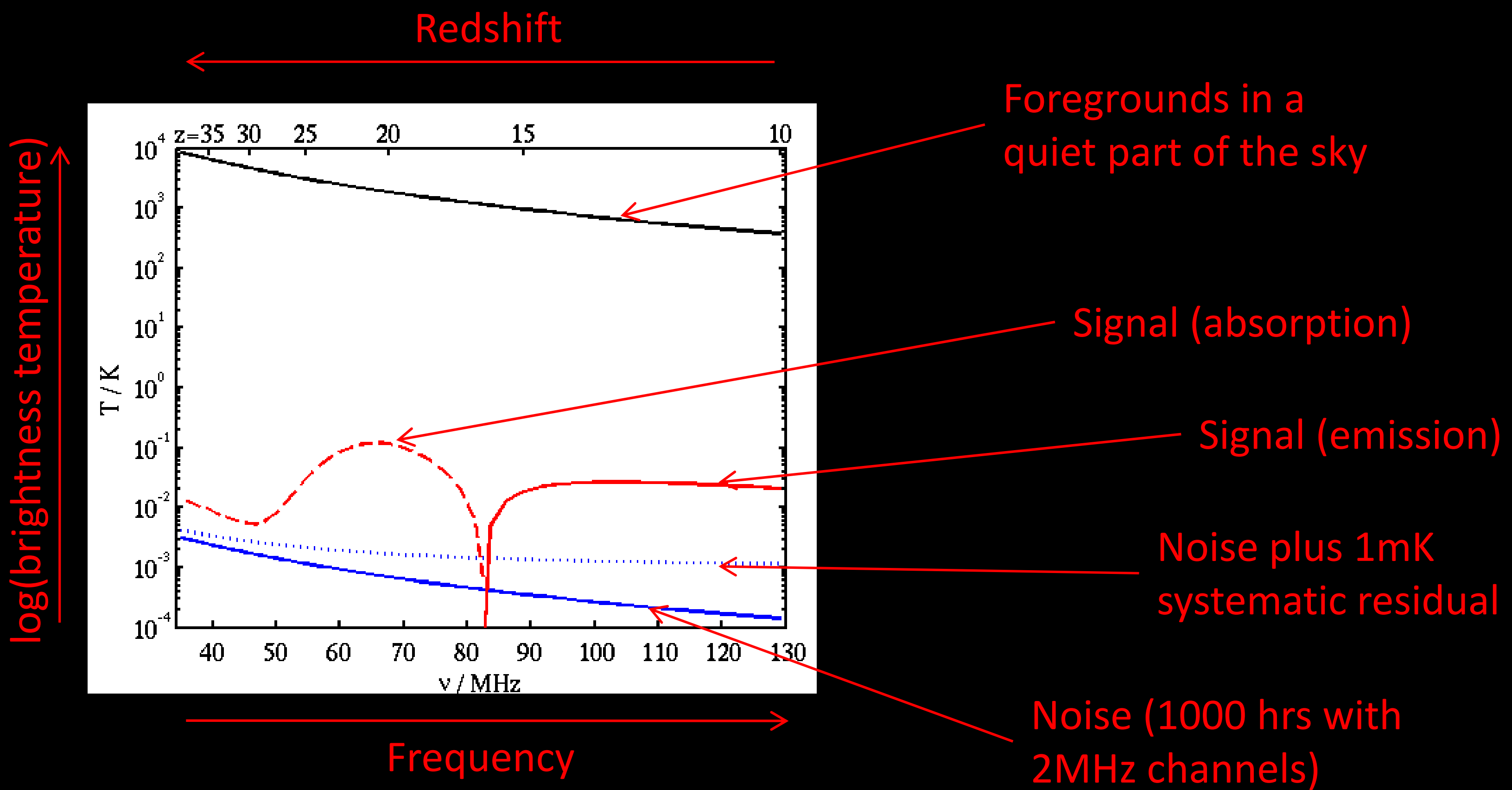
Modelling the foregrounds and the system response for DARE

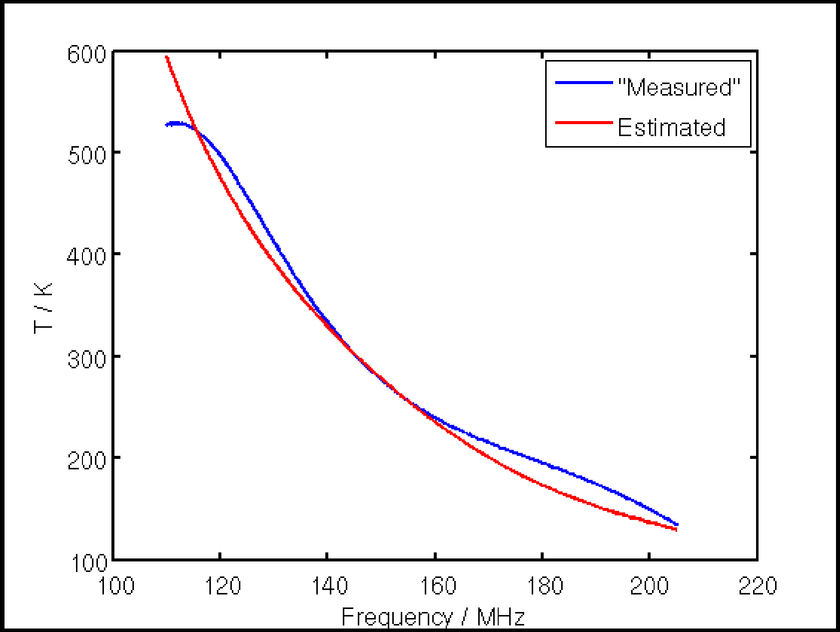
Geraint Harker
University of Colorado

Recap of the problem: foregrounds



Recap of the problem: instrumental response

EDGES data



- Measured spectrum is multiplied by an instrumental response.
  - This is not known sufficiently accurately in advance.
  - Must be estimated from data, so it must be sufficiently smooth (designed to be so for DARE).
  - Removes degrees of freedom from the extracted signal.
- Spurious additive features and non-Gaussian noise may also be introduced by the instrument, but are minimized by DARE's strategy and design.

Interaction of foregrounds and system response

- The basic strategy is to use the smoothness of the foregrounds, compared to the signal which has spectral features.
  - If the foregrounds really are smooth, their size relative to the foregrounds does not present an insurmountable problem.
  - A problem arises if the instrumental response compromises this smoothness.
- Similarly, if there were no foregrounds, an instrumental response modelled to within 1% would yield a signal accurate to 1%.
  - A 1% error combined with foregrounds 10^5 time larger than the signal yields errors 1000 times larger than the signal.
- The *combination* of large foregrounds with errors is what causes the problem.

Contributions to the foregrounds

- Spectrally smooth, diffuse foregrounds (around 2000K at 65 MHz).
  - Galactic synchrotron (~ 72%)
  - Galactic free-free (~1%)
  - A sea of unresolved extragalactic sources with synchrotron spectra (~27%)
- Radio recombination lines
  - Narrow, occur at known frequencies
  - Require sufficiently good spectral resolution (around 10 kHz) to detect and remove without discarding too much data.
- Local compact sources are diluted over the area of the beam
  - The Sun (tens of Kelvin except during bursts, when we can't observe)
  - Jupiter (a few millikelvin; bursts only occur at frequencies below 40 MHz)
- The Moon
  - Used as a secondary calibrator
  - Emits thermal radiation and reflects other foregrounds
  - Contribution is modulated by changing the orientation of the spacecraft.

System response (EDGES case)

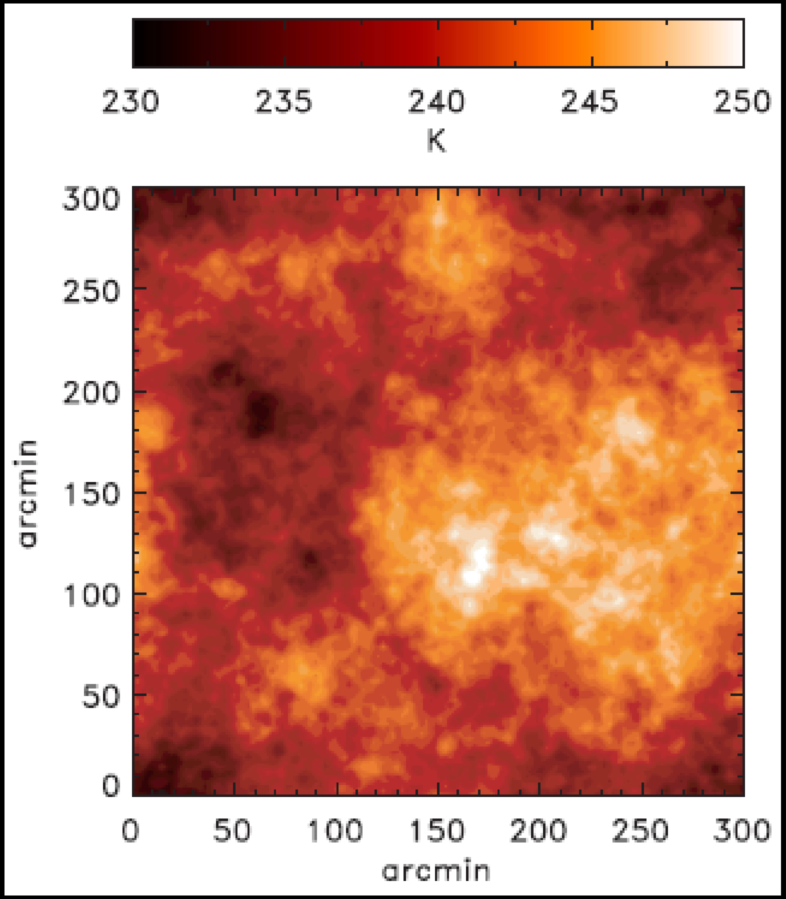
$$T_{\mathrm{ant}}(\nu) = \left[1 - |\Gamma(\nu)|^2\right] T_{\mathrm{sky}}(\nu) + \left[2\varepsilon|\Gamma|\cos(\beta) + \varepsilon^2|\Gamma|^2\cos^2(\beta) + (1-\varepsilon)^2|\Gamma|^2\right] T_{\mathrm{rcv}}(\nu) + \ldots$$

Rogers & Bowman (2010)

- T_{ant} : antenna temperature, calibrated by switching between loads
- T_{sky} : sky temperature
- T_{rcv} : temperature of receiver noise propagated back towards the antenna
- $|\Gamma|^2$: power reflection coefficient between antenna and receiver
- β : phase shift due to electrical path length
- ε : voltage correlation coefficient

Different foreground models: stochastic physical model

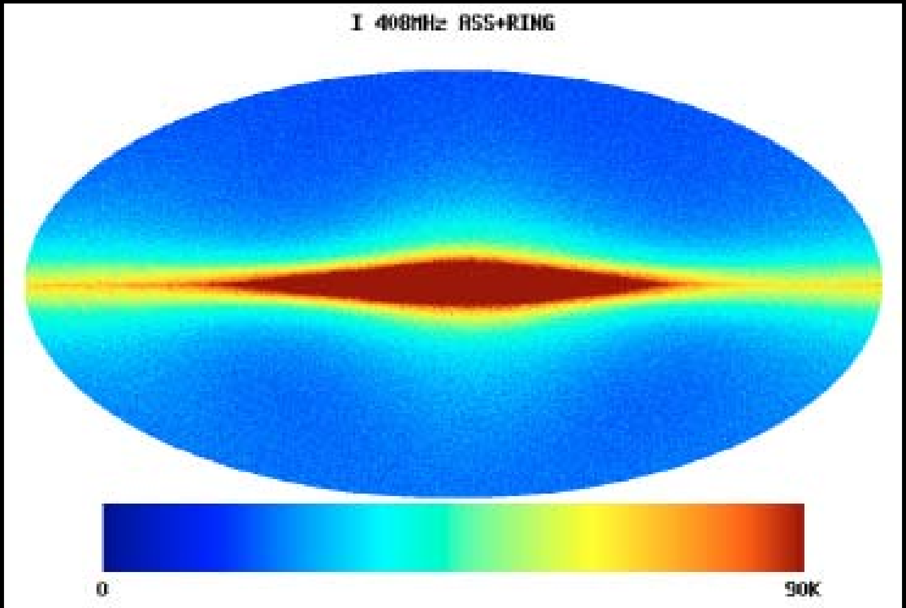
- Use a physical model for synchrotron, free-free, radio clusters, etc.
- Stochastic model using observationally motivated prescriptions for the power spectrum, frequency dependence, etc.
- More suitable for modelling small areas of the sky, e.g. for upcoming LOFAR observations.



Jelić et al. (2008)
120 MHz

Different foreground models: all-sky physical models

- Make a model of emission from the whole sky using 3D models of the Galaxy.
- Can use e.g. polarization data and information from other wavebands to constrain the 3D model.
- Doesn't include the more complicated or hard-to-model aspects of the sky.

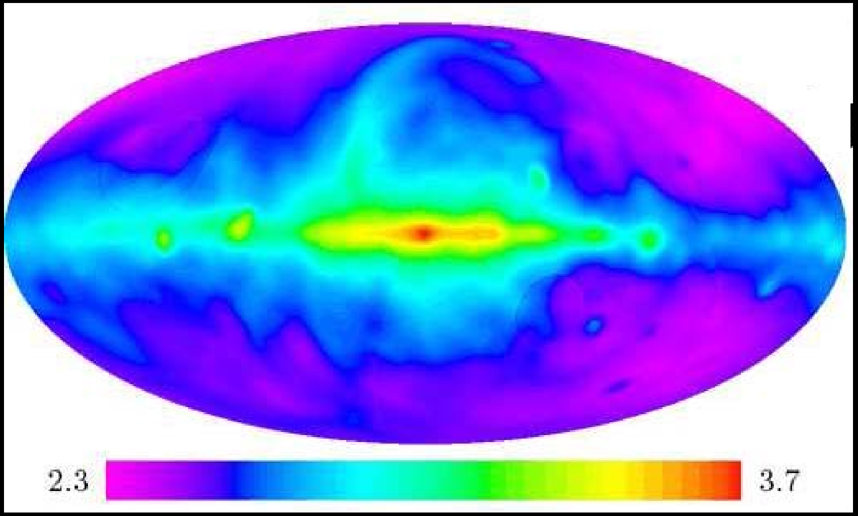


Sun at al. (2008)

Different foreground models: direct reconstruction from data

- Principal component analysis of a variety of low-frequency large-area surveys.
- Includes all relevant diffuse foregrounds by construction.
- Inhomogeneous coverage.
- Even here, some modelling is involved to produce the maps.

de Oliveira-Costa et al. (2008)
150 MHz



$\log_{10}$ (T / K)

Modelling the models

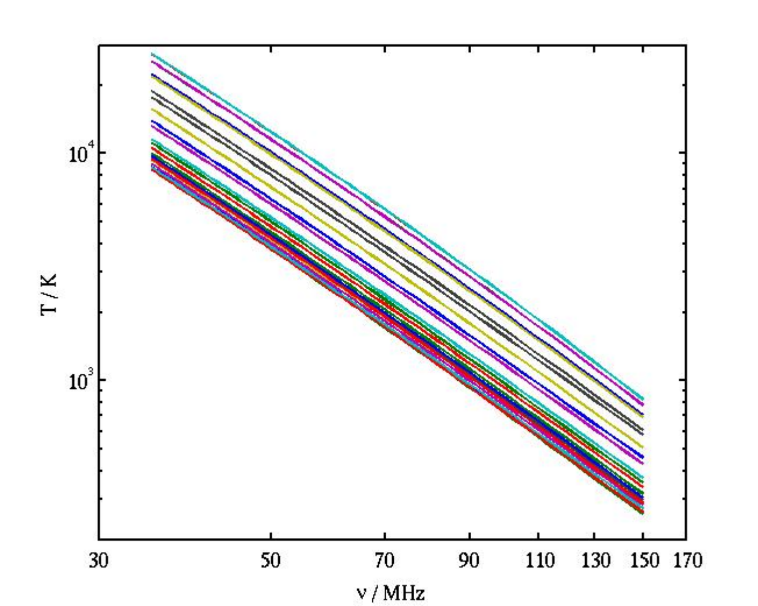
- Use the all-sky model as a test for our fitting routines.
- We fit a smooth function of frequency to simulated observations incorporating the foreground models, and subtract this off to leave noise + signal.
- At present, fitting the following form:

$$\log T_{\rm sky} = \log T_0 + a_1 \log(\nu/\nu_0) + a_2[\log(\nu/\nu_0)]^2 + a_3[\log(\nu/\nu_0)]^3$$

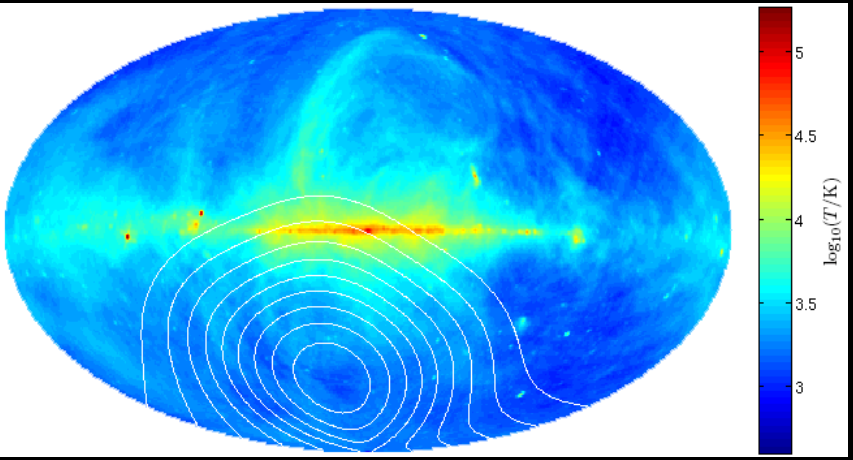
Pritchard & Loeb (2010)

  – Power law, with a running of the spectral index, and a running of the running

Foregrounds as a function of position



•log-log plot of spectra from 24 different sky areas
•Overall normalization changes by a factor of a few.



•de Oliveira-Costa sky model at 70 MHz with contours of a dipole beam overlaid.
•Contour spacing of 10%
•Of order 10 independent beam areas on the sky.

Using spatial variation of the foregrounds

- Averaged over a DARE beam, the 21-cm signal from any part of the sky is the same, while the foreground vary. This should be exploited in removing foregrounds! Several methods have been proposed.
- Independent component analysis
  - Extracts statistically independent components of the spectrum.
  - Non-parametric, but loses scaling information unless more assumptions are made.
- Matched filtering
  - Use information from different sky areas to construct a correlation matrix for the foreground spectrum.
  - Use this, plus a signal model, to construct a matched filter which optimally extracts a signal conforming to that model from the integrated spectrum.
- Maximum likelihood analysis (motivated in part by COBE FIRAS analysis)
  - Parametrize all the relevant components of the spectrum for all the sky areas under consideration.
  - Perform a search of a high-dimensional space to find the best combination of parameters.

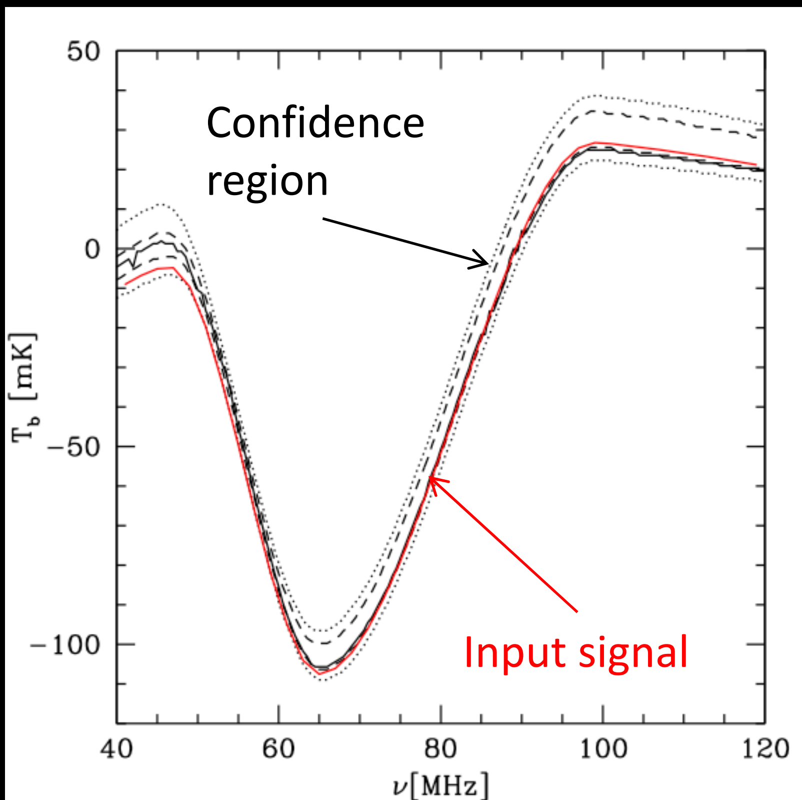
Using spatial variation of the foregrounds

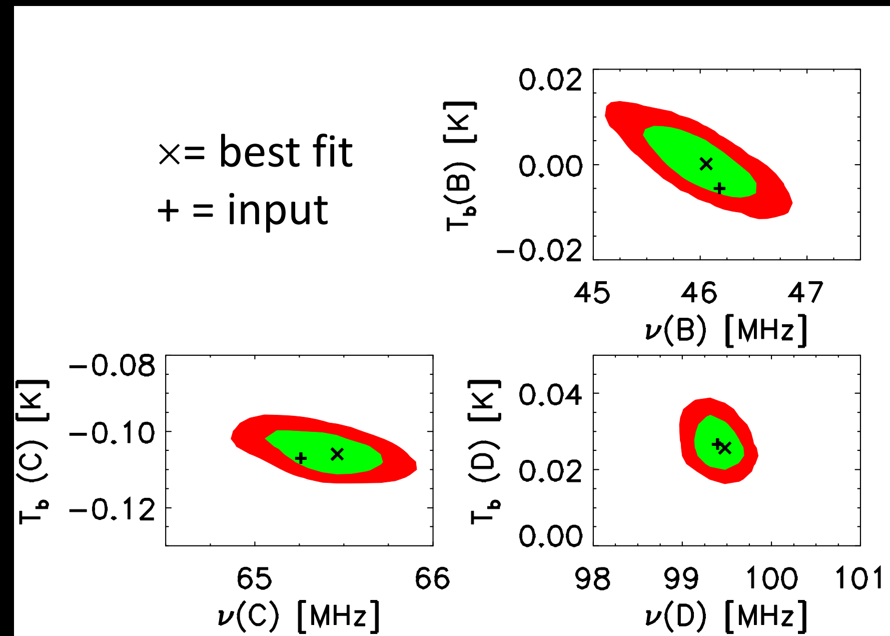
- Averaged over a DARE beam, the 21-cm signal from any part of the sky is the same, while the foreground vary. This should be exploited in removing foregrounds! Several methods have been proposed.
- Independent component analysis
  - Extracts statistically independent components of the spectrum.
  - Non-parametric, but loses scaling information unless more assumptions are made.
- Matched filtering
  - Use information from different sky areas to construct a correlation matrix for the foreground spectrum.
  - Use this, plus a signal model, to construct a matched filter which optimally extracts a signal conforming to that model from the integrated spectrum.
- Maximum likelihood analysis (motivated in part by COBE FIRAS analysis)
  - Parametrize all the relevant components of the spectrum for all the sky areas under consideration.
  - Perform a search of a high-dimensional space to find the best combination of parameters.

Parameters to be fit

Description of contribution	Number of parameters
Frequency and amplitude of turning points of 21-cm signal	$2\times3 = 6$
Foreground spectrum in *n* different regions of the sky	4n
Instrumental frequency response (fit with a low-order polynomial)	~6
Spectrum of the quiet Sun and Jupiter (relatively weak)	~8
Spectrum and reflectivity of the Moon	?
Total	**>20 + 4n**

Preliminary results





- Confidence regions for the parameters of the turning points from a Markov Chain Monte Carlo analysis
- 1000 hrs of total integration
- Parameters of the turning points and the diffuse foregrounds are included in the analysis

Future work

- Include the rest of the parameters shown in the table into the analysis
  - Need to work out how to parametrize some parts!
  - Speed up the MCMC code to deal with such a high-dimensional minimization
- Revisit other approaches
  - Independent component analysis
  - Matched filtering
- Fit foregrounds directly from the sky model, rather than parametrizing them before inputting them.

Summary

- We choose to use a model for the foregrounds constructed from a principal components analysis of low-frequency data.
- Instrument modelling is motivated by that done for EDGES, along with models for the DARE antenna.
- By using spatial variation of the foregrounds, we end up with tens of parameters to fit: our fiducial method uses a Markov Chain Monte Carlo technique to map the likelihood surface.
- Good constraints are achieved on the signal for a mission with the projected lifetime of DARE, but much work still needs to be done on modelling all the relevant physical effects.