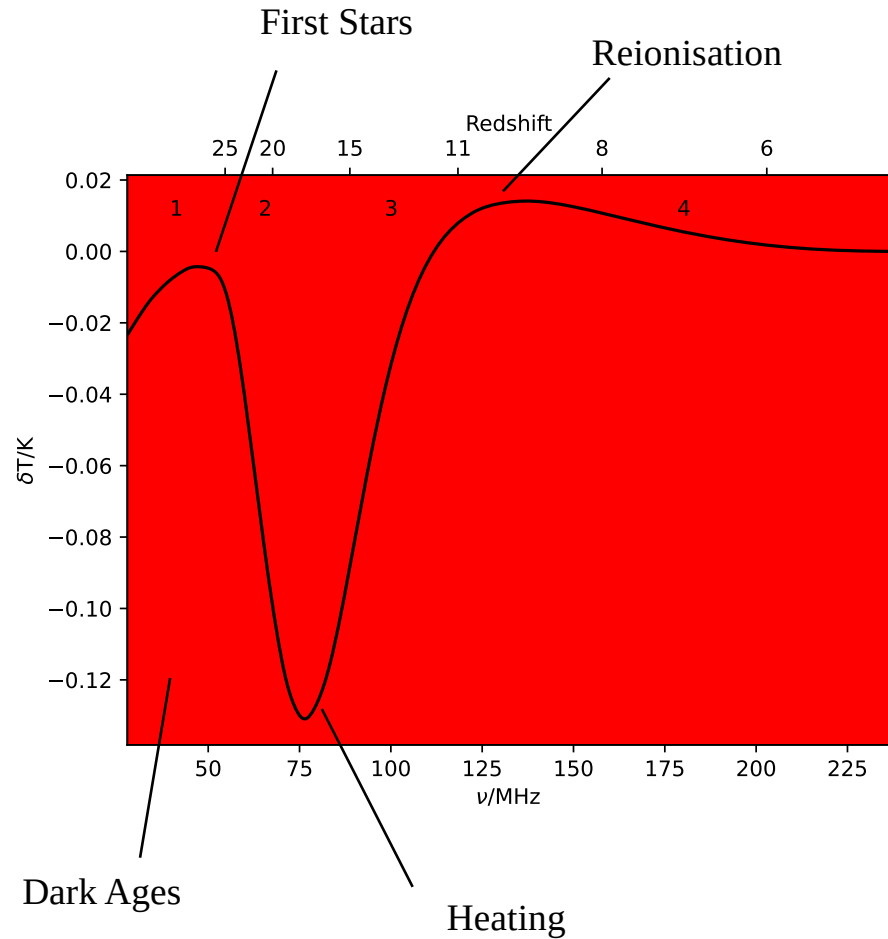
# Other Effects

- Collisions  $T_S \rightarrow T_K$

- Ionisation  $\bar{x}_i \rightarrow 1$

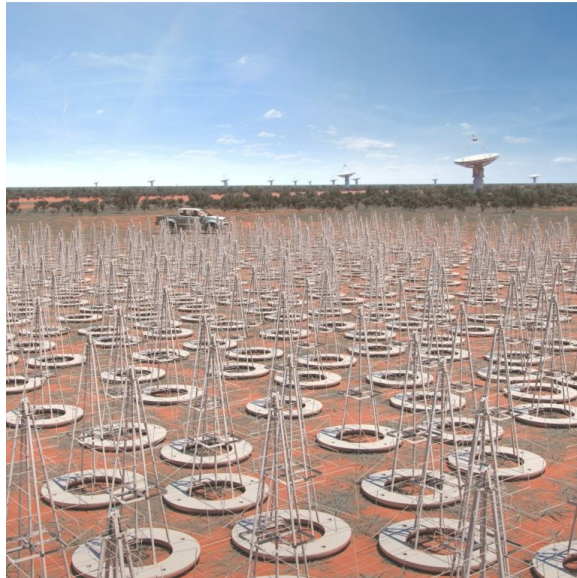
$$\delta T_b \approx 27 (1 - \bar{x}_i) \left( \frac{T_S - T_{\text{CMB}}}{T_S} \right) \left( \frac{1+z}{10} \right)^{\frac{1}{2}} \text{mK}$$

# The 21-cm Signal



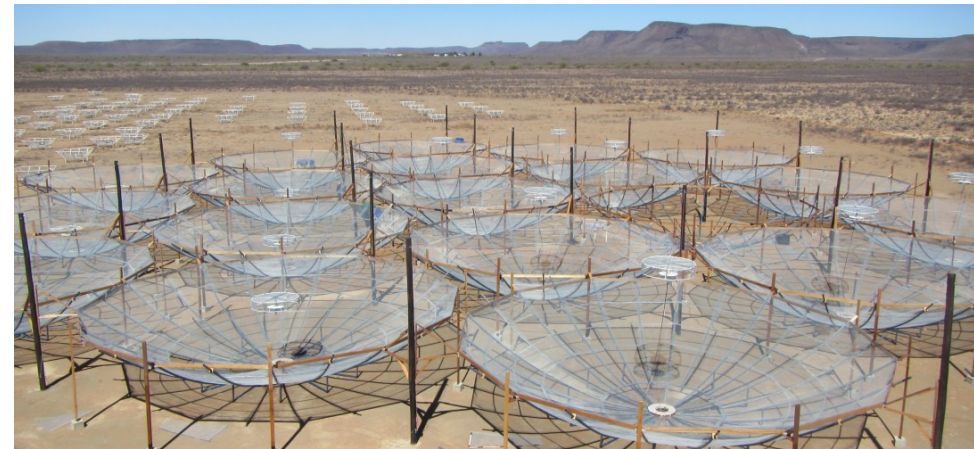
# Interferometers

Square Kilometre Array  
(SKA)



<https://www.skatelescope.org/layout/>

Hydrogen Epoch of Reionization Array (HERA)



<https://www.reionization.org>



# Global Experiments

- Experiment to Detect the Global EoR Signature (EDGES)
- Shaped Antenna measurement of the background RAdio Spectrum (SARAS)
- Large-Aperture Experiment to Detect the Dark Ages (LEDA)
- Radio Experiment for the Analysis of Cosmic Hydrogen (REACH)
- Probing Radio Intensity at high-Z from Marion (PRI<sup>2</sup>M)
- Etc.

# EDGES

## LETTER

doi:10.1038/nature25792



### An absorption profile centred at 78 megahertz in the sky-averaged spectrum

Judd D. Bowman<sup>1</sup>, Alan E. E. Rogers<sup>2</sup>, Raul A. Monsalve<sup>1,3,4</sup>, Thomas J. Mozdzen<sup>1</sup> & Nivedita Mahesh<sup>1</sup>

After stars formed in the early Universe, their ultraviolet light is expected, eventually, to have penetrated the primordial hydrogen gas and altered the excitation state of its 21-centimetre hyperfine line. This alteration would cause the gas to absorb photons from the cosmic microwave background, producing a spectral distortion that should be observable today at radio frequencies of less than 200 megahertz<sup>1</sup>. Here we report the detection of a flattened absorption profile in the sky-averaged radio spectrum, which is centred at a frequency of 78 megahertz and has a best-fitting full-width at half-maximum of 19 megahertz and an amplitude of 0.5 kelvin. The profile is largely consistent with expectations for the 21-centimetre signal induced by early stars; however, the best-fitting amplitude of the profile is more than a factor of two greater than the largest predictions<sup>2</sup>. This discrepancy suggests that either the primordial gas was much colder than expected or the background radiation temperature was hotter than expected. Astrophysical phenomena (such as radiation from stars and stellar remnants) are unlikely to account for this discrepancy; of the proposed extensions to the standard model of cosmology and particle physics, only cooling of the gas as a result of interactions between dark matter

The absorption profile is found by fitting the integrated spectrum with the foreground model and a model for the 21-cm signal simultaneously. The best-fitting 21-cm model yields a symmetric U-shaped absorption profile that is centred at a frequency of  $78 \pm 1$  MHz and has a full-width at half-maximum of  $19^{+4}_{-2}$  MHz, an amplitude of  $0.5^{+0.5}_{-0.2}$  K and a flattening factor of  $\tau = 7^{+5}_{-3}$  (where the bounds provide 99% confidence intervals including estimates of systematic uncertainties; see Methods for model definition). Uncertainties in the parameters of the fitted profile are estimated from statistical uncertainty in the model fits and from systematic differences between the various validation trials that were performed using observations from both instruments and several different data cuts. The 99% confidence intervals that we report are calculated as the outer bounds of (1) the marginalized statistical 99% confidence intervals from fits to the primary dataset and (2) the range of best-fitting values for each parameter across the validation trials. Fitting with both the foreground and 21-cm models lowers the residuals to an r.m.s. of 0.025 K. The fit shown in Fig. 1 has a signal-to-noise ratio of 37, calculated as the best-fitting amplitude of the profile divided by the statistical uncertainty of the amplitude fit, including the cova-

# EDGES

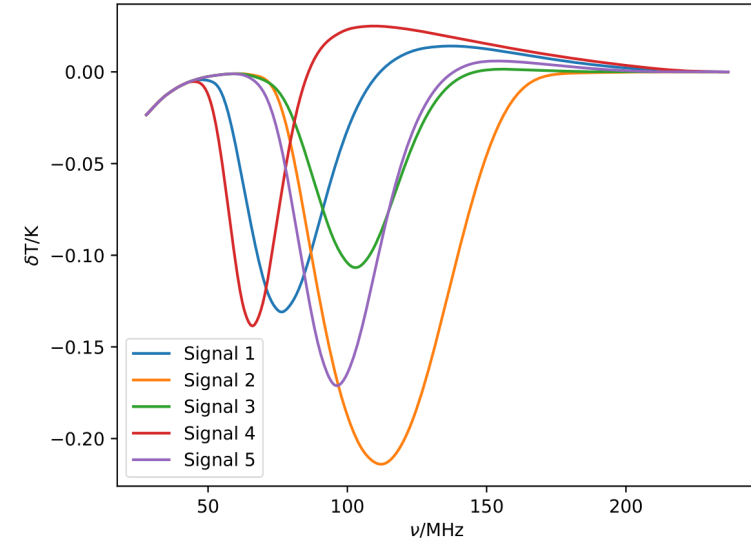
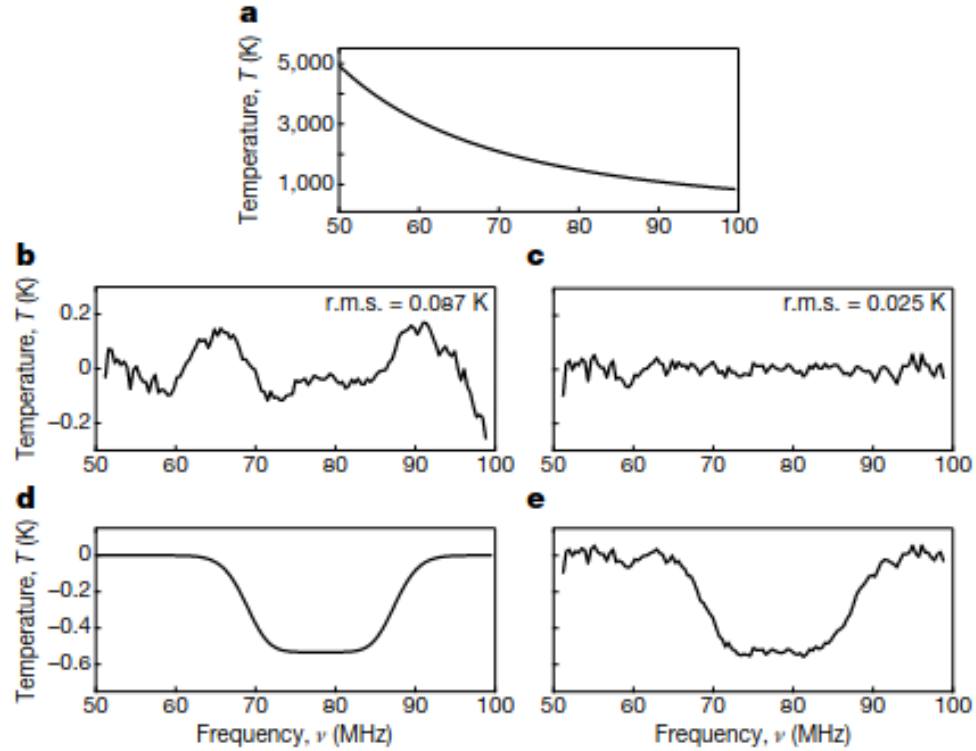


Figure 1 of Bowman et al. (2018)

# Possible Explanations

## Increased Radio Background

- AGNs
- Radiative decay (e.g. Bolliet et al. 2021)
- Annihilating dark matter (e.g. Fraser et al. 2018)
- Superconducting cosmic strings (e.g. Brandenberger et al. 2019)

## Lowered Gas Temperature

- Dark matter interactions (e.g. Liu et al. 2019, Barkana et al. 2018, Berlin et al. 2018)

# Doubts in Data Analysis

- Hills et al. (2018)
- Singh and Subrahmanyan (2019)
- Bevins et al. (2021)

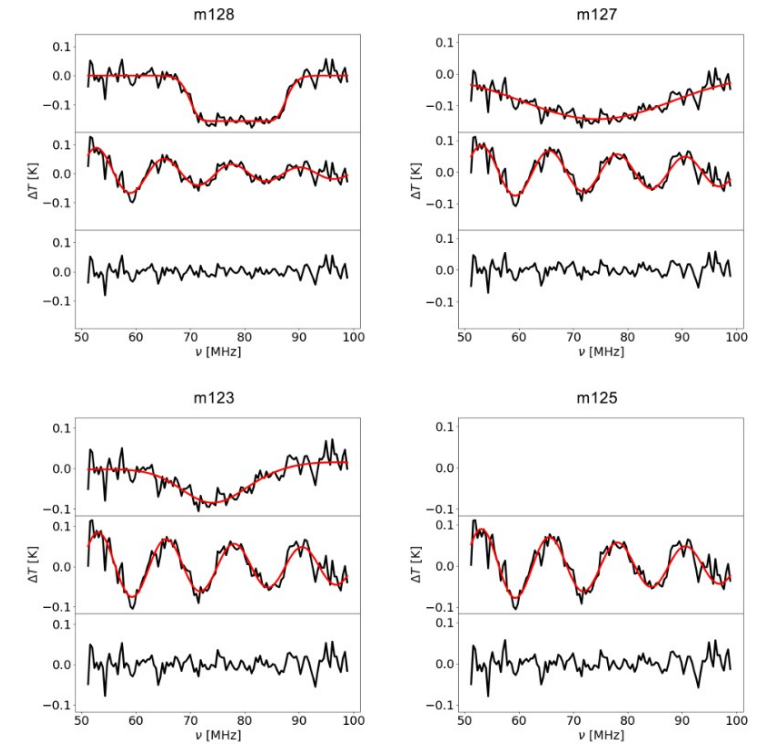
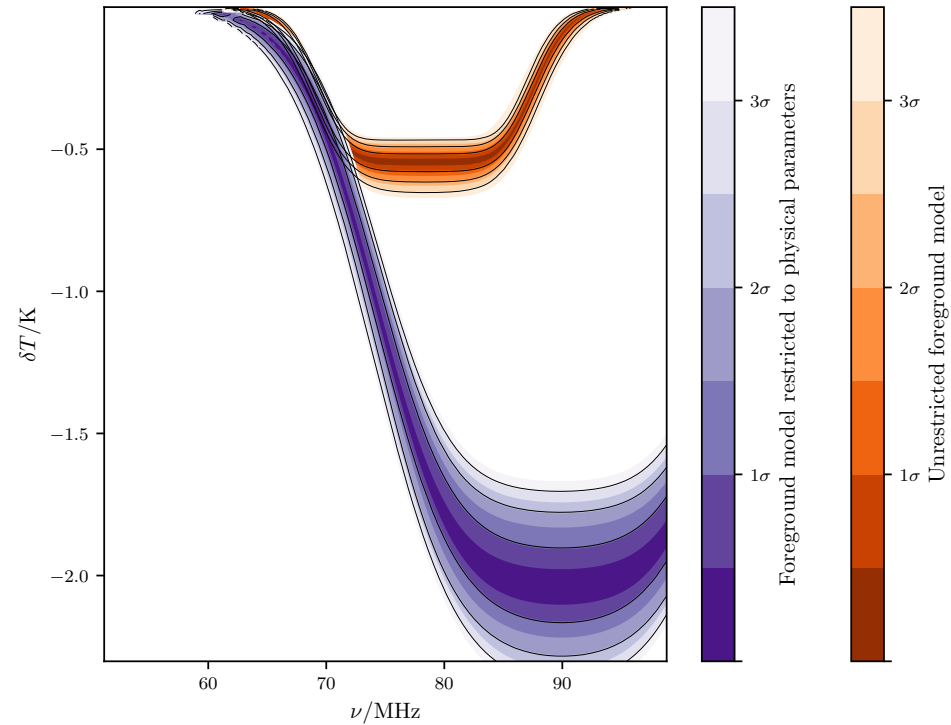


Figure 2 from Sims and Pober (2019)

# SARAS Results



## On the detection of a cosmic dawn signal in the radio background

Authors: Saurabh Singh<sup>1,2,3\*</sup>, Jishnu Nambissan T.<sup>1,4</sup>, Ravi Subrahmanyam<sup>1,5</sup>,  
N. Udaya Shankar<sup>1</sup>, B. S. Girish<sup>1</sup>, A. Raghunathan<sup>1</sup>, R. Somashekar<sup>1</sup>, K. S. Srivani<sup>1</sup> &  
Mayuri Sathyanarayana Rao<sup>1</sup>

<sup>1</sup> Raman Research Institute, C V Raman Avenue, Sadashivanagar, Bangalore 560080, India.

<sup>2</sup> Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A 2T8, Canada.

<sup>3</sup> McGill Space Institute, McGill University, 3550 rue University, Montréal, QC H3A 2A7, Canada.

<sup>4</sup> International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia

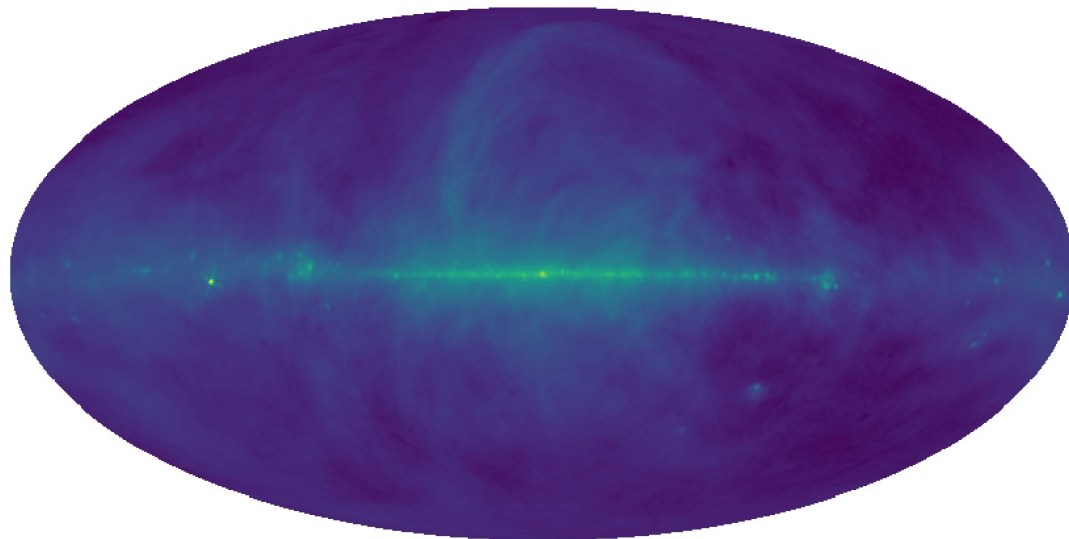
<sup>5</sup> Space & Astronomy, CSIRO, PO Box 1130, Bentley, WA 6102, Australia

\* Corresponding author. Email: saurabhs@rri.res.in

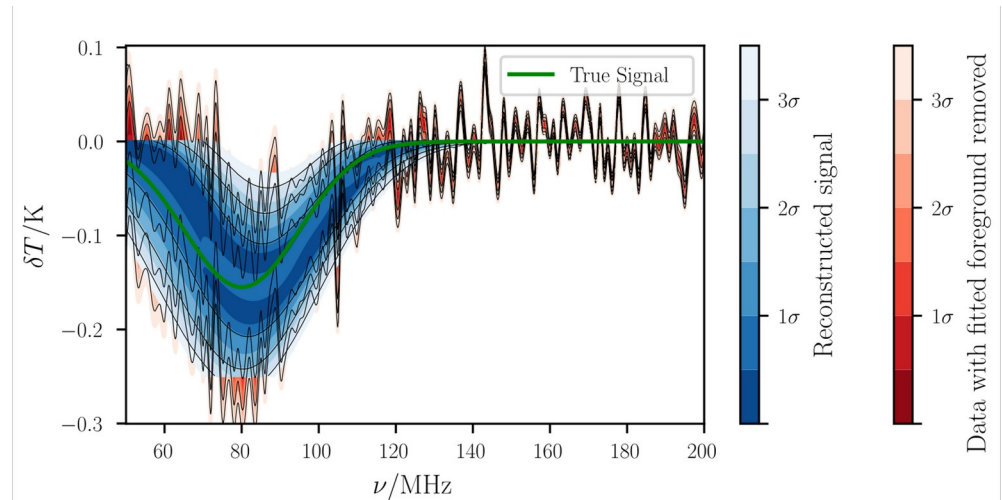
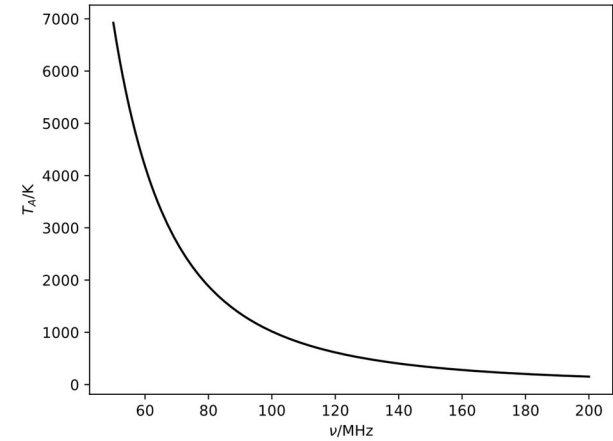
o-ph.COJ 13 Dec 2021

# Bright Foregrounds

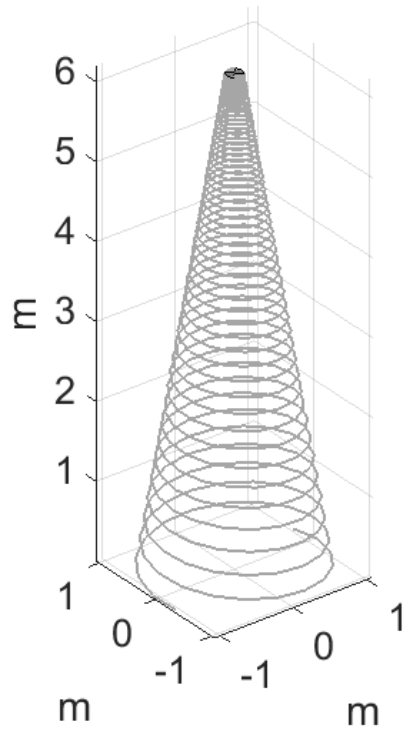
408 MHz



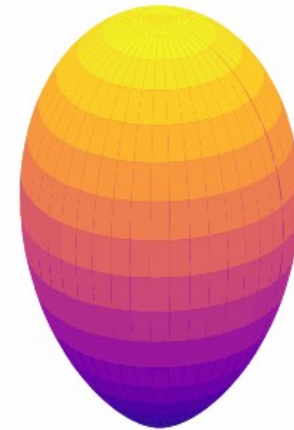
Remazeilles et al. (2015)



# Chromatic Distortion



Log Spiral



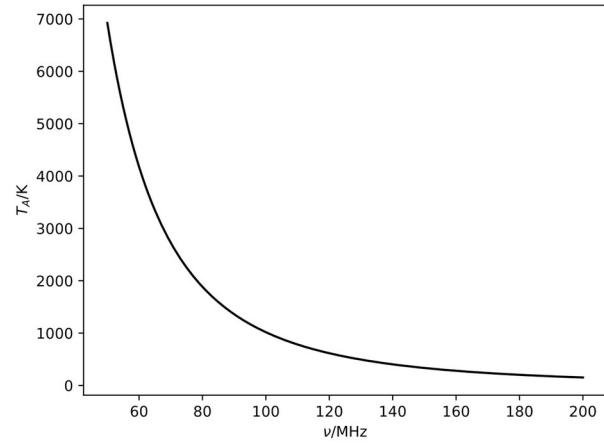
Antenna simulation and image  
provided by Quentin Guerring

$\nu=50\text{MHz}$

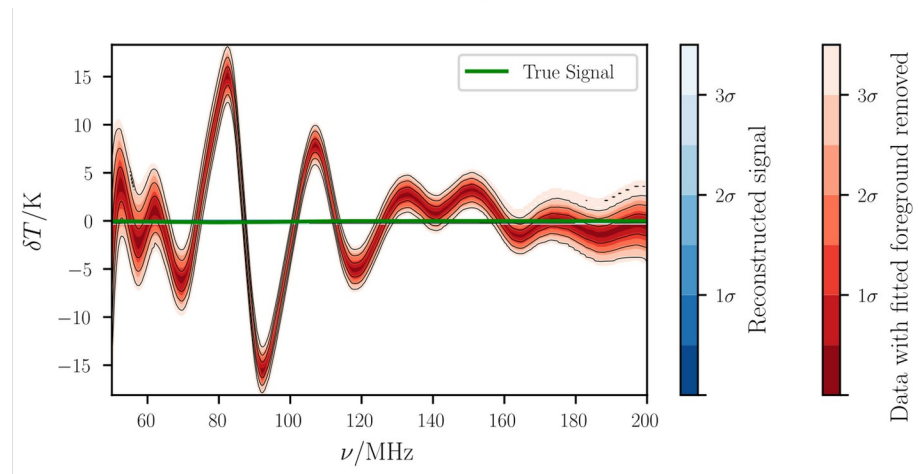
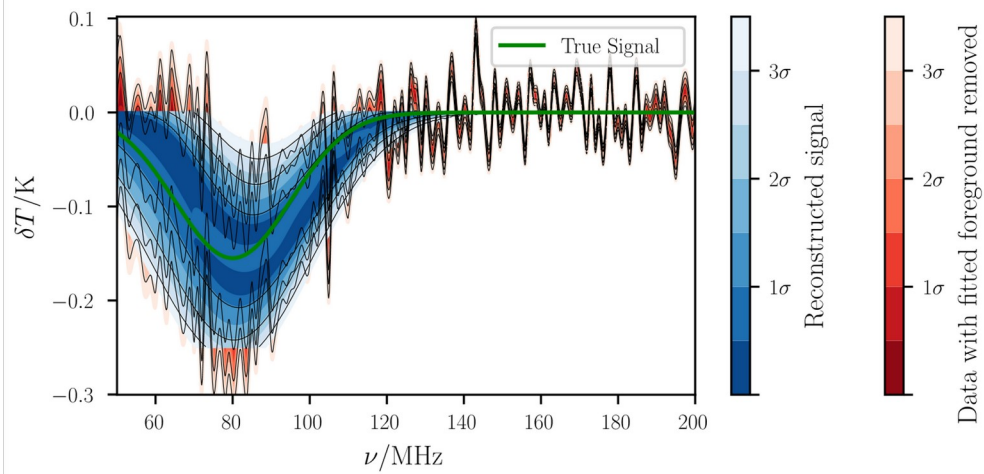
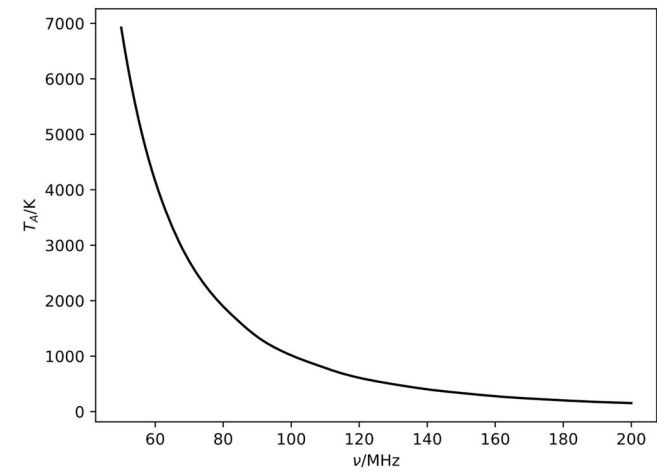


# Chromatic Distortion

Achromatic antenna

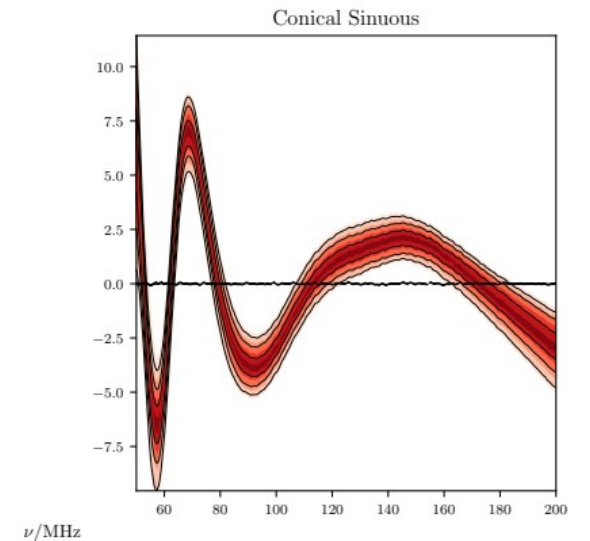
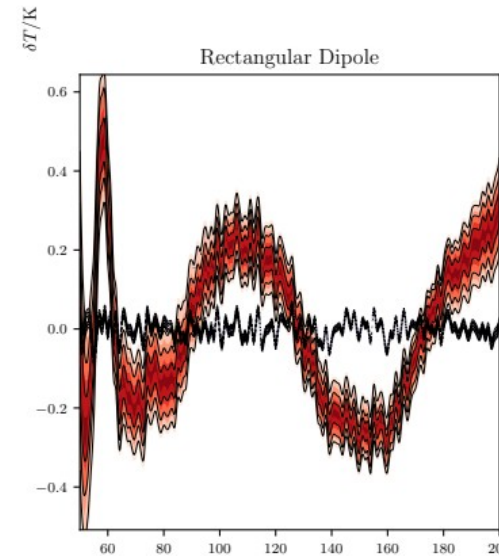
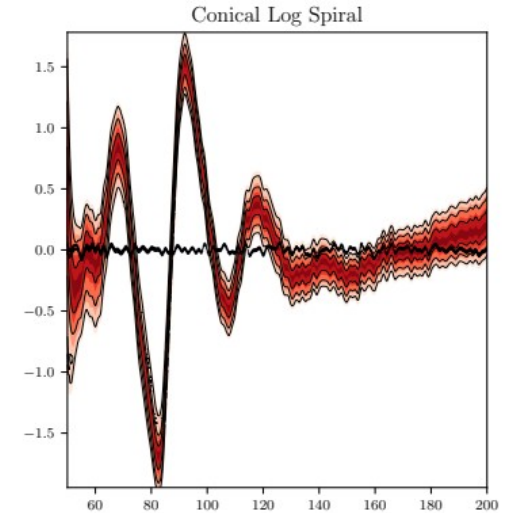
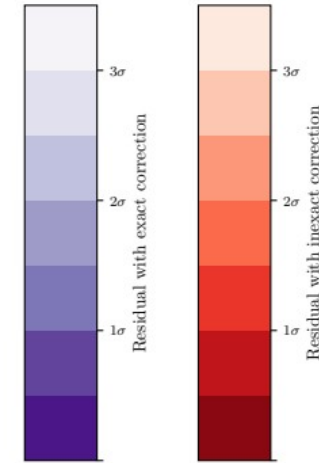


Chromatic antenna

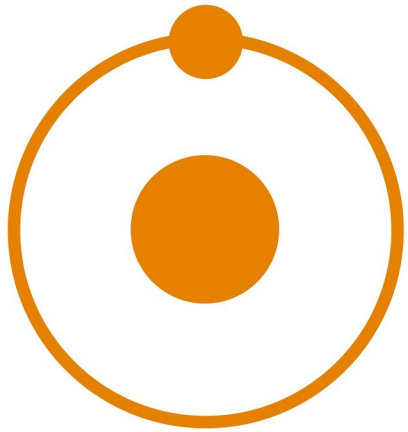


# Chromatic Correction Methods

$$B_{\text{factor}}(\nu, t) = \frac{\int D(\Omega, \nu) T_{\text{sky}}(\Omega, t) d\Omega}{\int D(\Omega, \nu_{\text{ref}}) T_{\text{sky}}(\Omega, t) d\Omega}$$



# REACH

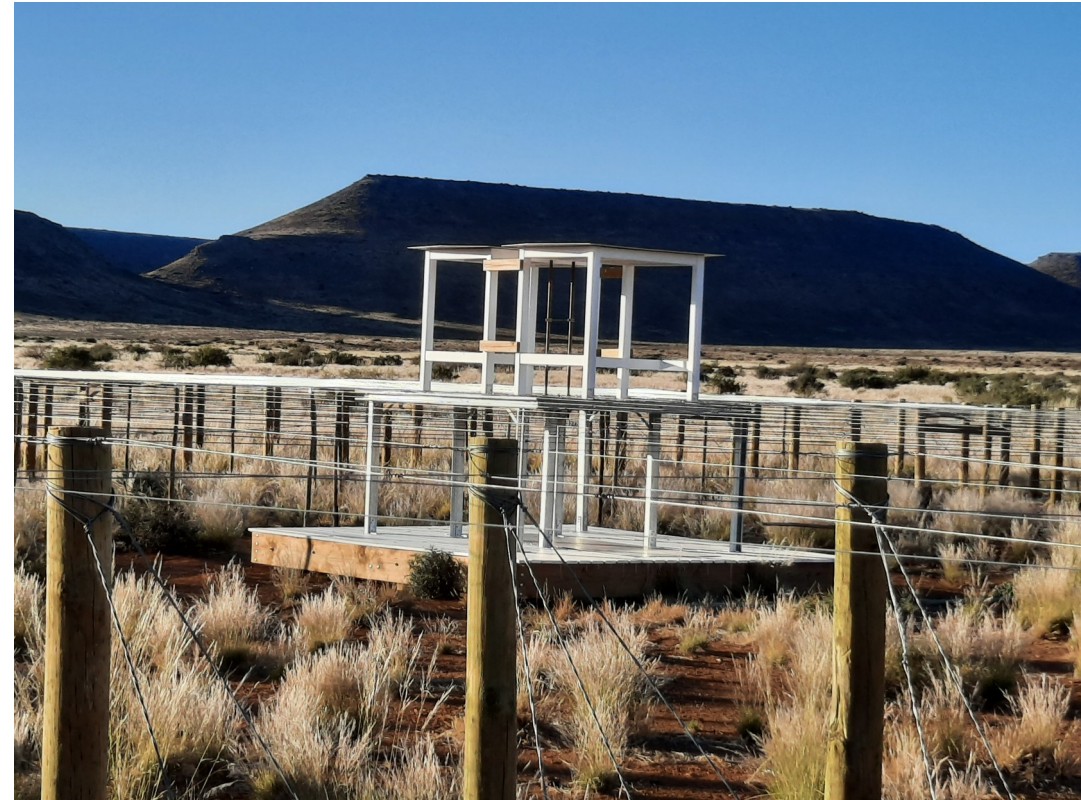


# REACH



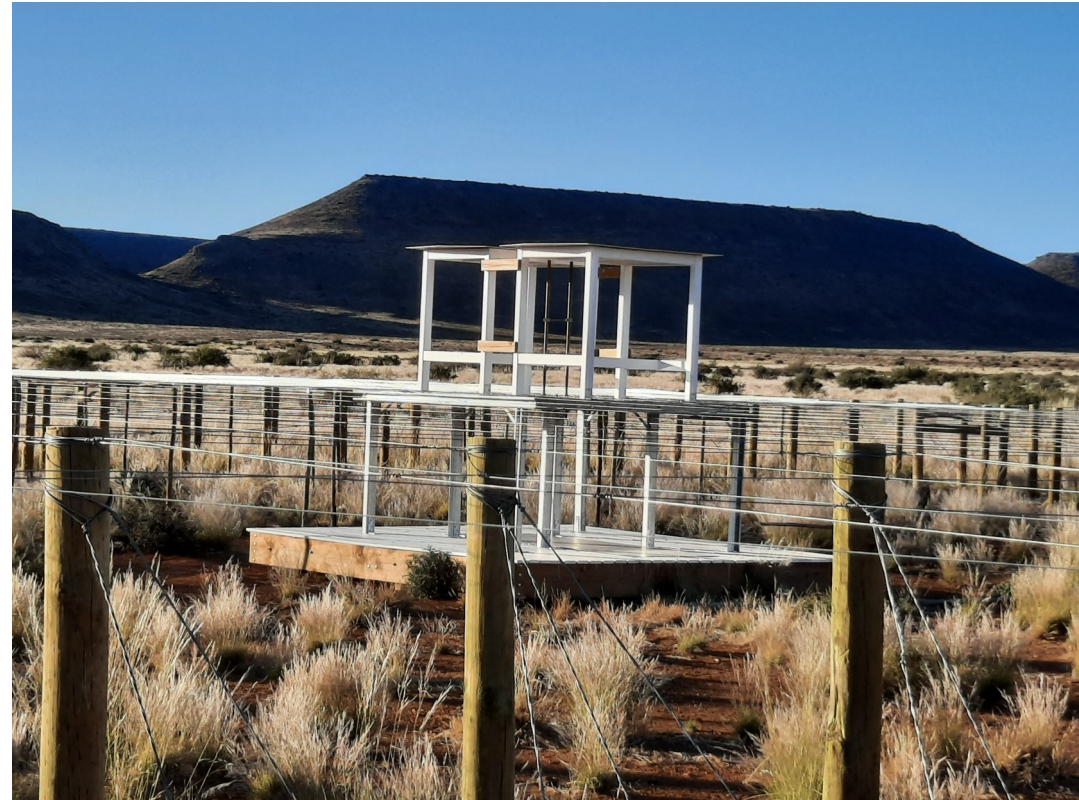
# REACH

- Located in the Karoo radio reserve in South Africa
- Very wide-band
- In-field calibration
- Raised Ground Plane



# REACH

- Designed with Bayesian calibration and data analysis in mind
- Use physics-rooted, physically interpretable models to understand and account for systematics
- Simultaneous observations with two different antennae



# REACH Data Analysis

- Generate a model of the sky radio emission based on physical parameters

$$T_{\text{sky}}(\Omega, \nu, \theta_{\text{F}})$$

- Convolve this with a model of the antenna beam to produce a parameterised foreground model that includes chromatic distortions

$$T_{\text{F}}(\nu, \theta_{\text{F}}) = \frac{1}{4\pi} \int_0^{4\pi} D(\Omega, \nu) T_{\text{sky}}(\Omega, \nu, \theta_{\text{F}}) d\Omega$$

- Use this model in a Nested Sampling fit

$$\log \mathcal{L} = \sum_i \left[ -\frac{1}{2} \log(2\pi\theta_\sigma) - \frac{1}{2} \left( \frac{T_{\text{data}}(\nu_i) - T_{\text{F}}(\nu_i, \theta_{\text{F}}) - T_{\text{S}}(\nu_i, \theta_{\text{S}})}{\theta_\sigma} \right)^2 \right]$$

# Bayesian Data Analysis

- Frequentist probability: The fraction of events that show a given outcome in the limit of a large number of identical repeats

$$P(\mathcal{D}|\mathcal{M}) = \mathcal{L}$$

Likelihood

- Bayesian Probability: The confidence in a given statement

$$P(\mathcal{M}|\mathcal{D}) = \mathcal{P}$$

Posterior

# Bayesian Data Analysis

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad P(\theta_{\mathcal{M}}|\mathcal{D}, \mathcal{M}) = \frac{P(\mathcal{D}|\theta_{\mathcal{M}}, \mathcal{M})P(\theta_{\mathcal{M}}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}$$

$$\mathcal{P} = \frac{\mathcal{L}\pi}{\mathcal{Z}}$$

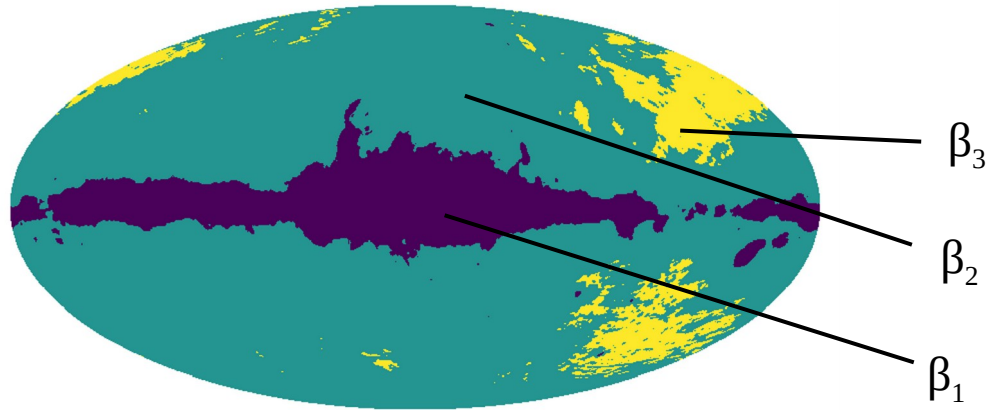
$$P(\mathcal{M}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{M})P(\mathcal{M})}{P(\mathcal{D})} = \mathcal{Z} \frac{P(\mathcal{M})}{P(\mathcal{D})}$$



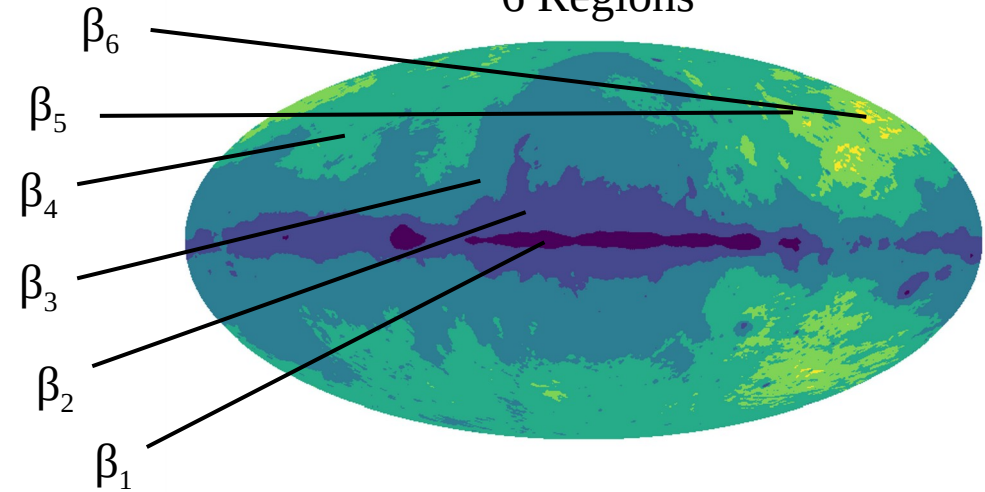
# REACH Data Analysis

Parameterise by spectral index

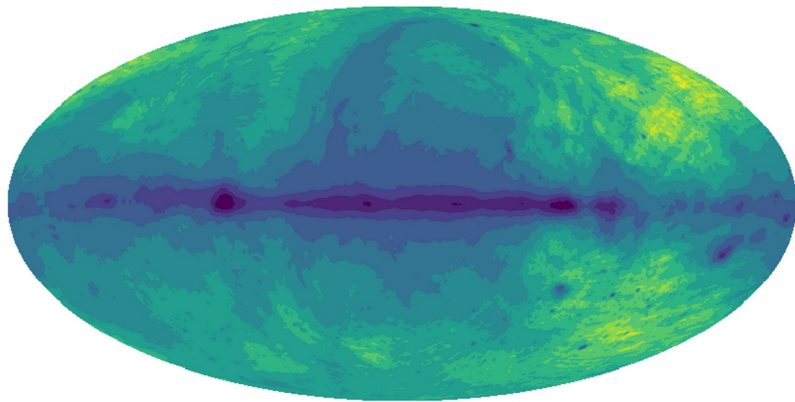
3 Regions



6 Regions

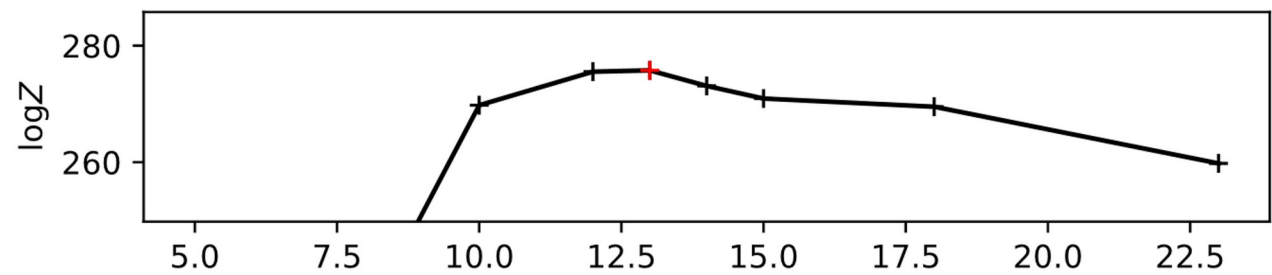
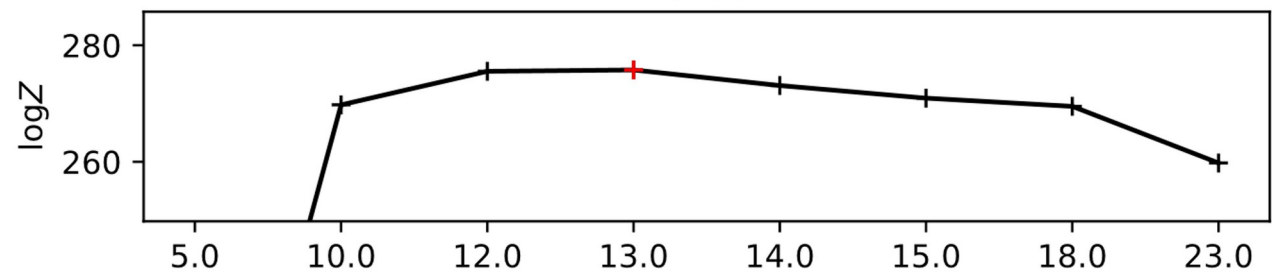
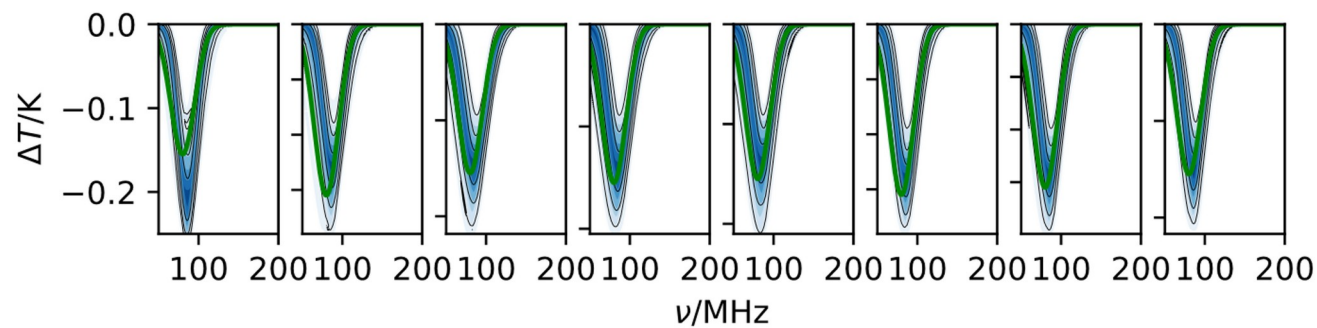
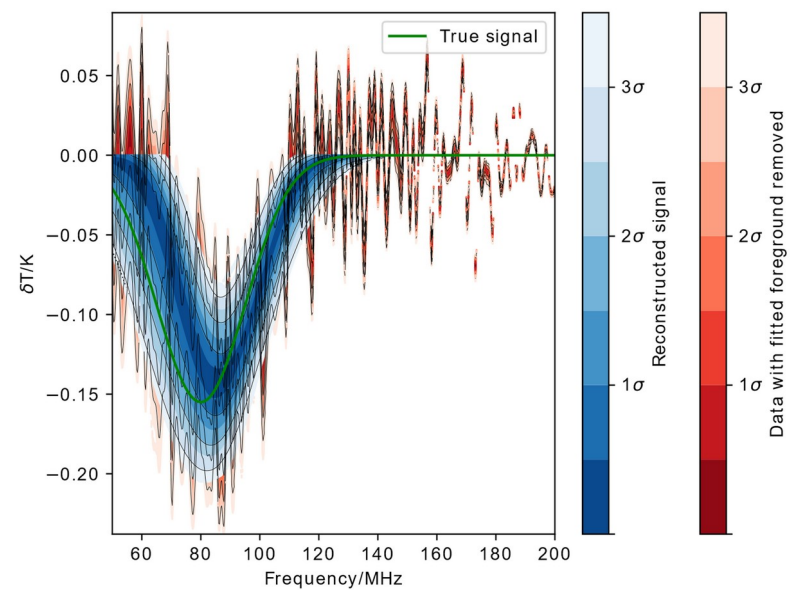
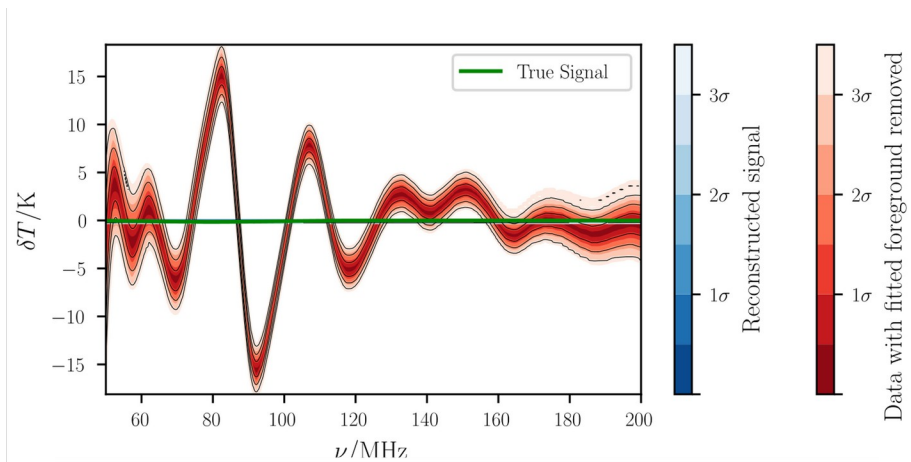


12 Regions



$$T_{\text{sky}}(\Omega, \nu) = \left[ \sum_{n=1}^N M_n(\Omega) (T_{\text{base}}(\Omega) - T_{\text{CMB}}) \left( \frac{\nu}{\nu_{\text{base}}} \right)^{-\beta_n} \right] + T_{\text{CMB}}$$

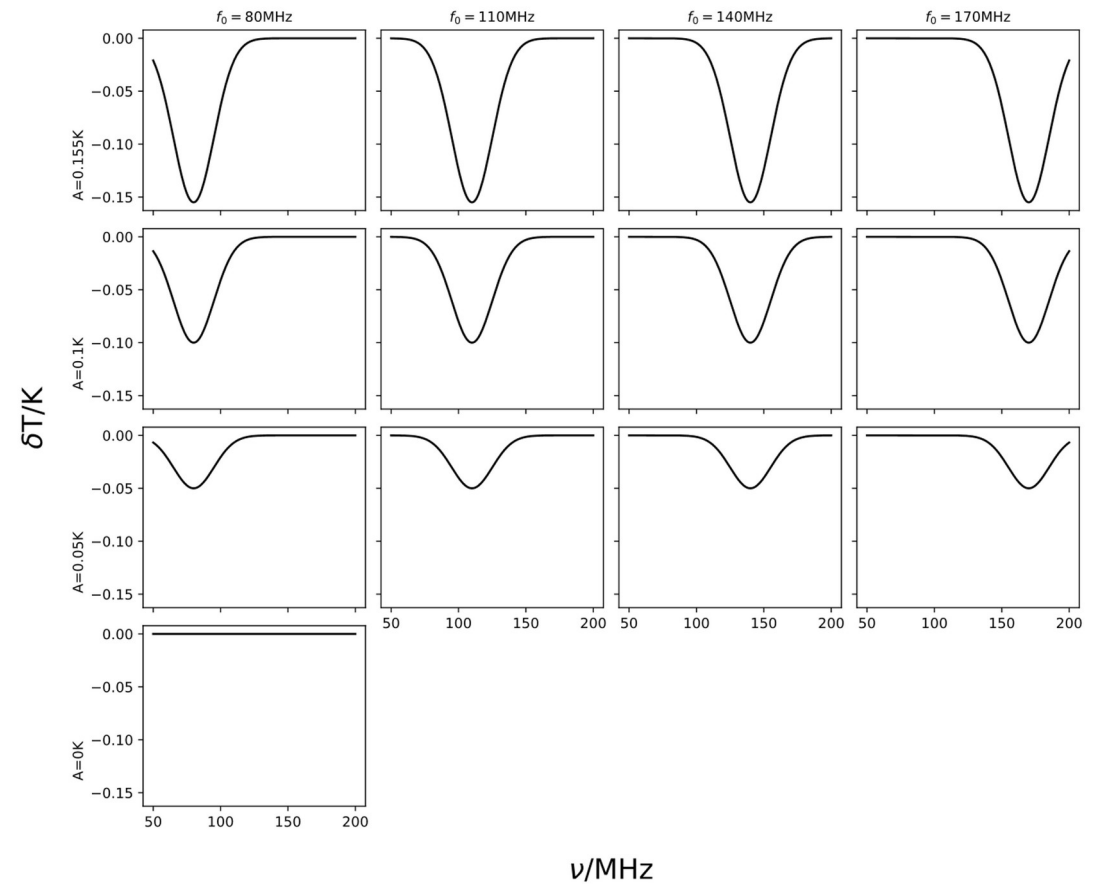
# Performance



# Antenna Comparison

- Simulate observation data for a range of potential antenna designs
- Insert a range of mock 21cm signals into the data
- Attempt to recover each signal using the described pipeline

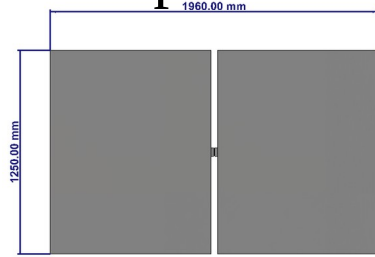
This allows the accuracy of each recovered signal to be quantified relative to the true signal. We can also quantify the statistical confidence with which the signal is detected.



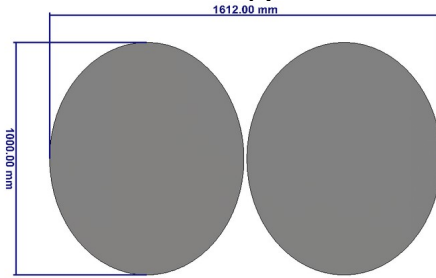
Anstey et al. 2021 (b)

# Antenna Comparison

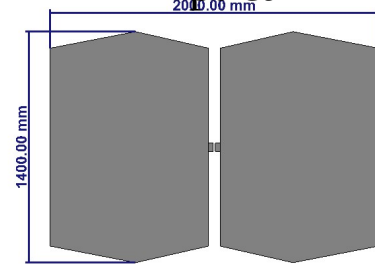
Dipoles:



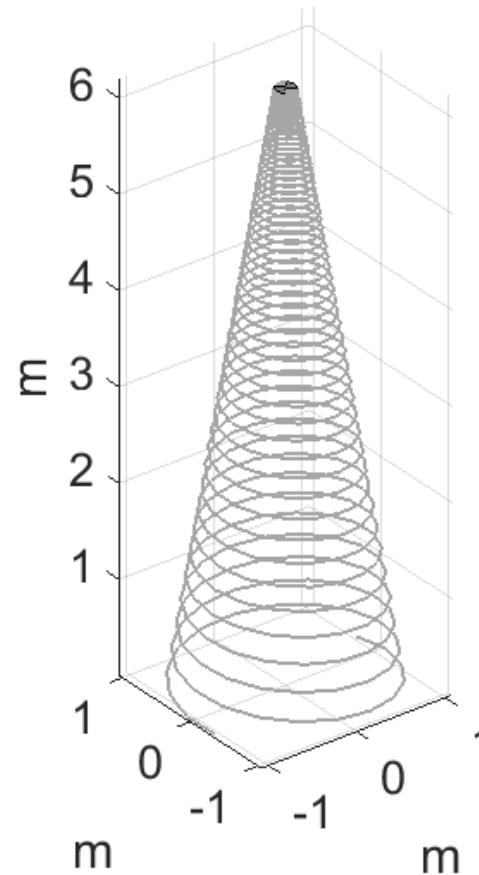
Rectangular



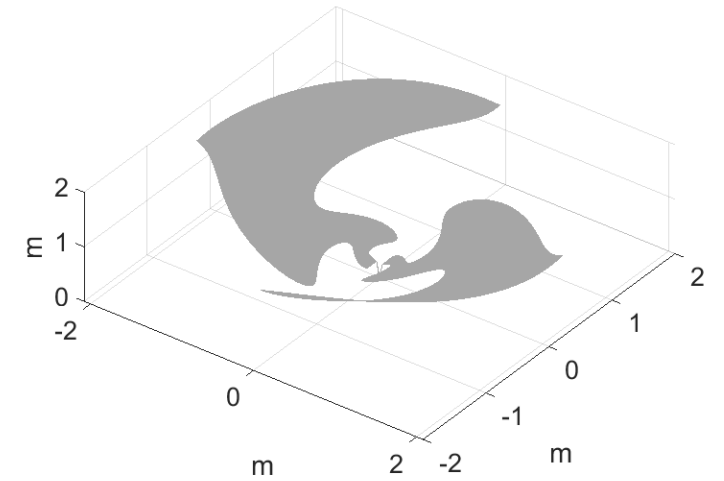
Elliptical



Polygonal



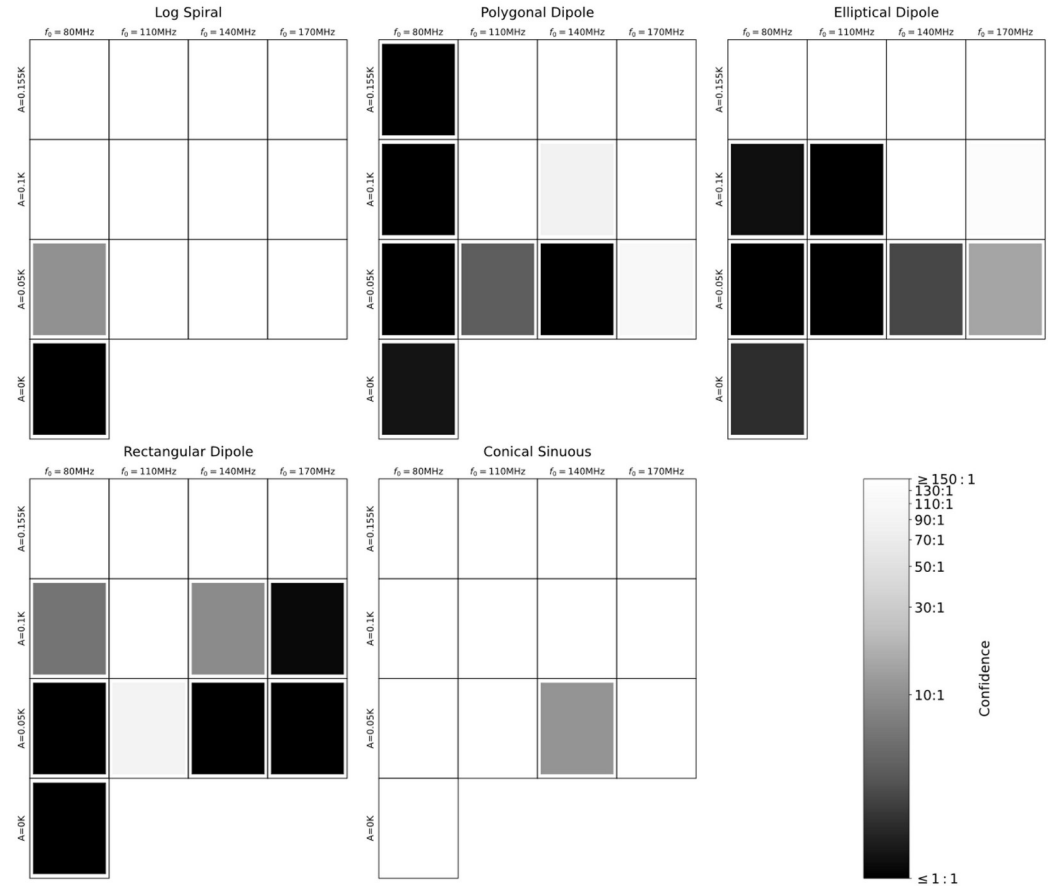
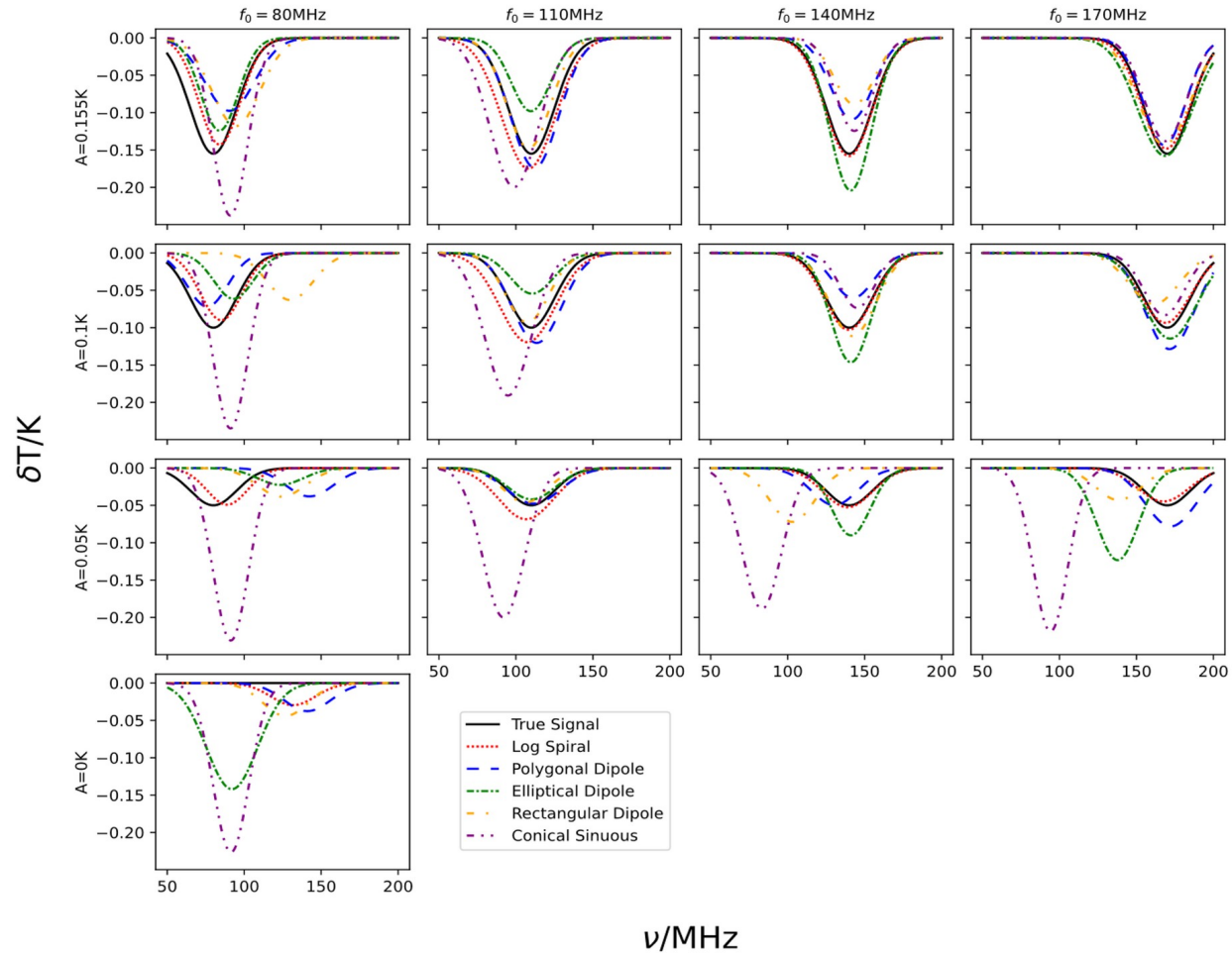
Conical  
Logarithmic Spiral



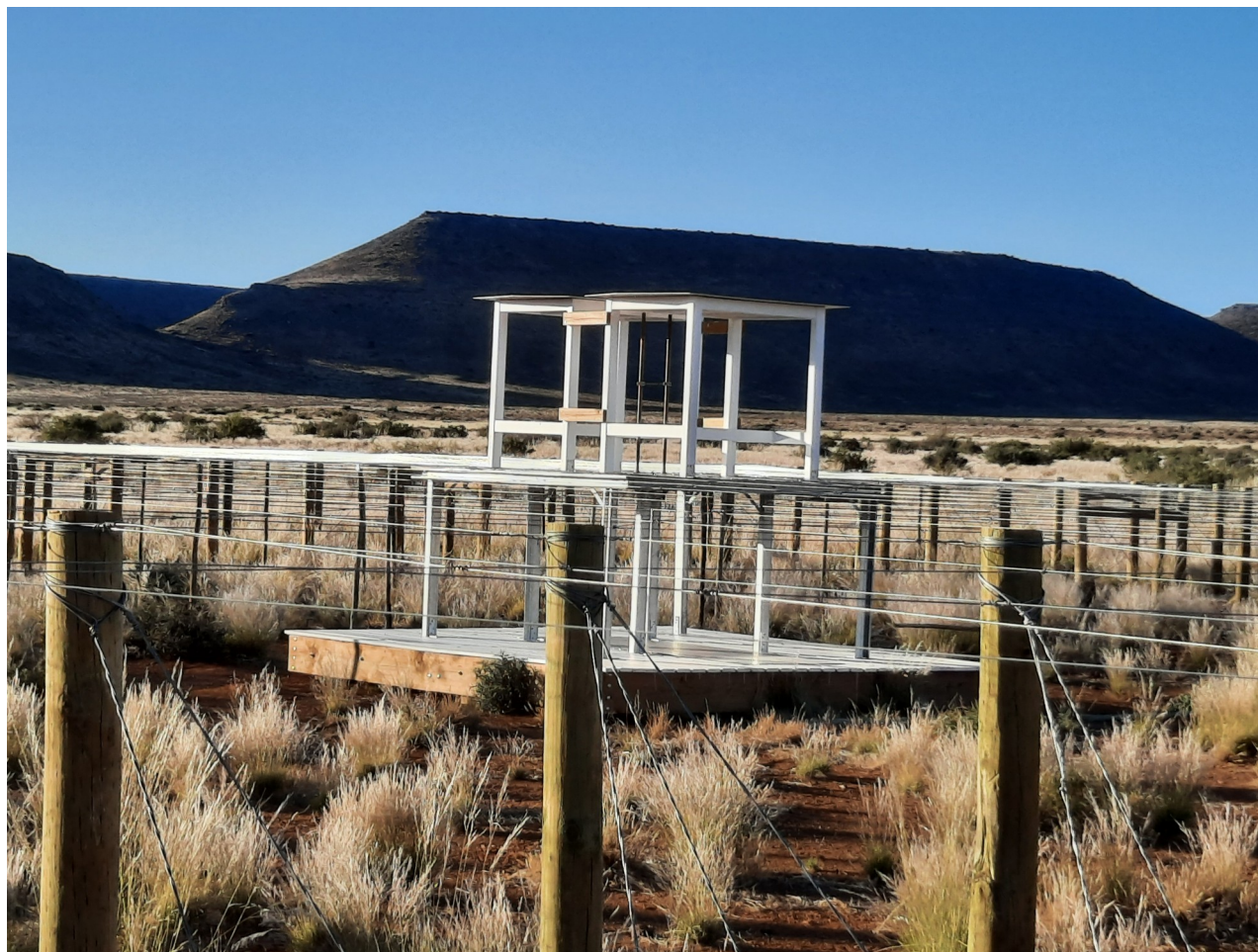
Inverse Conical  
Sinuous

Antenna simulations and images provided by John Cumner and Quentin Guerring

# Antenna Comparison

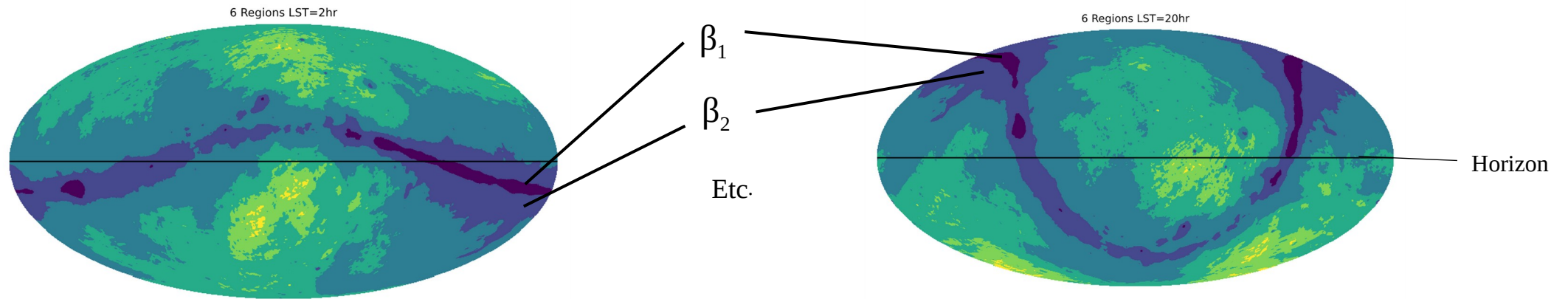


# Antenna Comparison



# Using Multiple Data Sets

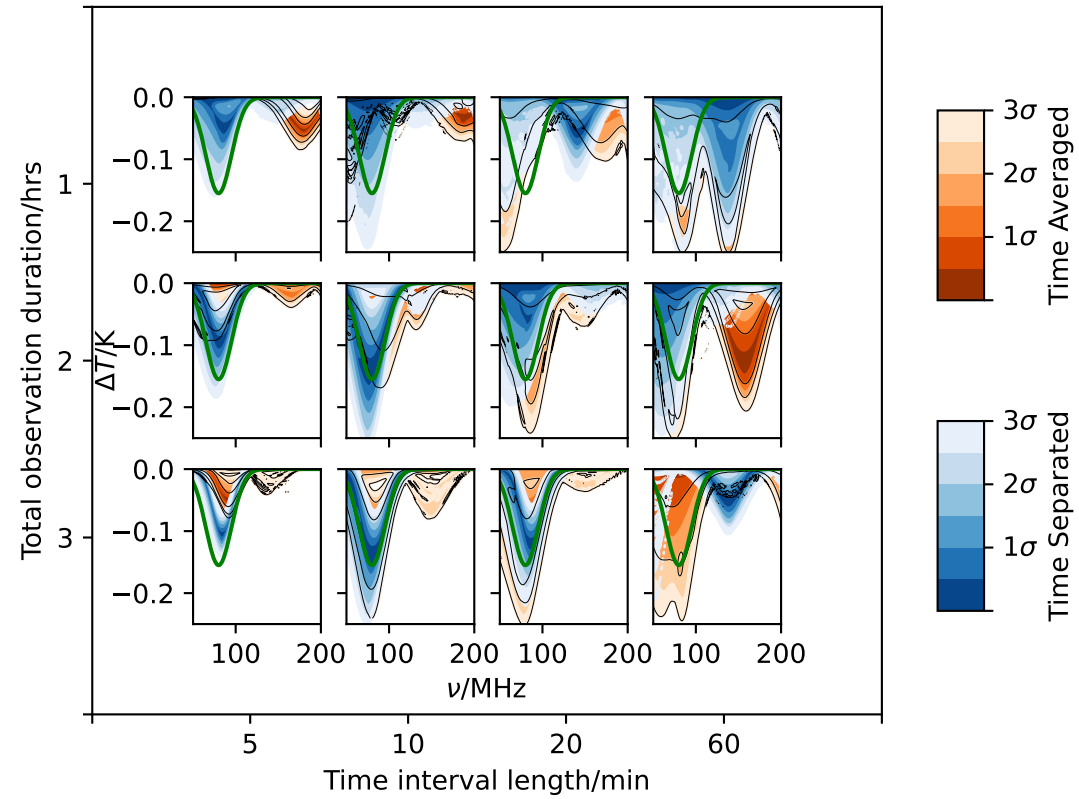
$$\log \mathcal{L} = \sum_i -\frac{1}{2} \log (2\pi\sigma_n^2) - \frac{1}{2} \left( \frac{\frac{1}{n_j} \sum_j [T_{\text{data}}(\nu_i, t_j)] - \left( \frac{1}{n_j} \sum_j [T_{\text{F}}(\nu_i, t_j, \theta_{\text{F}})] + T_{\text{S}}(\nu_i, \theta_{\text{S}}) \right)}{\sigma_n} \right)^2$$



$$\log \mathcal{L} = \sum_i \sum_j \left[ -\frac{1}{2} \log (2\pi\theta_\sigma) - \frac{1}{2} \left( \frac{T_{\text{data}}(\nu_i, t_j) - T_{\text{F}}(\nu_i, t_j, \theta_{\text{F}}) - T_{\text{S}}(\nu_i, \theta_{\text{S}})}{\theta_\sigma} \right)^2 \right]$$

# Results

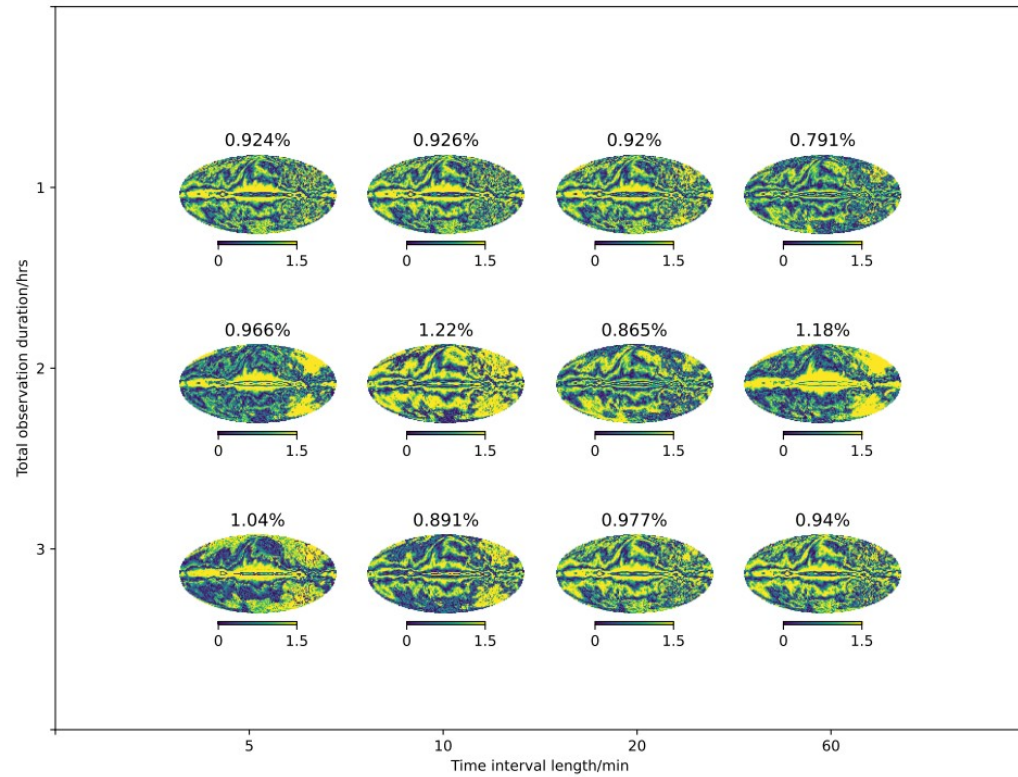
## Polygonal Dipole



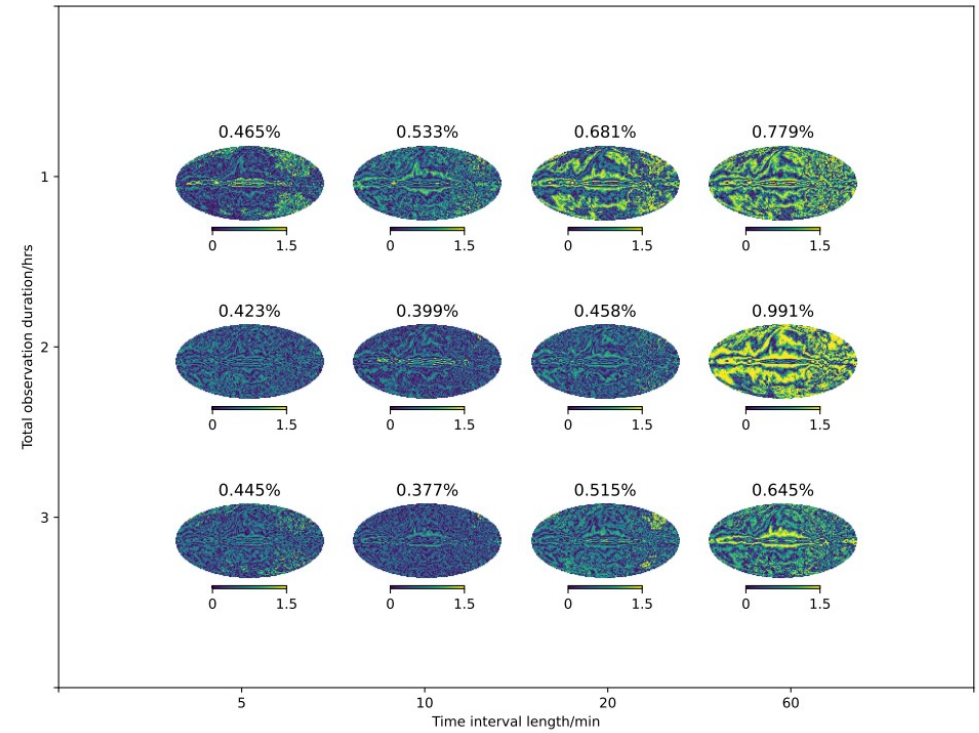


# Results

## Time-averaged

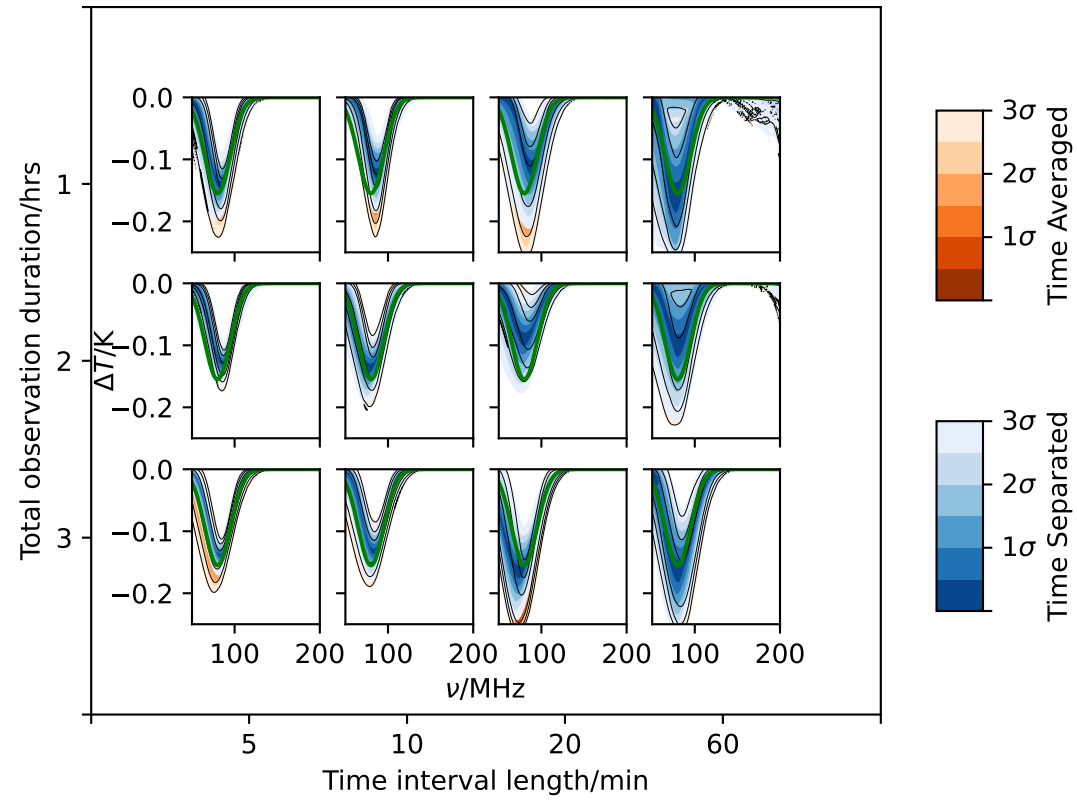


## Time-separated



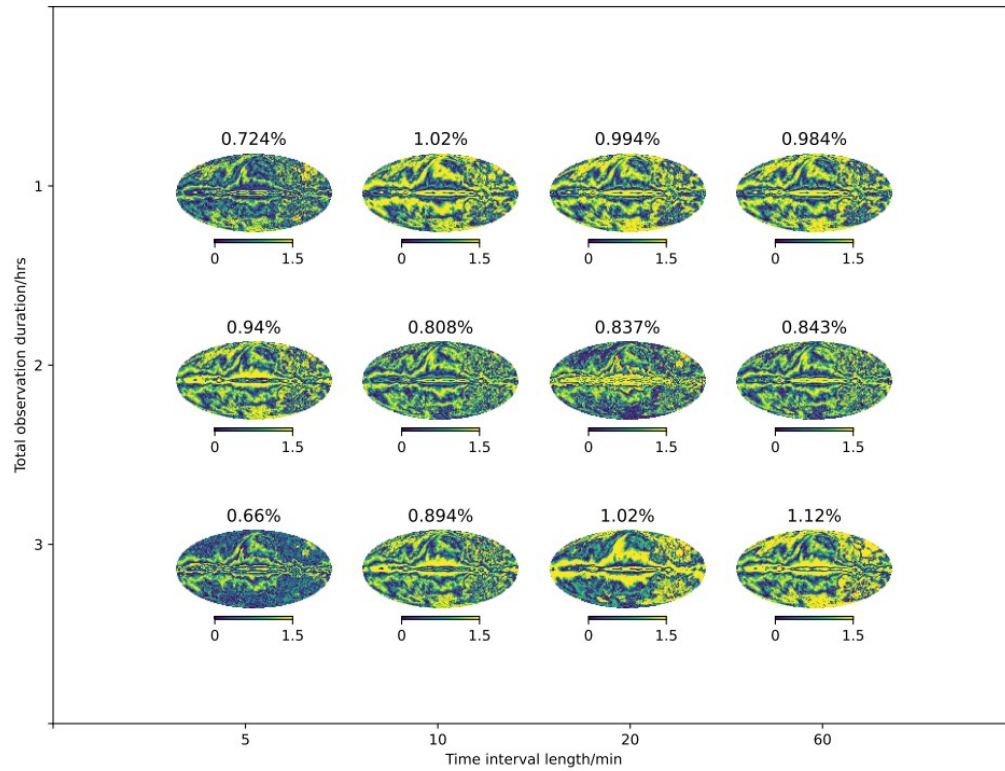
# Results

Log Spiral

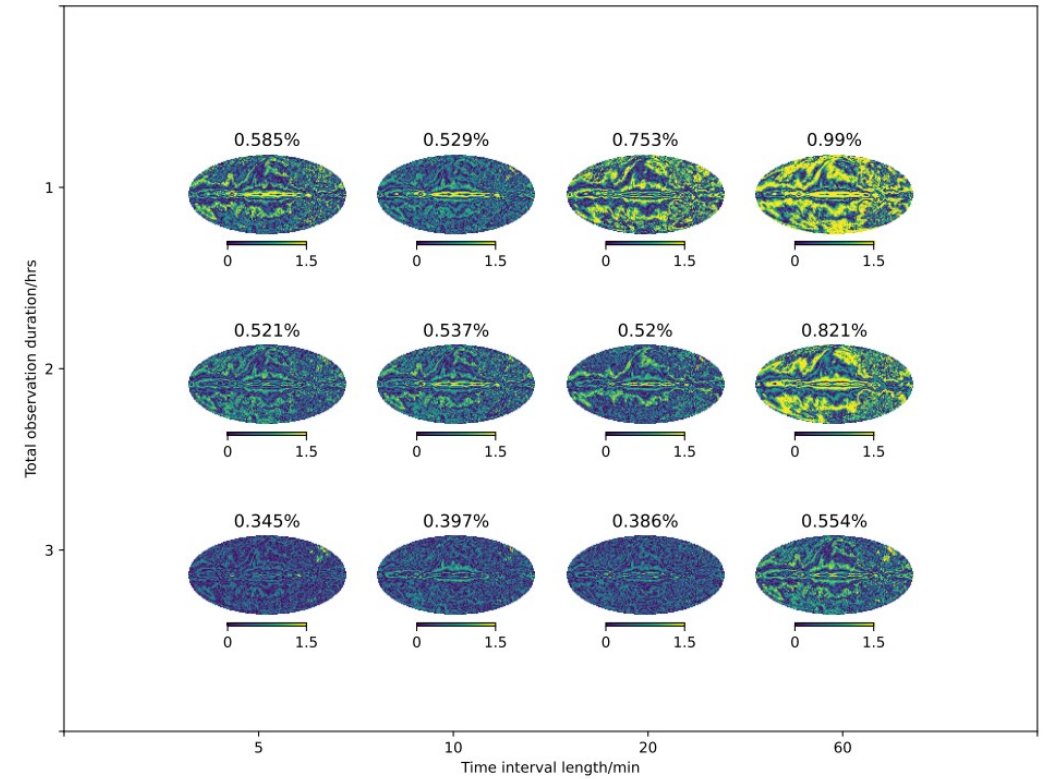


# Results

## Time-averaged

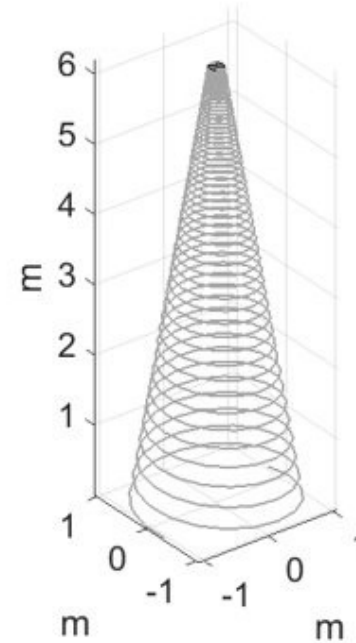
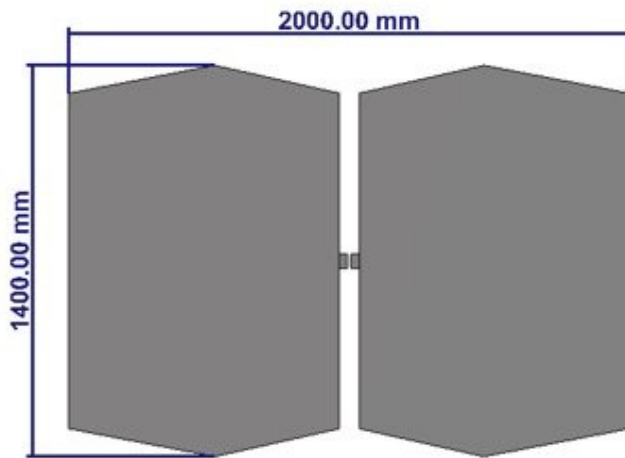


## Time-separated

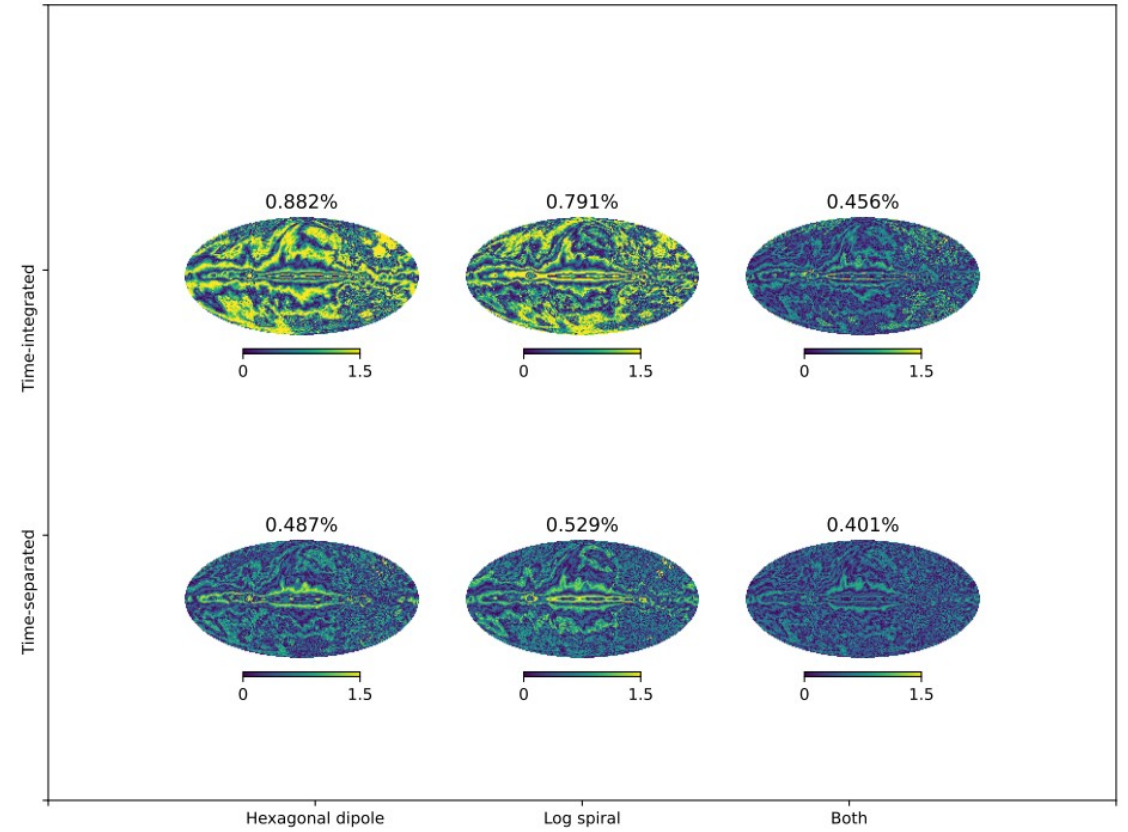
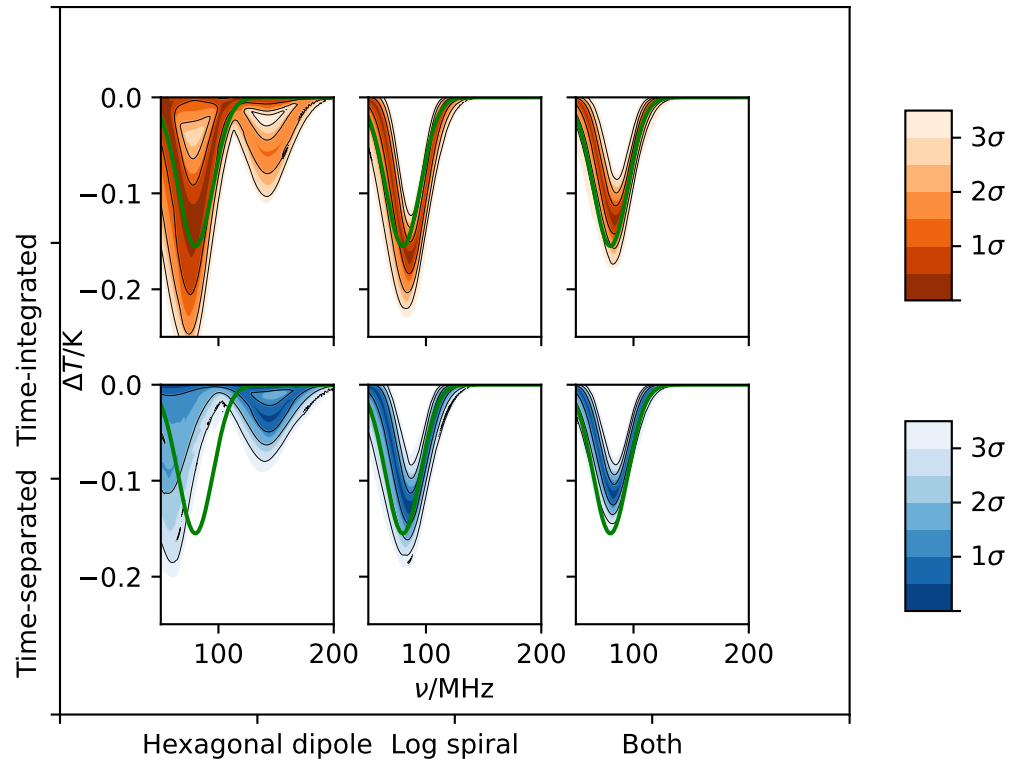


# Using Multiple Antennae Sets

$$\log \mathcal{L} = \sum_i \sum_j \sum_k -\frac{1}{2} \log (2\pi\sigma_n^2) - \frac{1}{2} \left( \frac{T_{\text{data}k}(\nu_i, t_j) - (T_{\text{F}k}(\nu_i, t_j, \theta_{\text{F}}) + T_{\text{S}}(\nu_i, \theta_{\text{S}}))}{\sigma_n} \right)^2$$



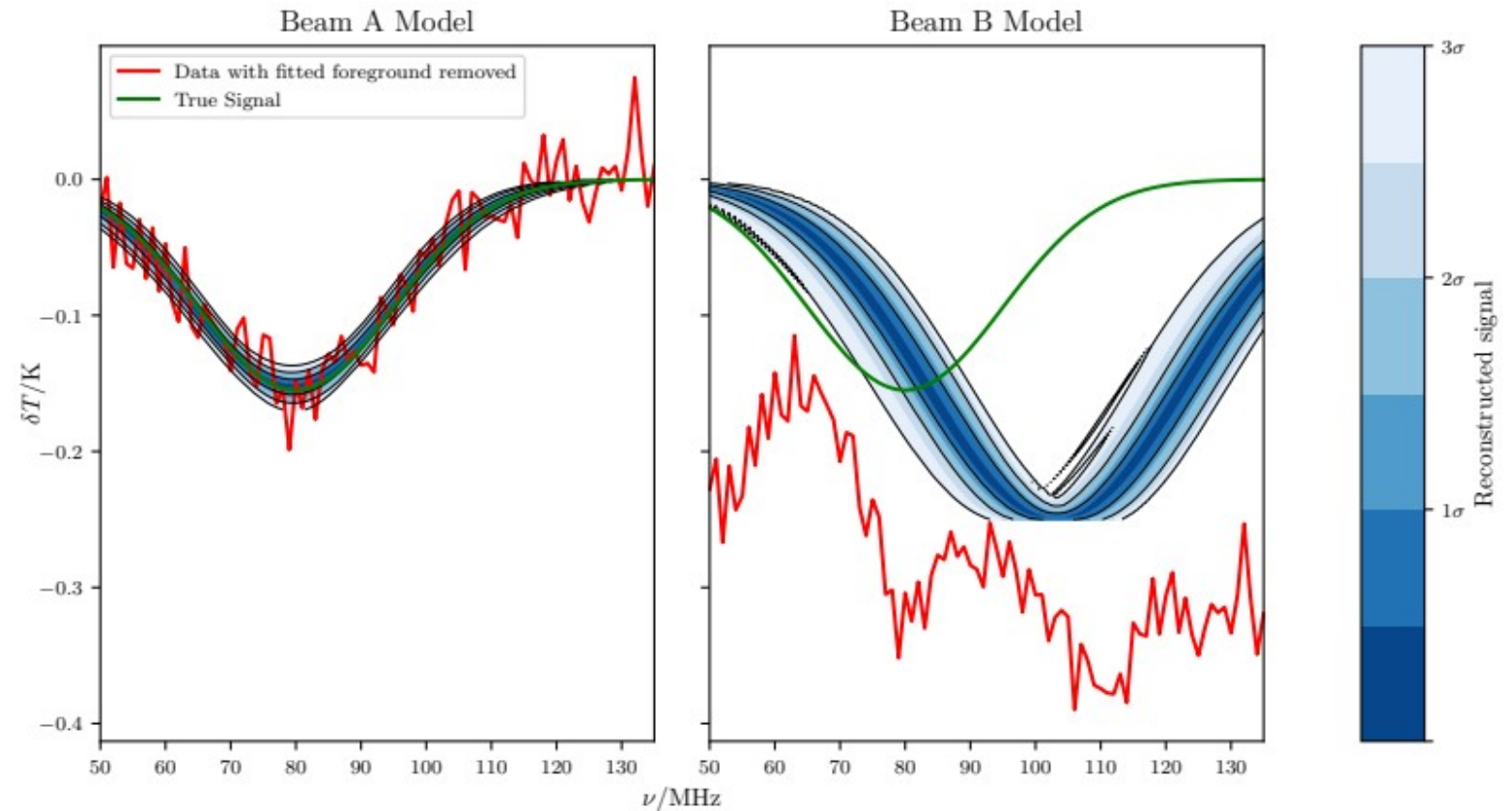
# Results



# Beam Errors

The antenna beam pattern may not be known exactly in practice:

- Soil permittivity
- Horizon effects
- Imperfections in construction
- Etc.



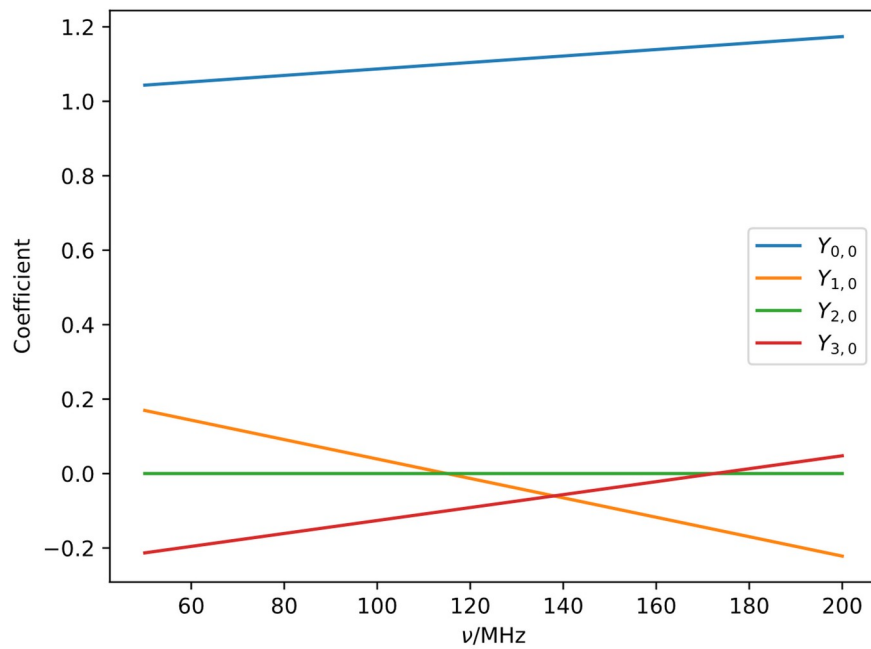
# Modelling Beam

$$D(\Omega, \nu, \bar{\theta}) = \sum_{j=1}^M \theta_j X_j(\Omega, \nu)$$

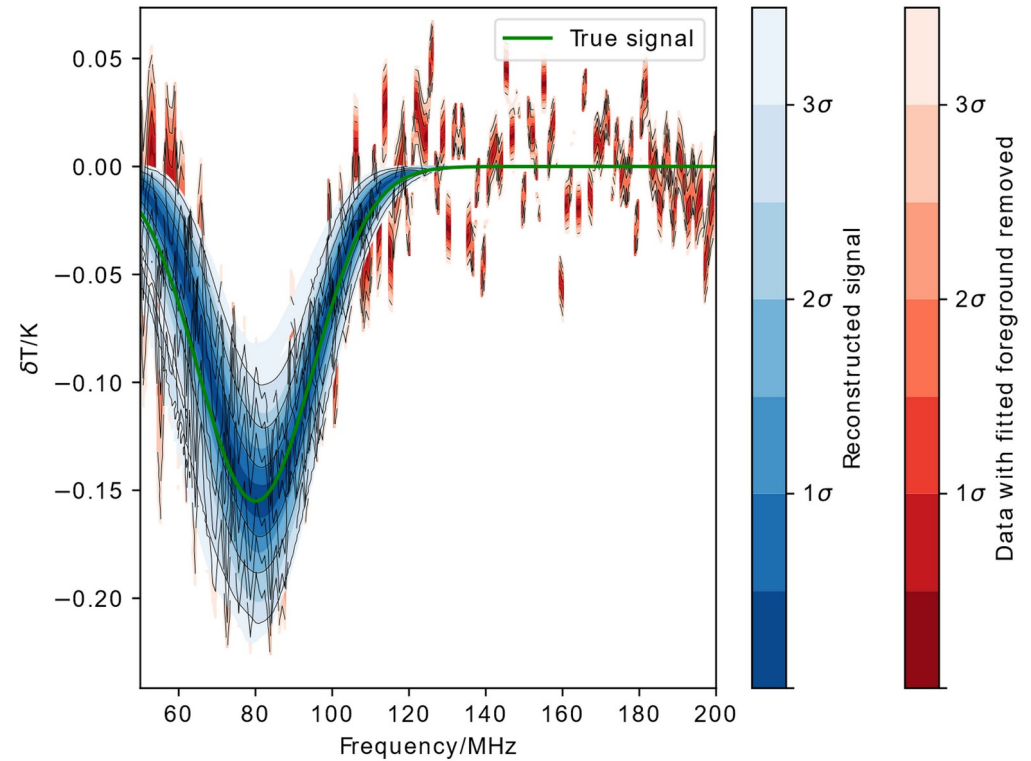
$$D(\Omega, \nu, \bar{\theta}) = \sum_{j=1}^M \Gamma(\nu, \bar{\theta}_j) Y_j(\Omega)$$

# Modelling Beam

## Model Beam Coefficients



## Fit Results





# Conclusions

- 21 cm Cosmology is one of the most promising methods of studying the dark ages, cosmic dawn and the epoch of reionisation
  - There is currently no confirmed detection of the 21 cm signal
  - Producing a detection has numerous challenges
- 
- REACH
  - Core REACH data analysis pipeline
  - Uses and extensions to this pipeline