Foreground Removal

Geraint Harker
University of Colorado at Boulder

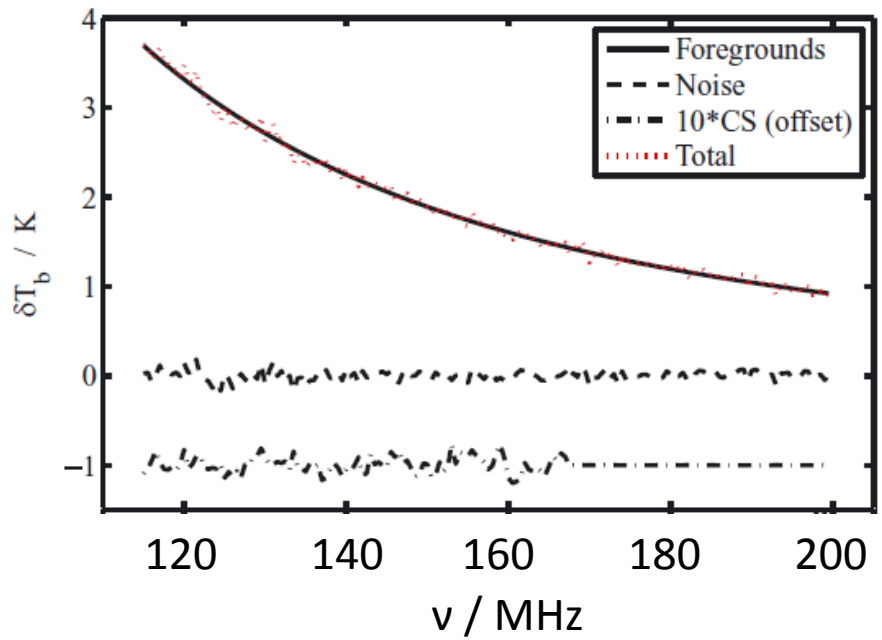
Outline

- Introduction to foreground fitting for 21-cm experiments.
- Different approaches:
  - Parametric vs. non-parametric fitting;
  - (u,v,ν) space vs. (θ_x,θ_y,ν) space.
- How should different approaches be evaluated?
- Results and interpretations from foreground fitting simulations for LOFAR.
  - Effects on plane-of-sky and line-of sight modes.

Diffuse, unpolarized foregrounds

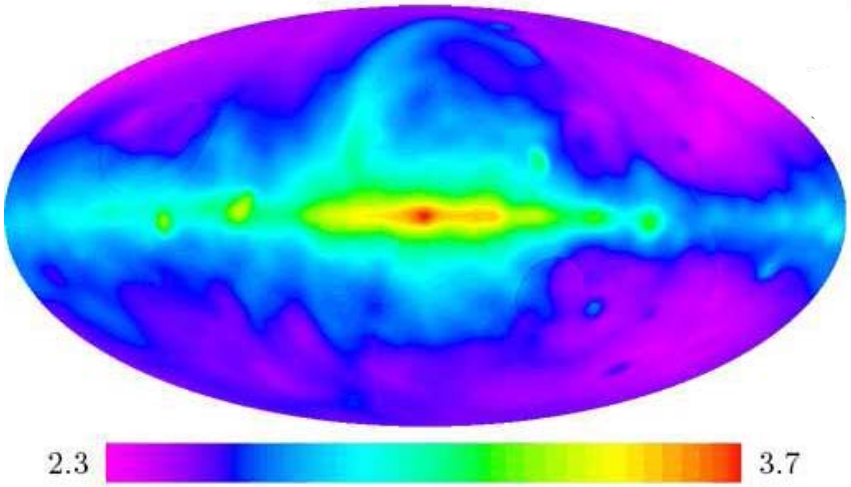
- Signal RMS ≈ a few mK (maybe up to ≈20mK).
- Noise ≈ a few tens of mK for integration times of a few hundred hours.
- Foregrounds (mainly synchrotron at large scales and unsubtracted point sources at small scales) ≈ a few K, but smooth as a function of frequency.
- Won't deal with point source subtraction and polarization here.

Typical line of sight



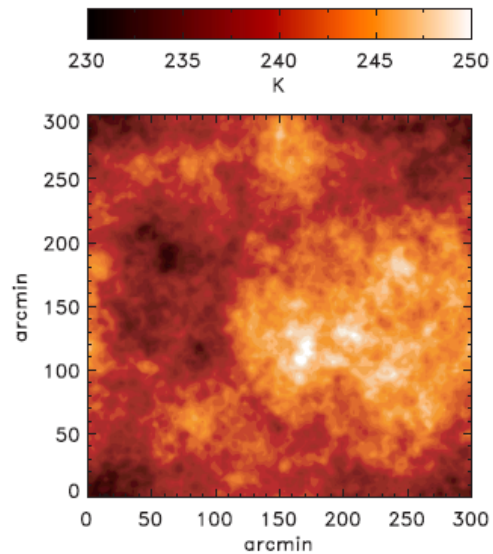
Foreground models

de Oliveira-Costa at al. (2008), 150 MHz



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

- Principal components analysis of existing radio surveys.
- Useful for observations covering a very large area.



Jelić et al. (2008)
120 MHz

- Uses a physical model for synchrotron, free-free, radio clusters etc.
- More representative of sky regions for planned LOFAR observations

Statistical approaches

Parametric

- Good results if the model really represents the foregrounds well.
- Somewhat inflexible.
- Fits well with estimating parameters of a signal model simultaneously.
- Usually computationally cheap.
- Some hope of choosing a parametrization from physics.

Non-parametric

- Big choice of techniques.
- Many techniques overfit: need to choose carefully.
- May still be some ‘parameters’ to choose, but this can add flexibility.
- Need only specify some general properties of the foregrounds, rather than a specific model.
- Sometimes computationally expensive.

Different spaces

(θ_x,θ_y,ν) space

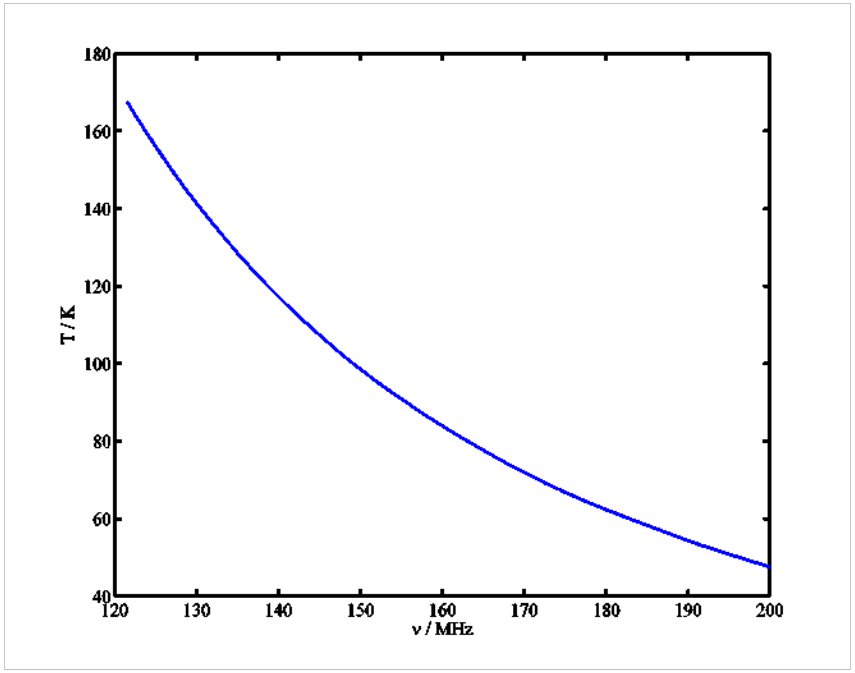
- Uniform noise properties across each image.
- Only need to fit a real function, but may be more pixels to fit.
- Hard to take into account correlation properties of adjacent pixels.
- Trouble with e.g. a chromatic p.s.f.

(u,v,ν) space

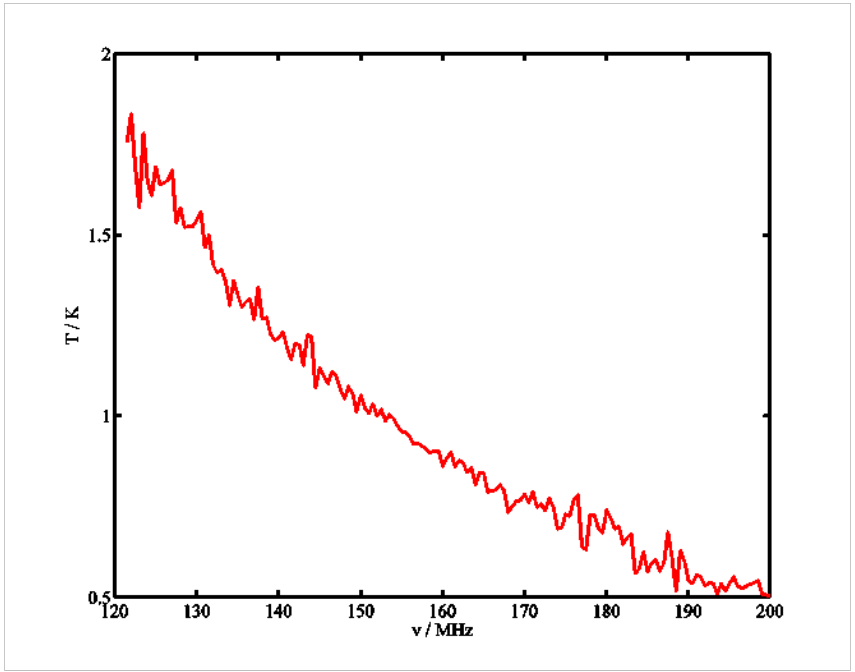
- Must fit a complex function at each point.
- Reasonable uv binning copes well with correlation properties.
- Better results for chromatic effects.
- Noise varies across the uv plane: need something adaptive.

Different uv ‘lines of sight’

Centre



Edge



Evaluating foreground fitting techniques

- For known foregrounds:
  - Size of fitting errors for individual lines of sight.
  - Residual RMS at a given frequency between different lines of sight: under-/over-fitting?
    - Looking at power spectra may be a more sophisticated version of the same thing.
  - Correlation of residuals with known foregrounds.

- For a realistic case:
  - Comparison between statistics where foreground residuals have different effects, e.g. power spectra from cross-correlation and from autocorrelation with estimated noise, and perhaps higher-order statistics.

Wp smoothing: non-parametric foreground fitting

- Model data points (x_i ,y_i) by:

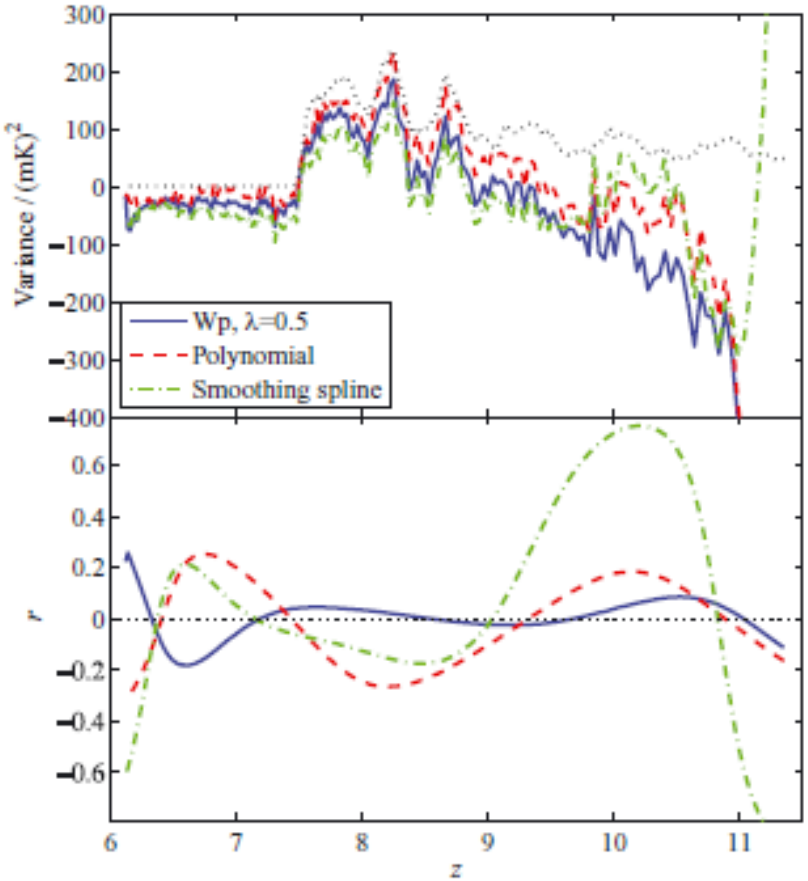
$$y_i = f(x_i) + \varepsilon_i, \;\; i = 1, \ldots, n$$

- Then we wish to solve the following problem:

$$\min_f \left\{ \underbrace{\sum_{i=1}^{n} \rho_i(y_i - f(x_i))}_{\text{“Least squares”}} + \underbrace{\lambda R[f]}_{\text{Roughness penalty}} \right\}$$

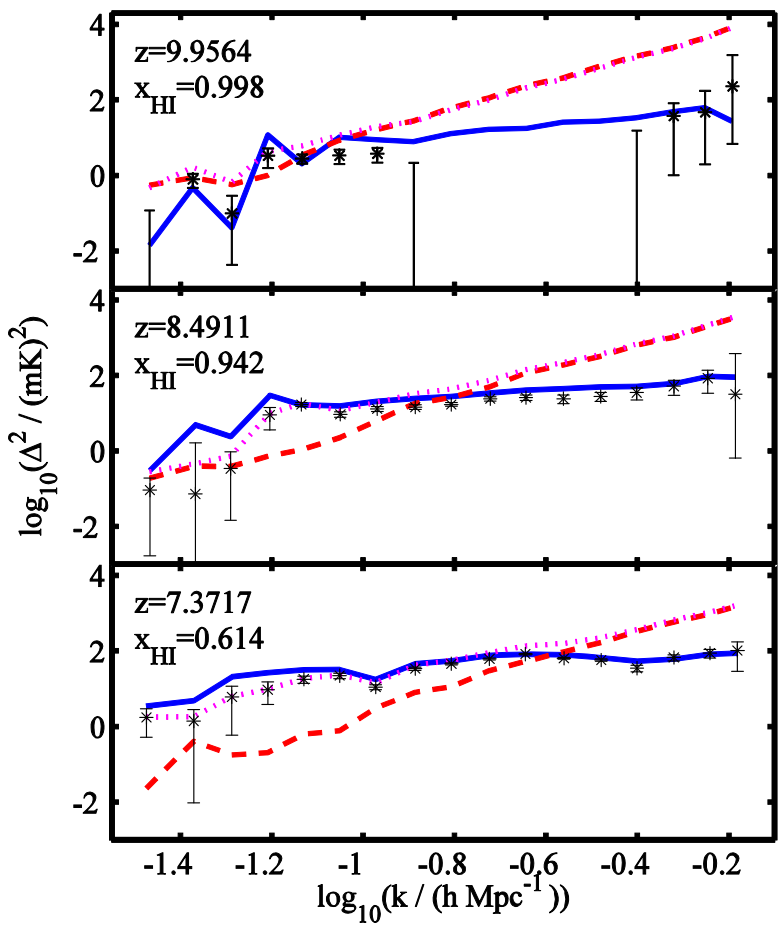
- Here the roughness penalty measures the integrated change in curvature ‘apart from inflection points’; inflection points are the primary measure of roughness.
- The solution of this minimization is the solution of a boundary value problem derived by Mächler.
- ‘Wp’ stands for ‘Wendepunkt’.

Simulated LOFAR results



- RMS as a function of frequency shows similar results for Wp smoothing and polynomial fitting.
- Polynomial residuals more correlated with foregrounds, however.
- Smoothing splines (a very simple non-parametric approach) do worse than either.

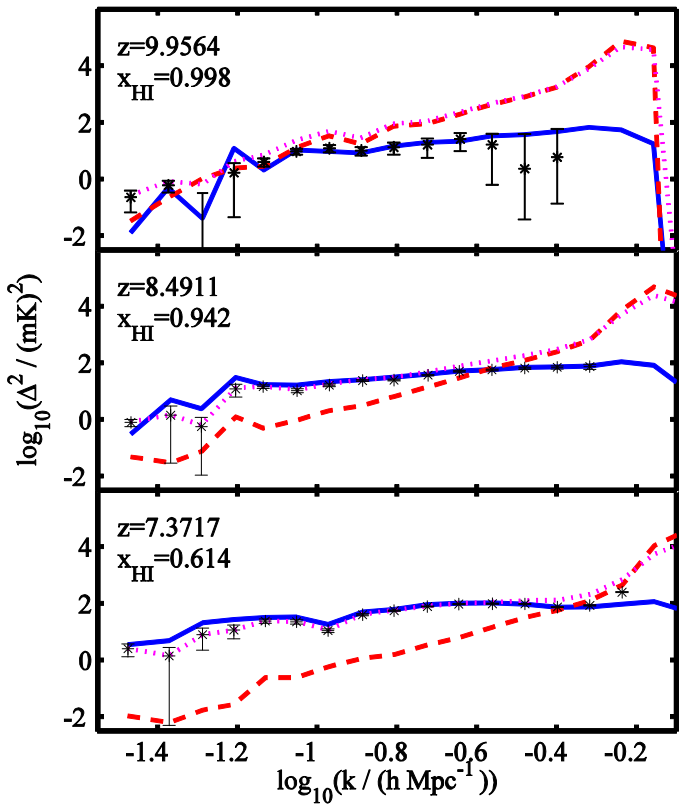
Example power spectra and errors



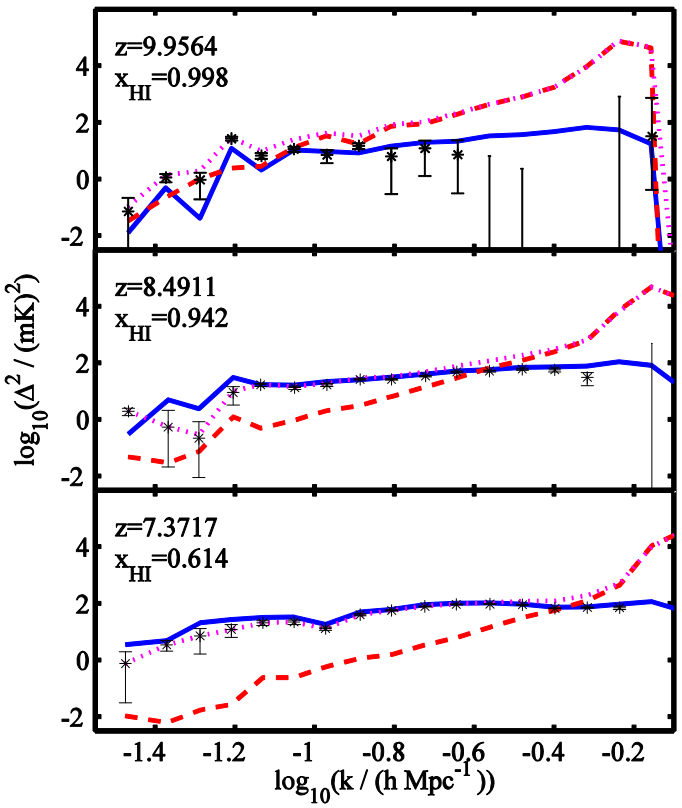
- Estimate the power spectrum by computing the autocorrelation of the foreground fitting residuals and subtracting a noise power spectrum
- 300 hrs, 1 beam, 1 window.
- Recover the power spectrum reasonably well at low redshift, but lose intermediate scales at high redshift.
- Some bias at large scales.

Fitting in the (*u*,*v*,ν) cube

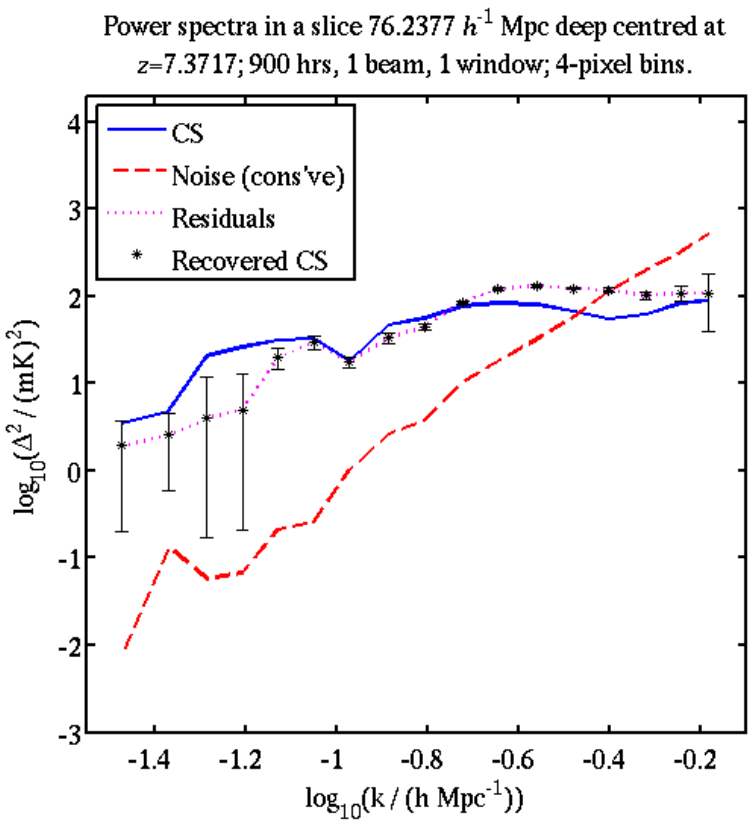
Wp smoothing



Polynomial fit



Cross-correlation



- Errors turn up in different combinations in a cross-correlation estimator.
- Other considerations:
  - Need to split data into epochs, and either fit foregrounds separately, or live with correlated errors between epochs.
  - Should yield positive power spectra.
  - Doesn't need such a good noise estimate.

Autocorrelation vs. cross-correlation

- Autocorrelation

$$P^r(k) = P^s(k) + P^n(k) + \underbrace{P^\epsilon(k) + \langle \epsilon(\boldsymbol{k})[s(\boldsymbol{k}) + n(\boldsymbol{k})]^* + [s(\boldsymbol{k}) + n(\boldsymbol{k})]\epsilon(\boldsymbol{k})^* \rangle_{|\boldsymbol{k}|=k}}_{\text{Fitting errors and cross terms (simulate?)}}$$

Residual power

Signal power

Noise power (estimate and subtract)

- Cross-correlation

$$\langle r_1 r_2^* \rangle = P^s + \underbrace{\langle s\epsilon_2^* \rangle + \langle \epsilon_1 s^* \rangle + \langle \epsilon_1 \epsilon_2^* \rangle}_{\text{Cross terms (no noise)}}$$

Cross-correlation of two epochs

Signal power

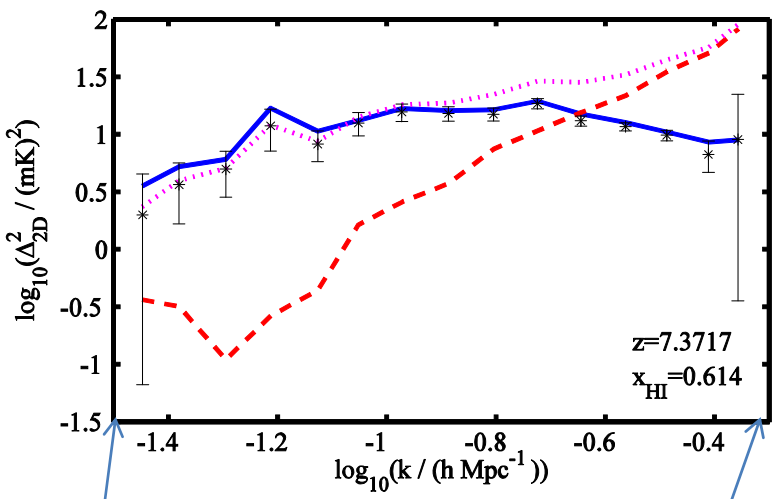
Angular and line-of-sight power spectra

Angular power spectrum

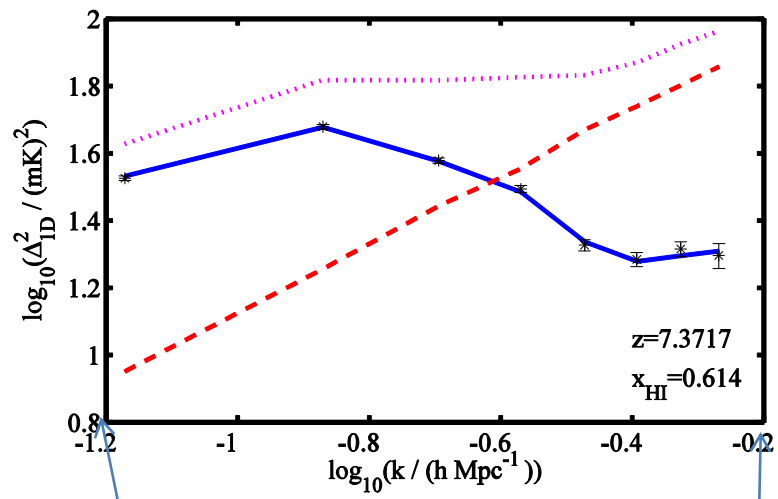
- Large-scale bias as for 3D power
- Goes to larger scales without risking evolution effects

Line-of-sight power spectrum

- Can reach smaller scales (depending on frequency resolution)
- No large-scale bias



-1.5 -0.3



-1.2 -0.2

Why does the bias show up in the angular power spectrum?

- We assume smoothness in the frequency direction, but the fitting leads to loss of power in angular modes.
- Along one line of sight, for a narrow frequency range, we are likely to make an error estimating the foregrounds which is roughly constant with frequency.
  - No change in the one-dimensional power spectrum over this frequency range.
  - For the angular power spectrum, this constant offset is likely to be different between different lines of sight, leading to bias in the power spectrum.
  - A similar offset between nearby lines of sight (because of large-scale correlation in the foregrounds) would lead to the offset being roughly constant within small regions, so the small-scale power loss would be small.

Summary

- Fair simulations of foregrounds are necessary to reasonably compare different approaches.
- Non-parametric methods may be a little more flexible and less model-dependent, but can be awkward to work with.
- Fitting in (u,v,v) space seems to work well and should probably be preferred to real space, since it helps overcome some nasty problems.
- Fitting can have different and complicated effects on angular and line-of-sight modes.
- It may be possible to tune power spectrum estimators to minimize the harmful effect of foreground fitting errors.
- Otherwise, what role can simulations play in estimating these errors in observations, in order to correct for them?