Extracting a signal from the epoch of reionization with LOFAR

Geraint Harker

Kapteyn Institute, Groningen, the Netherlands

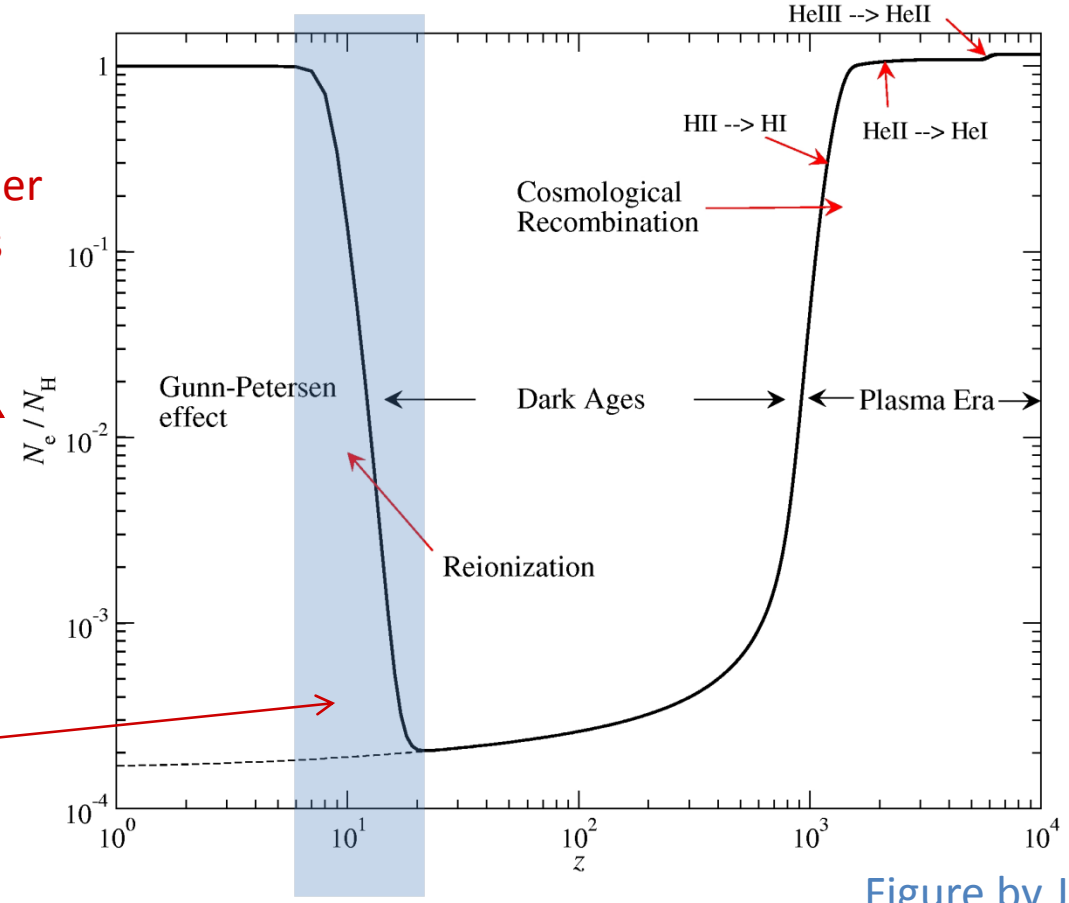
Collaborators: Ger de Bruyn;
Michiel Brentjens, Leon Koopmans, **Saleem Zaroubi**;
Gianni Bernardi, Benedetta Ciardi (MPA), Vibor Jelić, Panos Labropoulos, Garrelt Mellema (Stockholm), André Offringa, V. N. Pandey, Joop Schaye (Leiden), Rajat Thomas (IPMU, Japan), Sarod Yatawatta.

Outline

- Introduction
  - The physics of reionization and current constraints.
  - Observations: the 21cm line.
- Preparing for LOFAR EoR observations.
  - The project.
  - Analysis and extraction pipeline.
  - Synthetic data cubes.
- Foreground Fitting
- What do we want to measure?
  - Power spectrum.
  - Skewness?
- Results
- Summary, conclusions, further work.

Evolution of the ionization fraction: a rough picture

Number density of free electrons divided by number density of hydrogen atoms

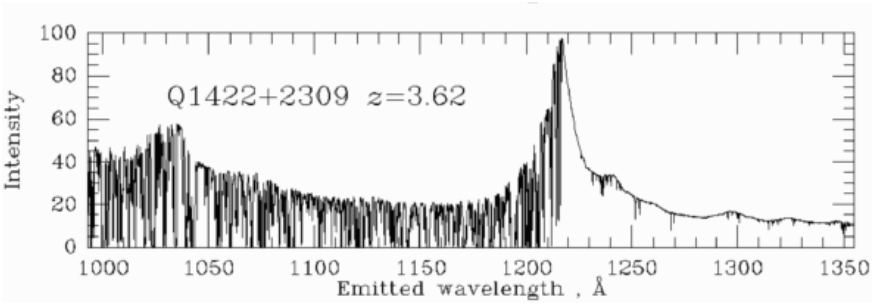


Epoch of reionization (EoR)

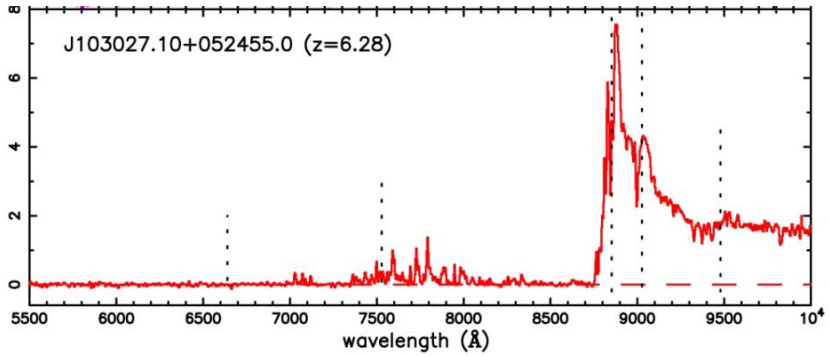
Figure by J. Chluba

Current constraints on the EoR: the Lyman α forest and SDSS quasars

- At low redshift, most hydrogen is ionized, and neutral patches show up as individual absorption features in quasar spectra.
- SDSS quasars: at $z \approx 6$ a sudden increase is seen in the flux decrement blueward of Lyman alpha (Becker et al. 2001; Djorgovski et al. 2001; Fan et al. 2002,2004): is this the tail end of reionization?



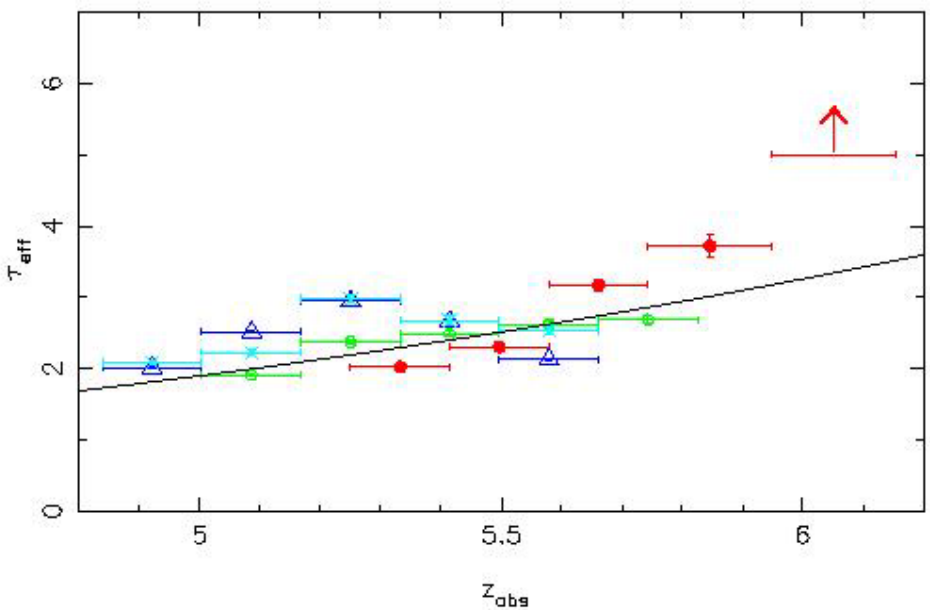
E.L. Wright, website



Becker et al. (2001)

Current constraints on the EoR: the Lyman α forest and SDSS quasars

- At low redshift, most hydrogen is ionized, and neutral patches show up as individual absorption features in quasar spectra.
- SDSS quasars: at $z \approx 6$ a sudden increase is seen in the flux decrement blueward of Lyman alpha (Becker et al. 2001; Djorgovski et al. 2001; Fan et al. 2002,2004): is this the tail end of reionization?

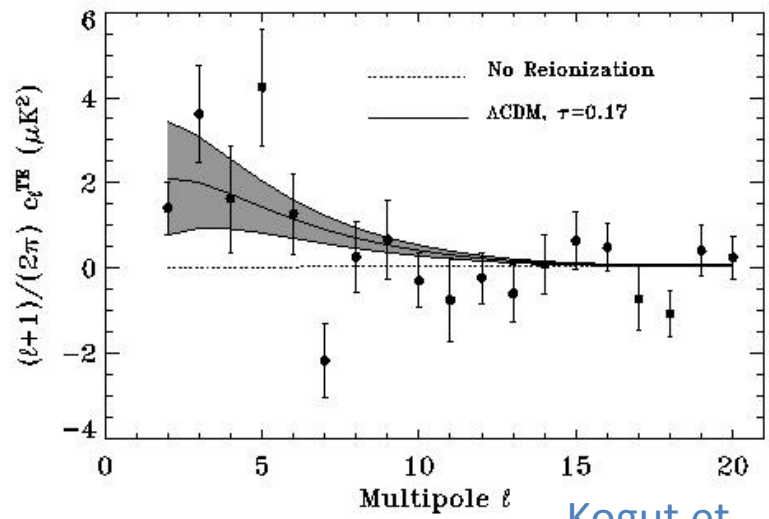


Becker et al. (2001)

Current constraints on the EoR: CMB polarization

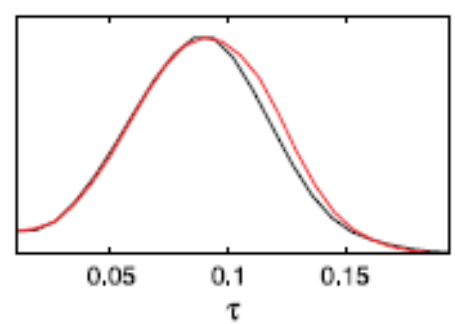
- Free electrons produced by reionization rescatter CMB photons and suppress small-scale anisotropies in temperature and polarization.
- The large-scale polarization anisotropies are *enhanced*, however.
- Page et al. (2007) give a redshift of instantaneous reionization $10.9^{+2.7}_{-2.3}$.

T-E cross-correlation



Kogut et al. (2003)

Likelihood



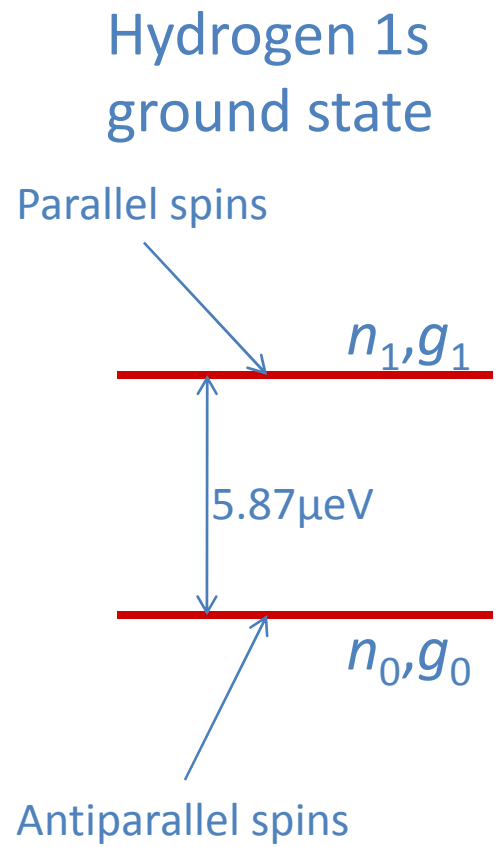
Page et al. (2007)

The hydrogen 21cm line

- The hydrogen 21cm (1420MHz) transition is a forbidden transition between the two ground-level states of hydrogen.
- The proportion of electrons in each of these states defines a 'spin temperature', T_{spin}, through:

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} \mathrm{e}^{-T_*/T_{\text{spin}}}$$

($T_* = 0.068$K)



Excitation mechanisms

- Collisions
  - H-H collisions (electron exchange)
  - H-e collisions (important near early X-ray sources)
- Wouthuysen-Field effect ('Lyman alpha pumping')

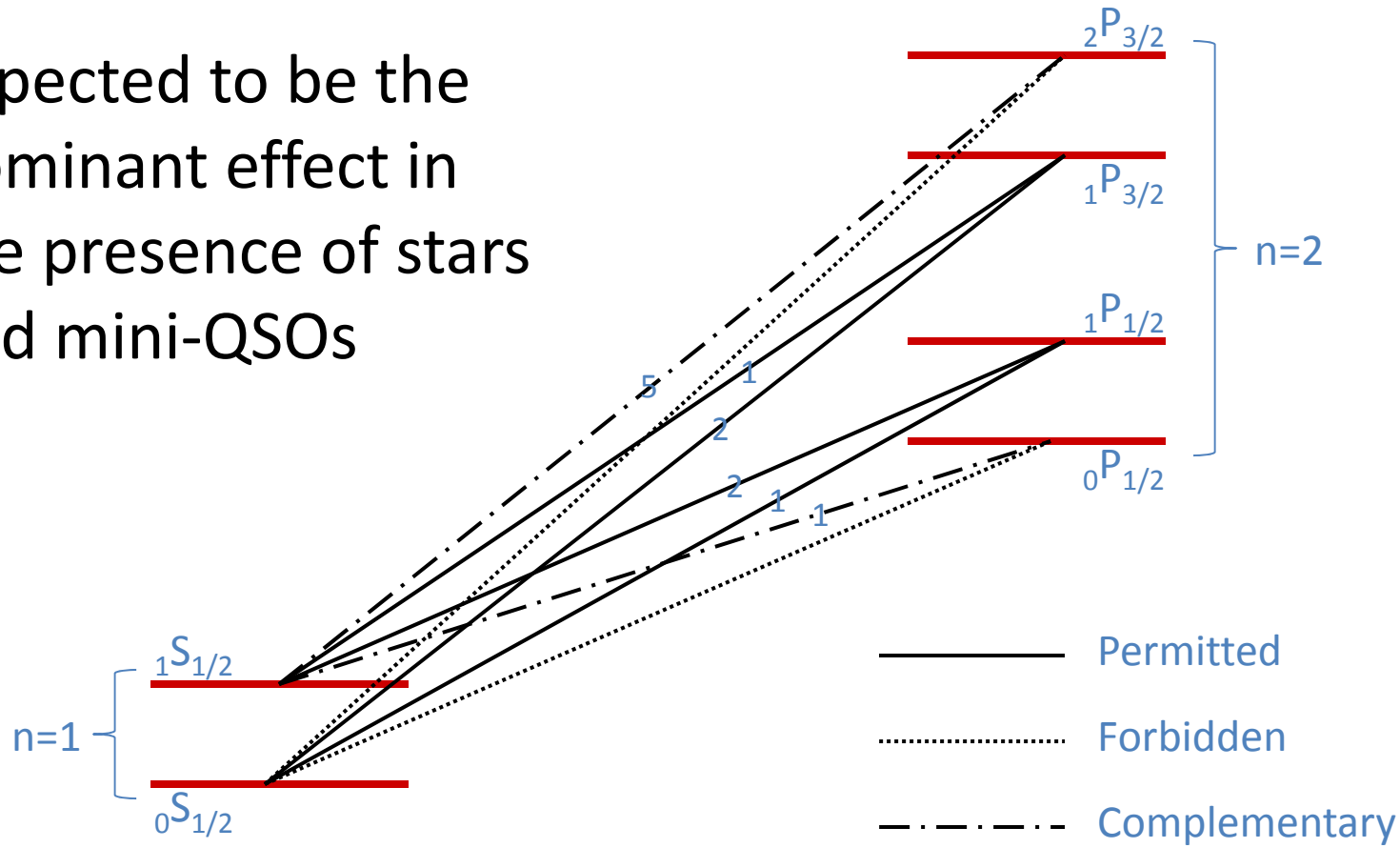
'Colour temperature', shown to be equal to T_k for the situation studied here

$$T_{\rm spin} = \frac{T_{\rm CMB} + y_\alpha T_\alpha + y_{\rm c} T_{\rm k}}{1 + y_\alpha + y_{\rm c}}$$

Field, 1958, Proc. IRE

Wouthuysen-Field effect

- Expected to be the dominant effect in the presence of stars and mini-QSOs



Differential brightness temperature

- To observe 21cm emission or absorption, the spin temperature needs to be decoupled from the CMB temperature.

$$\delta T_{\rm b} = \frac{T_{\rm spin} - T_{\rm CMB}}{1+z} \overbrace{(1 - {\rm e}^{-\tau_{\nu_0}})}^{\approx \tau_{\nu_0}}$$

$$\tau_{\nu_0} = \frac{3}{32\pi} \frac{hc^3 A_{10}}{k_{\rm B} T_{\rm spin} \nu_0^2} \frac{x_{\rm HI} n_{\rm H}}{(1+z)({\rm d}v_{\parallel}/{\rm d}r_{\parallel})}$$

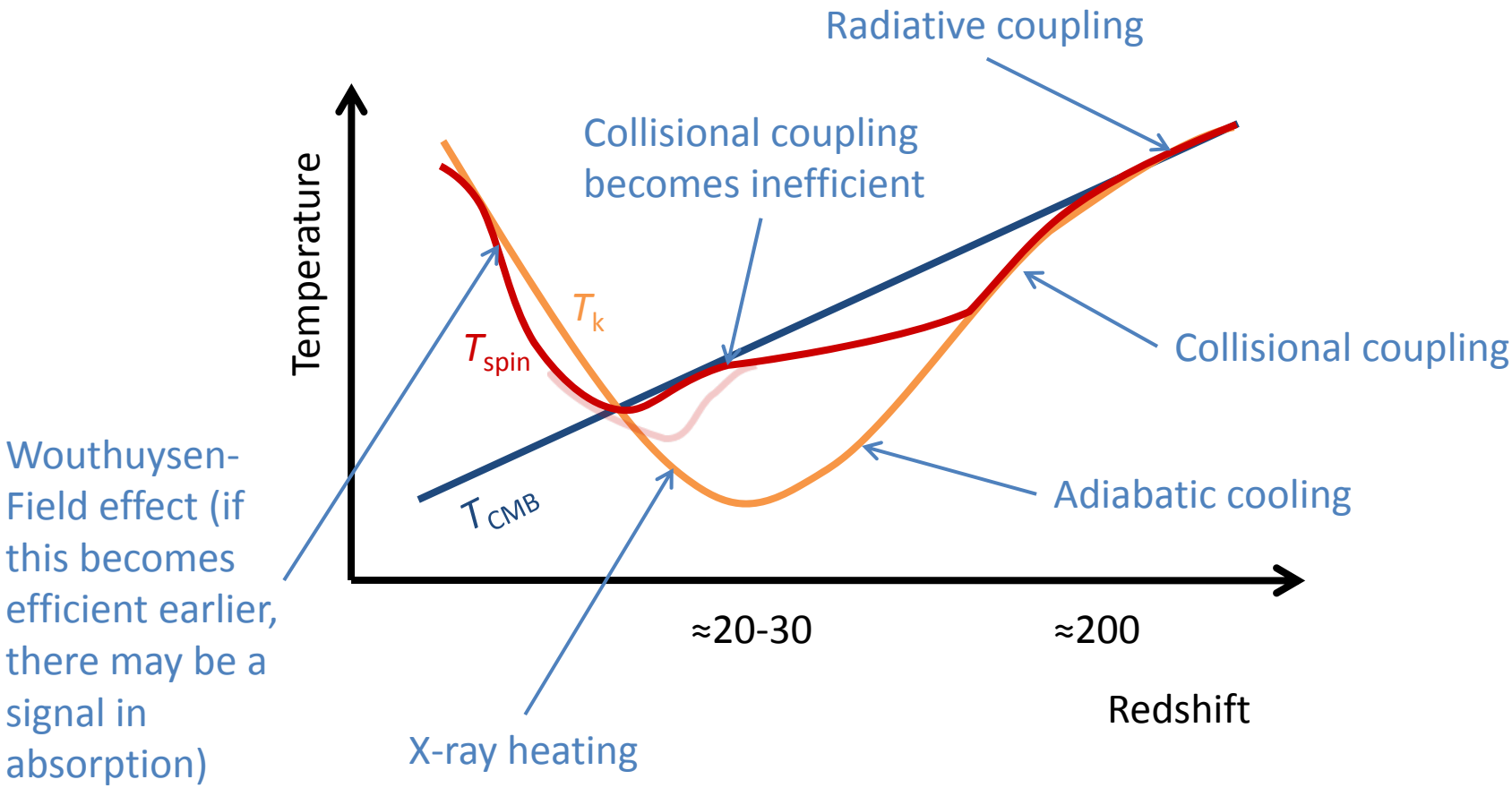
Differential brightness temperature

- δT_b depends on position and redshift, which in principle allows us to carry out tomography of high redshift neutral hydrogen.

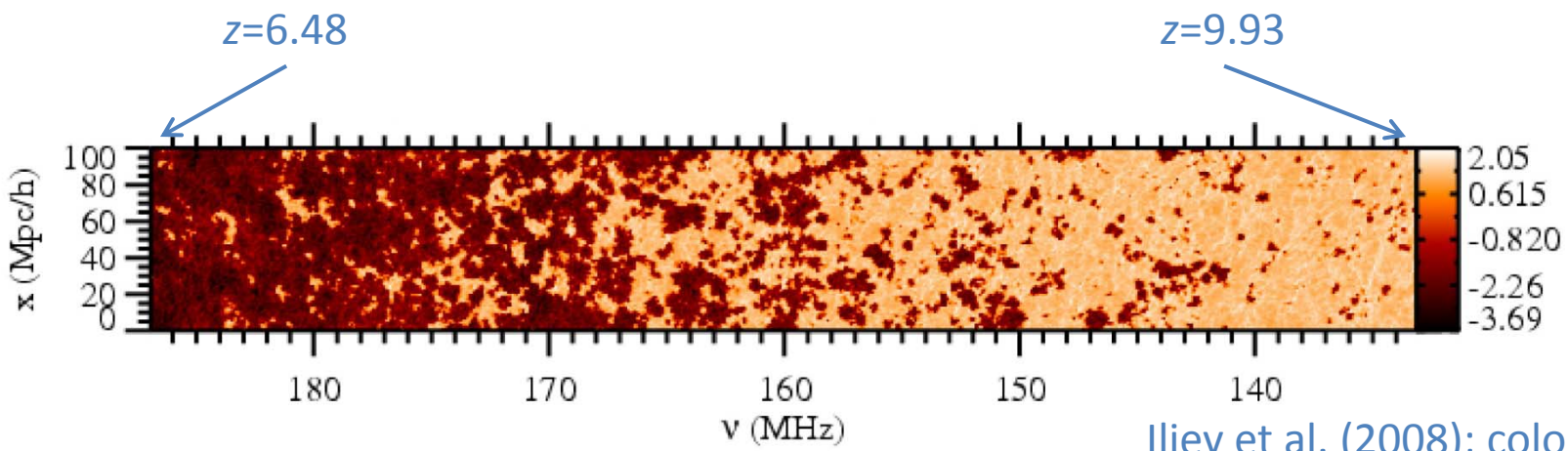
$$\frac{\delta T_{\rm b}}{\rm mK} = 39h(1+\delta)x_{\rm HI}\left(1-\frac{T_{\rm CMB}}{T_{\rm spin}}\right)\left(\frac{\Omega_{\rm b}}{0.042}\right)\left[\left(\frac{0.24}{\Omega_{\rm m}}\right)\left(\frac{1+z}{10}\right)\right]^{\frac{1}{2}}$$

- Contains information on:
  - the growth of structure (through 1+δ);
  - reionization (through *x*), e.g. growth of bubbles;
  - heating (through dependence on the spin temperature);
  - cosmology;
  - redshift-space distortions (through extra $\left[1-\frac{1+z}{H(z)}\frac{\partial v}{\partial r_v}\right]$ term).

Temperature evolution



Simulated cosmological signal



Iliev et al. (2008); colour scale shows $\log_{10}(\delta T_b)$

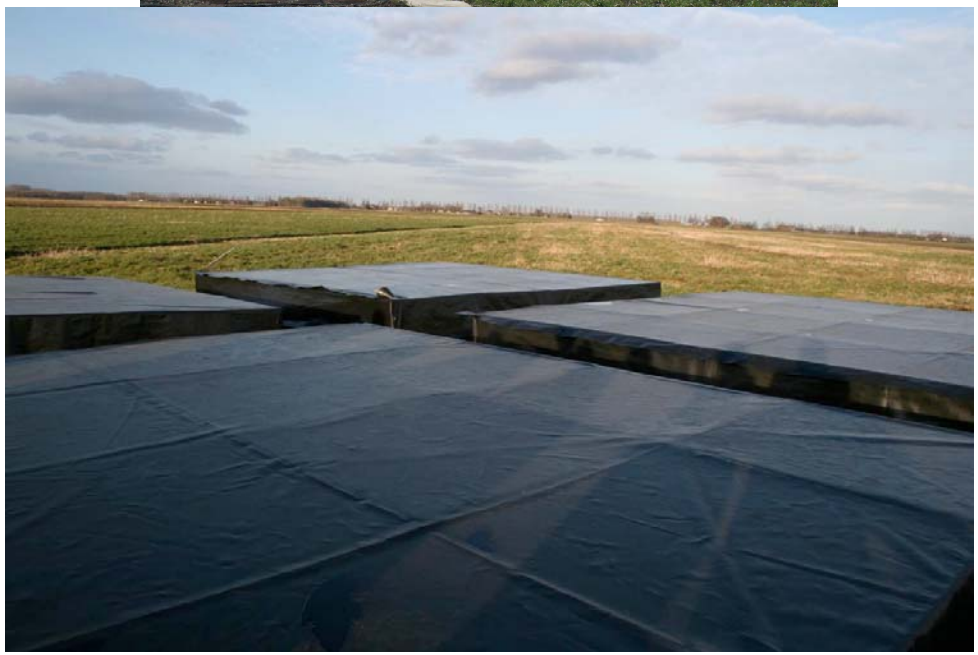
- 'Slice of sight' assuming that the spin temperature is large enough to saturate.
- Reionizing radiation here comes from stars.

The LOFAR EoR experiment

- The LOFAR EoR project is one of several (current or planned) experiments with some similar aims: MWA, GMRT, 21CMA, PAPER, SKA…
- LOFAR itself is:
  - an interferometer (as opposed to a single-antenna experiment such as EDGES).
  - an observatory, rather than a dedicated EoR experiment. There are four other key science projects:
    - Surveys
    - Transients
    - Cosmic rays
    - Magnetism
  - associated with some non-astronomical projects.

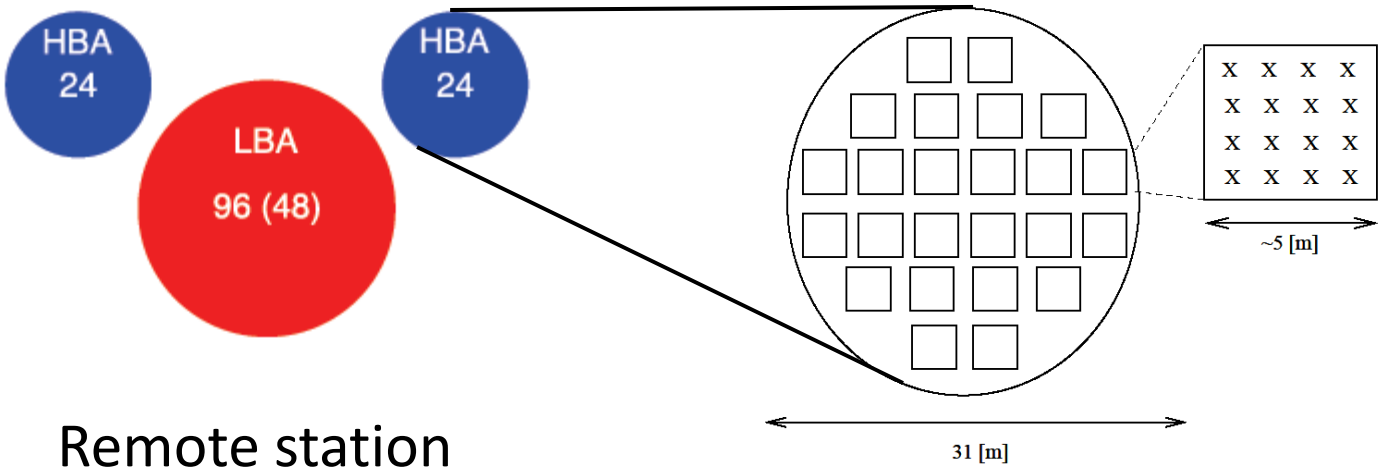
The LOFAR telescope

- Low-band antennas
  - 30-80 MHz
  - Not used for EoR work
- High-band antennas
  - 110-240MHz
  - 115MHz → z=11.35
- ≈18 stations in Dutch 'core' (<2km)
- ≈18 stations in 'extended array' (<100km); may use inner 'rings' for EoR.
- European baselines up to 1000km.

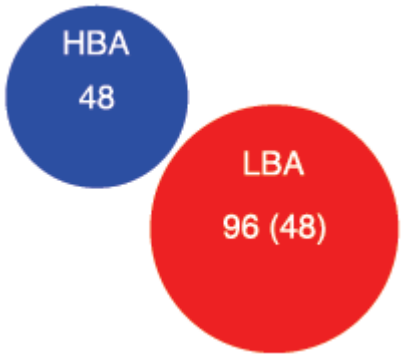


LOFAR stations

Core station

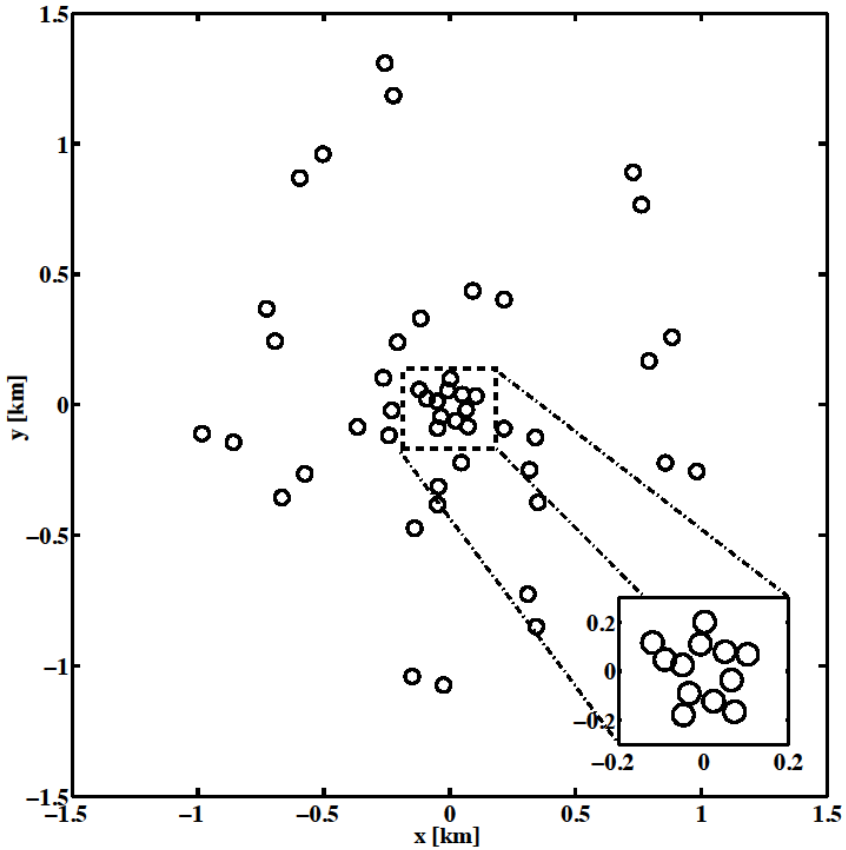


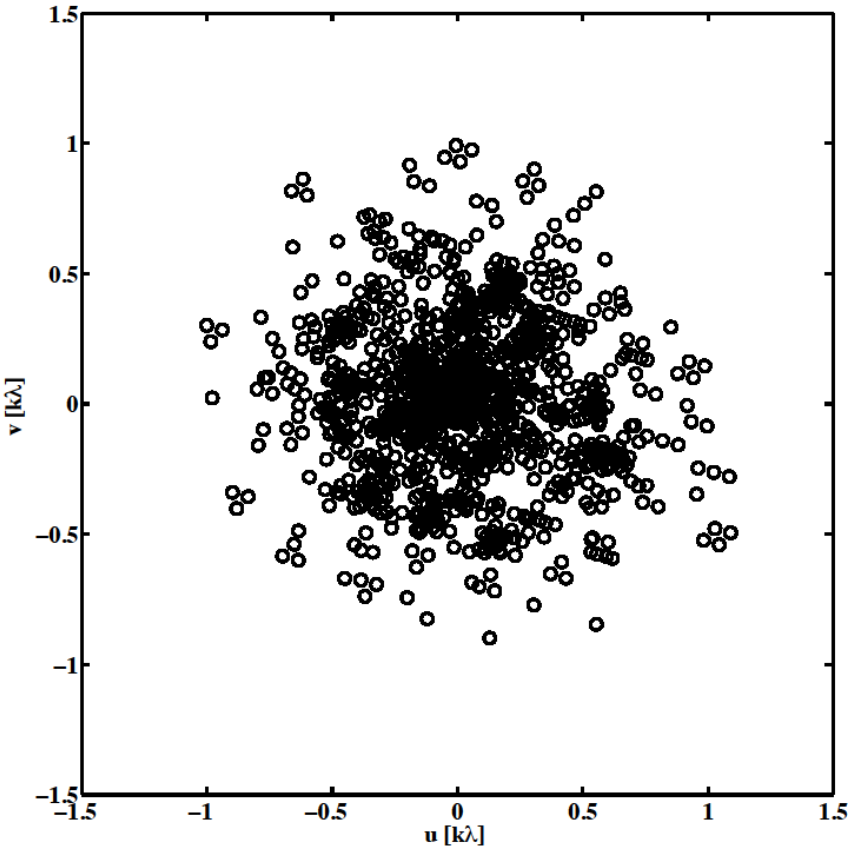
Remote station



- Effective area of one HBA 'semi-station' at 150MHz is 512 sq. m.
- Split stations give a baseline of approx. 100m.

Station layout in the core





Aims of the project

- Since tomography (signal to noise per resolution element greater than unity) is not expected to be possible with LOFAR, we aim to make a statistical detection of 21cm radiation from the EoR.
- As more data are collected, we will move beyond a mere detection to constrain models of reionization: most commonly, people envisage doing this through the power spectrum of δT_b fluctuations.
- Probes other than statistical measurements of the whole standalone data cube may be possible, though not discussed in detail here:
  - Environment of high redshift QSOs
  - Cross-correlation with other data sets
    - CMB (Planck)
    - Lyman alpha emitters
    - JWST, eventually?
  - 21cm forest, if there are sufficiently strong background sources.
- Extra science: physics of galactic emission; ionospheric studies; deep, long-baseline studies of foreground sources; faint transients.

History from an EoR point of view

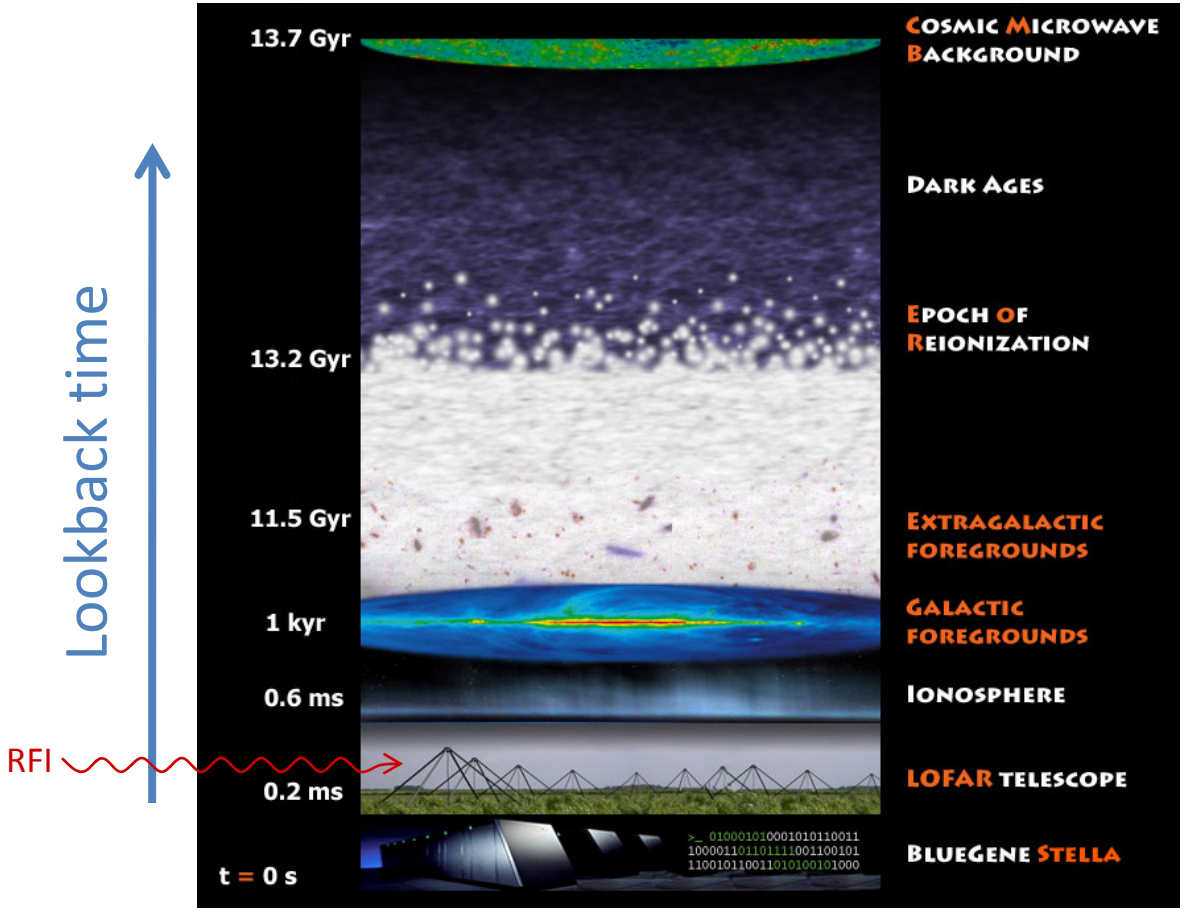
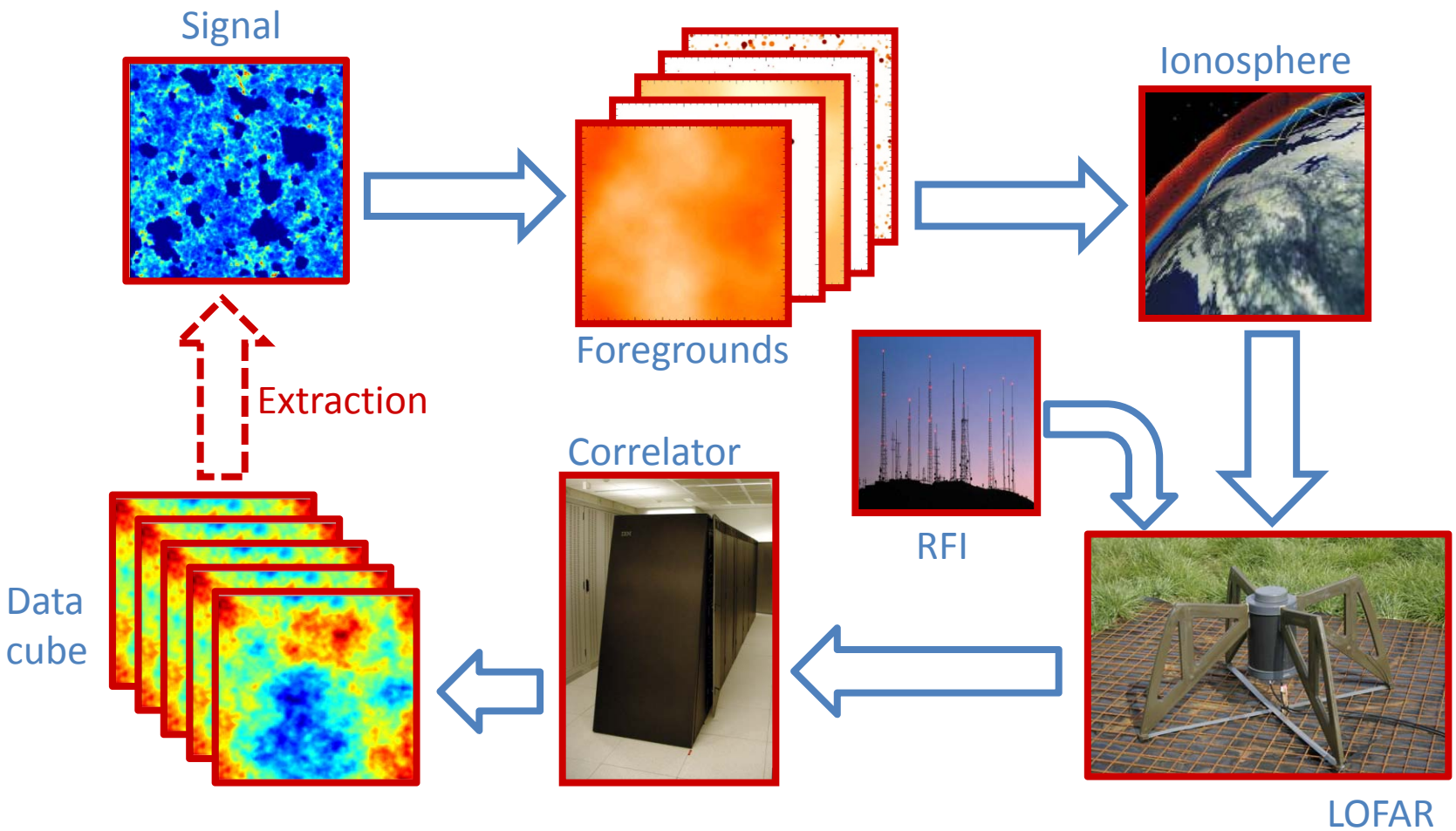
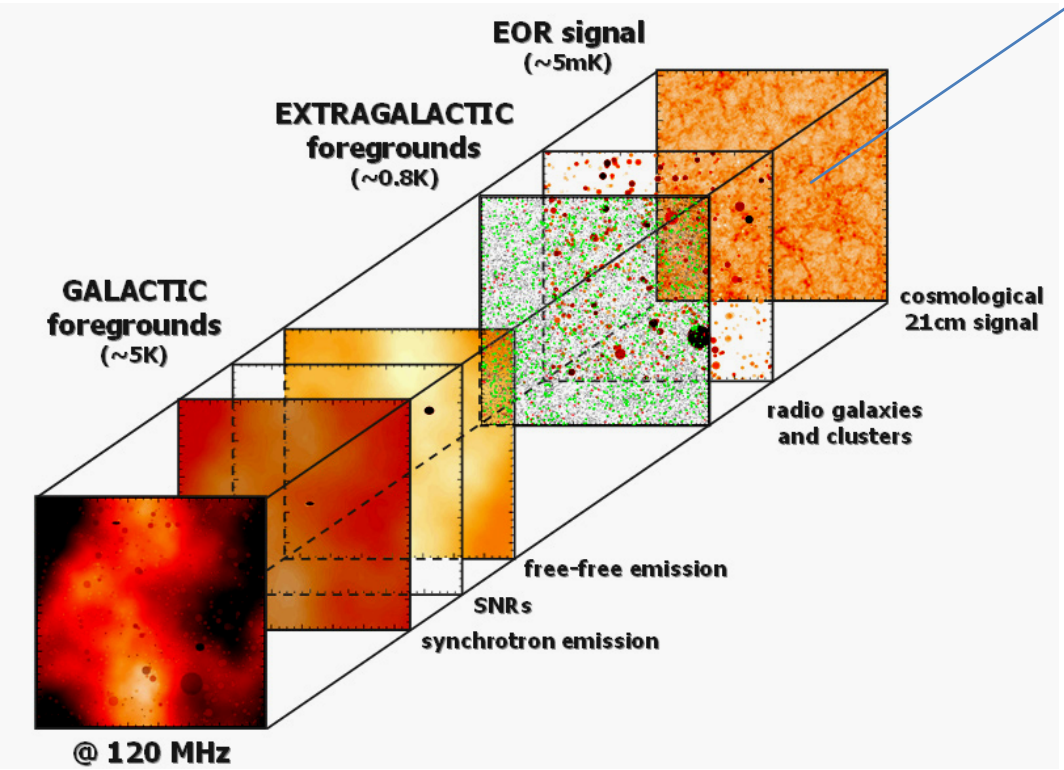


Figure by V. Jelić

Testing pipeline



Components of the data cubes



Foregrounds as in Jelić et al. (2008).

- Cosmological signal here from the f250C simulation of Iliev et al. (2008).
- Foregrounds and cosmological signal are convolved with the instrumental response.
- Uncorrelated noise in the *uv* plane: for one year with one beam, corresponds to an *rms* of 52mK at 150 MHz.
- Need foregrounds which are smooth as a function of frequency.

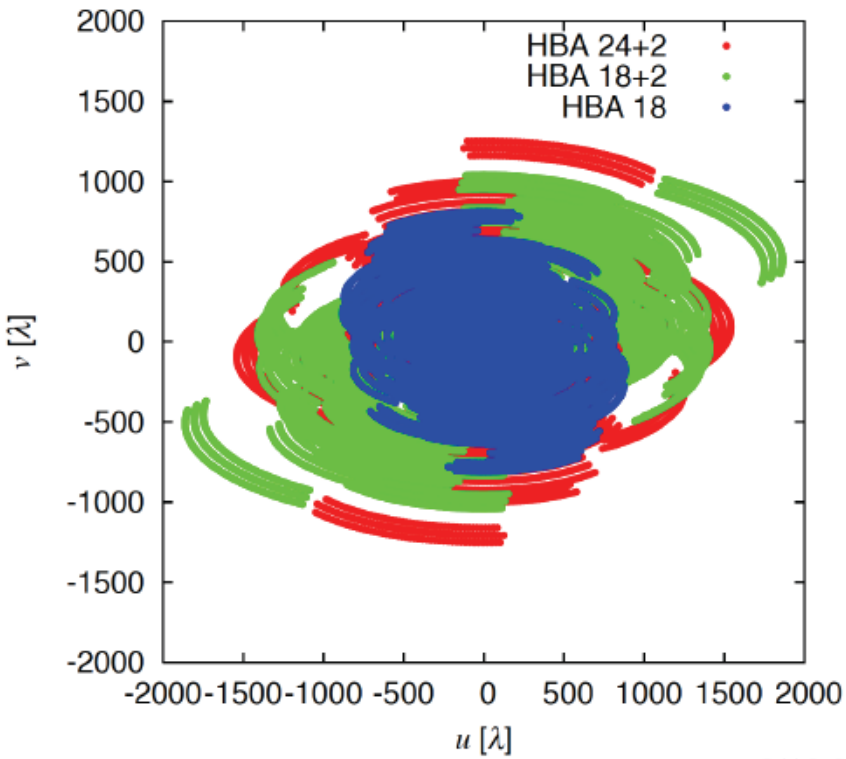
Signal simulations

- Results here use the simulation f250C of Iliev et al. (2008)
  - 100 Mpc/*h*
  - 1624^3 dark matter particles; 3D radiative transfer on a 203^3 grid (C2-Ray).
  - Stars provide the ionizing radiation.
  - Assume $T_{spin} >> T_{CMB}$
- To explore parameter space, we have also been using fast, approximate simulations using 1D radiative transfer (Thomas et al. 2008).
  - Study heating e.g. by X-rays
  - Different sources of reionization
- To produce a final data cube we interpolate between outputs at different redshifts.
- Observing window is roughly 5°x5° (over 600Mpc/*h* at redshift 10) and so to fill it we must then tile copies of the cosmological signal cube.
- Results restricted to scales <100Mpc/*h* for now, but bigger simulations are on the way.

Foregrounds

- Assume that bright point sources have been subtracted.
- Galactic foregrounds
  - Synchrotron (dominant component).
  - Free-free.
  - Supernova remnants.
- Extragalactic foregrounds
  - Radio galaxies (FRI and FRII).
  - Radio clusters.
- Foregrounds are assumed to be **smooth** in frequency; without this assumption, the situation is probably hopeless.

uv coverage

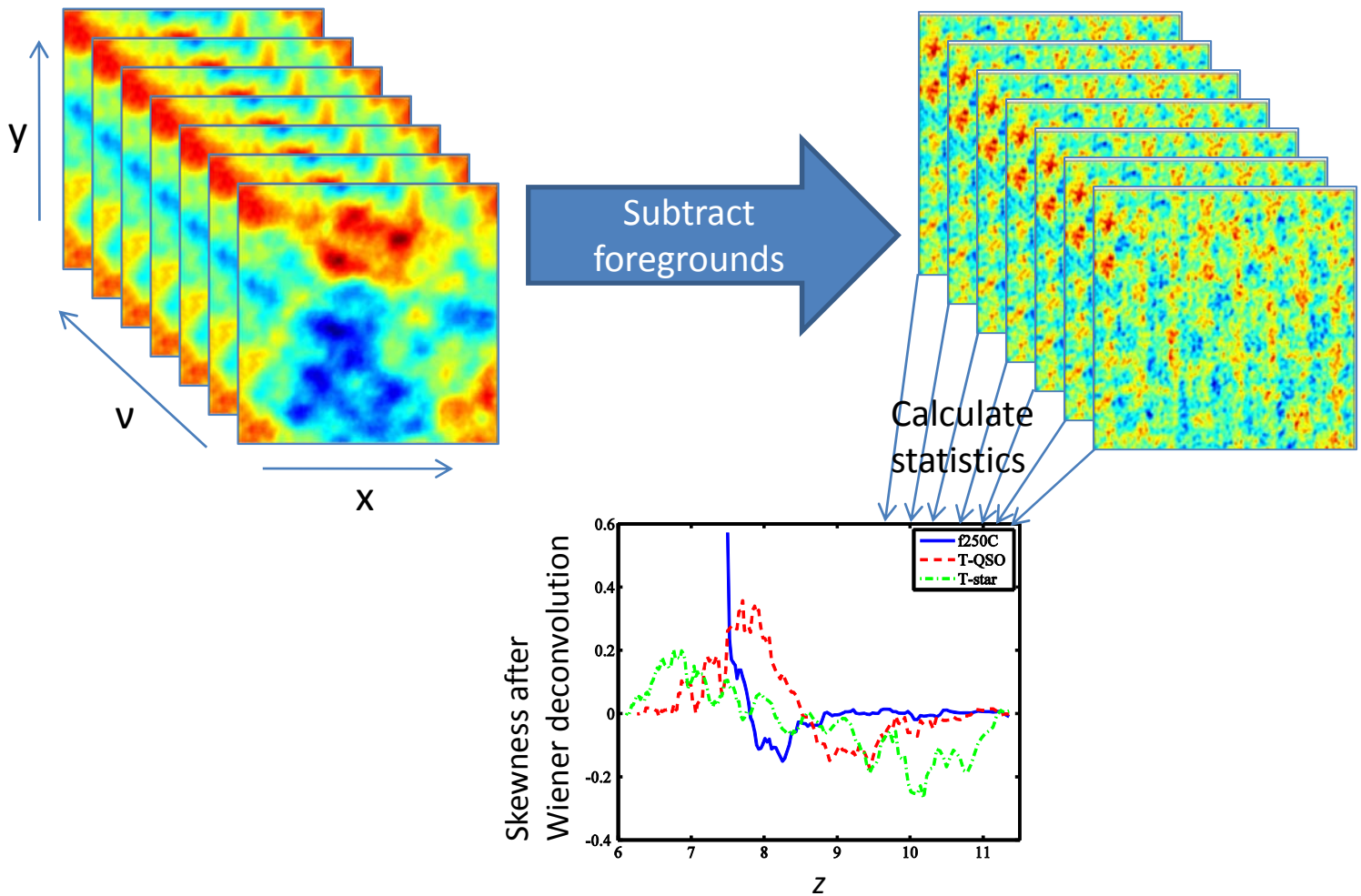


- Uniform *uv* coverage is important, else we can mix spatial power with spectral power since the coverage scales with frequency.
- Holes are especially damaging, but even without them the scaling of the *uv* coverage means a frequency-dependent PSF.
- Results here (mostly) assume the same coverage at all frequencies: this will change!

Noise

- Noise on individual baselines is independent; in other words, noise in the *uv* plane is uncorrelated.
- This implies that noise in the image plane is correlated on the scale of a resolution element.
- For one 'year' (300 hours) of observing time with LOFAR in one window and with one synthesized beam, we achieve noise of 52mK in a 1MHz band at 150MHz.
- Compare foreground fluctuations of order a few Kelvin, and fluctuations in the signal of order a few millikelvin.
- Only statistical detection of a signal is possible.

Signal extraction in two stages



Wish list for a foreground fitting algorithm

- Accuracy.
- Lack of bias.
- Avoidance of under-fitting or over-fitting.
- Make minimal assumptions about the functional form of the foregrounds; i.e., exploit their smoothness directly.
- Speed (less important if we only wish to subtract the foregrounds once, in post-processing).

Statistical approach

- Model data points (x_i ,y_i) by:

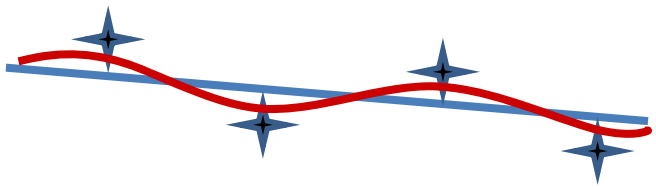
$$y_i = f(x_i) + \varepsilon_i, \;\; i = 1, \ldots, n$$

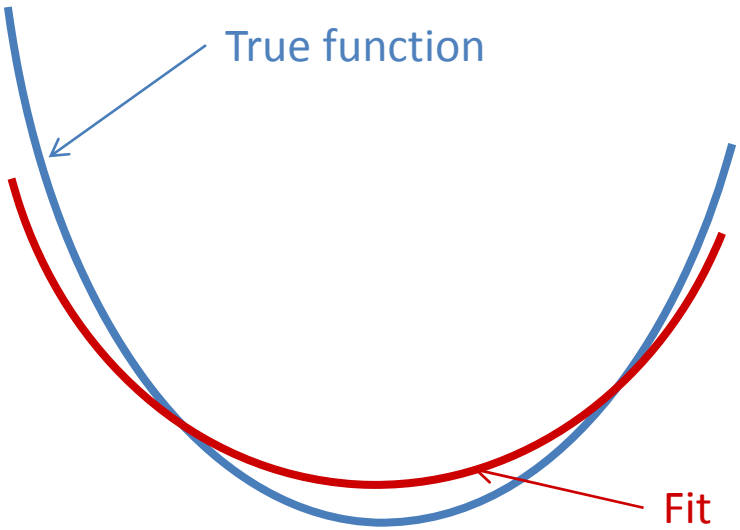
- Then we wish to solve the following problem:

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

Choosing a roughness penalty $R[f]$

- Require a roughness penalty that stops the curve wiggling towards individual data points, but avoids the problem of attrition.
- 'Smoothing splines' use integrated curvature as the roughness penalty, but in Wp smoothing the integrated *change* of curvature is used instead.





Wp smoothing

- An approximation to the change of curvature, *f'''/f''*, blows up at the inflection points *f''*=0.
- *R*[*f*] measures the change of curvature 'apart from the inflection points', w_i
- Perform the minimization with the position of the inflection points (and s_f) fixed.

$$R[f] = \int_{x_1}^{x_n} h'_f(t)\mathrm{d}t$$

$$f''(x) = p_{\mathbf{w}}(x)e^{h_f(x)}$$

$$\begin{aligned} p_{\mathbf{w}}(x) &= s_f(x-w_1)(x-w_2) \\ &\quad \times \ldots (x-w_{n_w}) \end{aligned}$$

Wp smoothing

- Mächler (1993,1995), who proposed the method, showed that the variational problem leads to the following differential equation:

$$h''_f = p_{\mathbf{w}} e^{h_f} \left[-\frac{1}{2\lambda} \sum_{i=1}^{n} (x - x_i)_+ \psi_i (y_i - f(x_i)) \right]$$

where $a_+ = \max(0,a)$, $\psi_i(\delta) = \frac{\mathrm{d}}{\mathrm{d}\delta}\rho_i(\delta)$, and the boundary conditions are

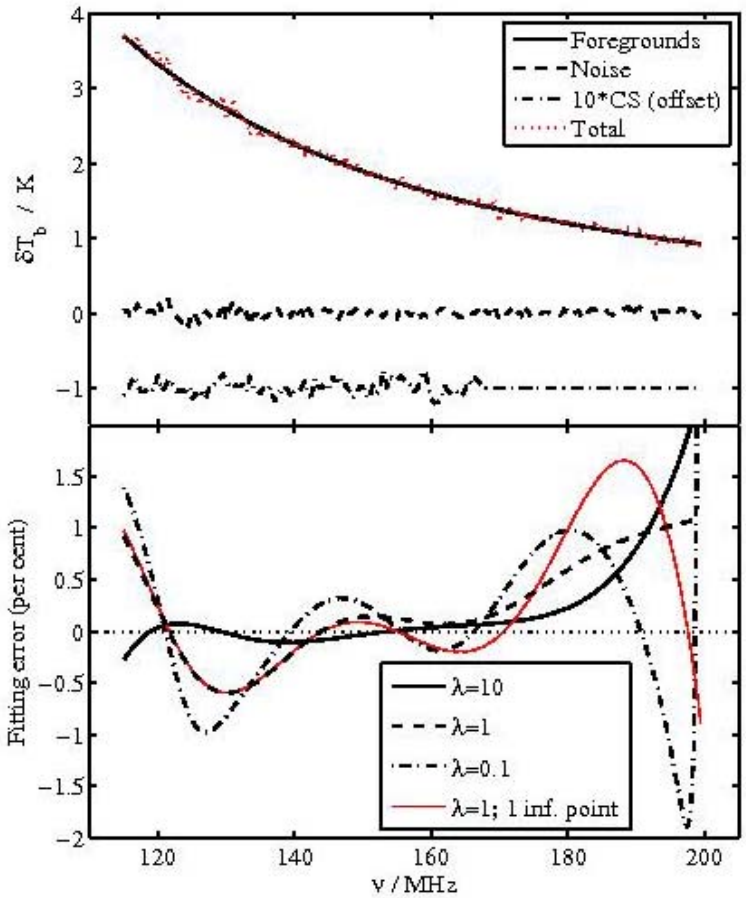
$$h'_f(x_1) = h'_f(x_n) = \sum_i \psi_i (y_i - f(x_i)) = \sum_i x_i \psi_i (y_i - f(x_i)) = 0$$

Implementation

- In general we need a method to find the number of inflection points, and need to perform a further minimization over their position.
- For the foreground fitting we find that it works well to have no inflection points (this would be the case anyway for a sum of negative-index power laws).
- The differential equation and the boundary conditions are in a nonstandard form:
  - Can rewrite as a system of 5*n*-4 coupled first-order equations and use a standard BVP solver.
  - Alternatively, convert to a finite difference equation and perform a multidimensional function minimization (seems better so far).
- Either approach requires a reasonable initial guess for the solution; we fit a power law since this has no inflection points.

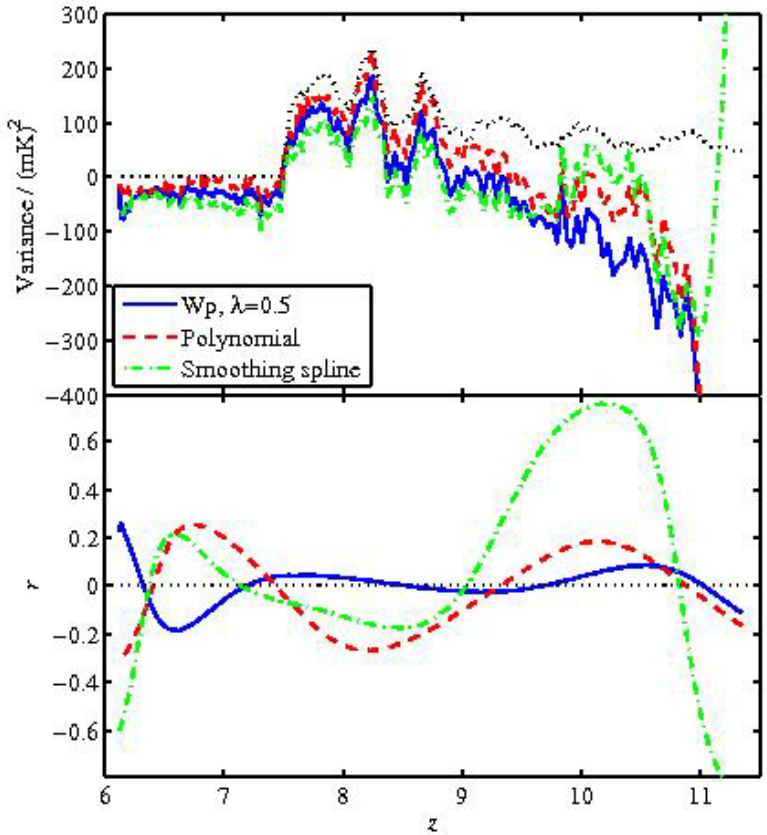
An example line of sight

- Since an interferometer cannot measure a mean, the signal along an individual line of sight can be positive or negative.
- Foregrounds typically a few Kelvin (smaller FG and noise at high frequencies).
- Fitting errors typically less than 1% using Wp smoothing.



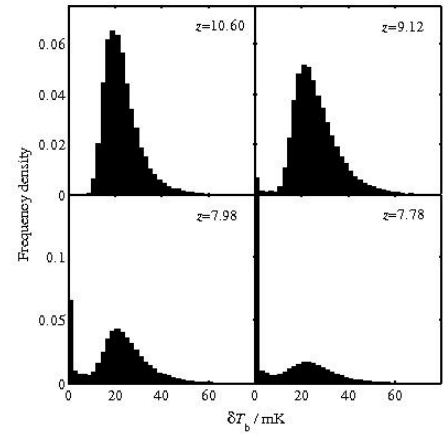
Statistics of residual maps

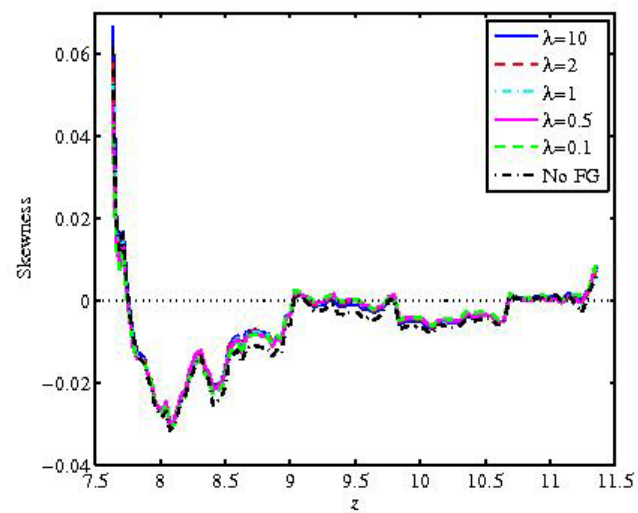
- Perhaps the most obvious way of detecting a signal is to find excess variance over and above the expected noise level in the residual maps.
- The complete variance of the cosmological signal is not recovered because of over-fitting.
- Under-fitting induces correlations between the fitting errors and the foregrounds.
- Wp smoothing seems to minimize this under-fitting while still allowing a detection of excess variance.



An alternative statistic: skewness of the one-point distribution.

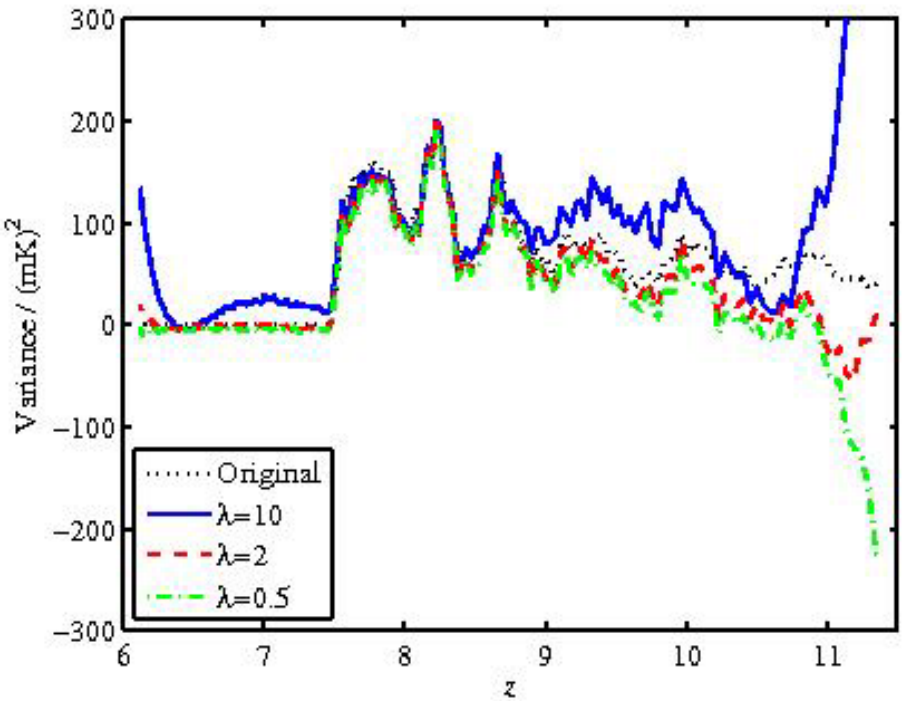
- The growth of structure and bubbles induces skewness in the one-point distribution of maps of δT_b.
- The redshift evolution of skewness may show characteristic features from reionization (especially negative skewness)
- This can be extracted and may confirm a detection of reionization using other statistics.
- Extraction involves a deconvolution step that requires knowledge or an estimate of the power spectrum.



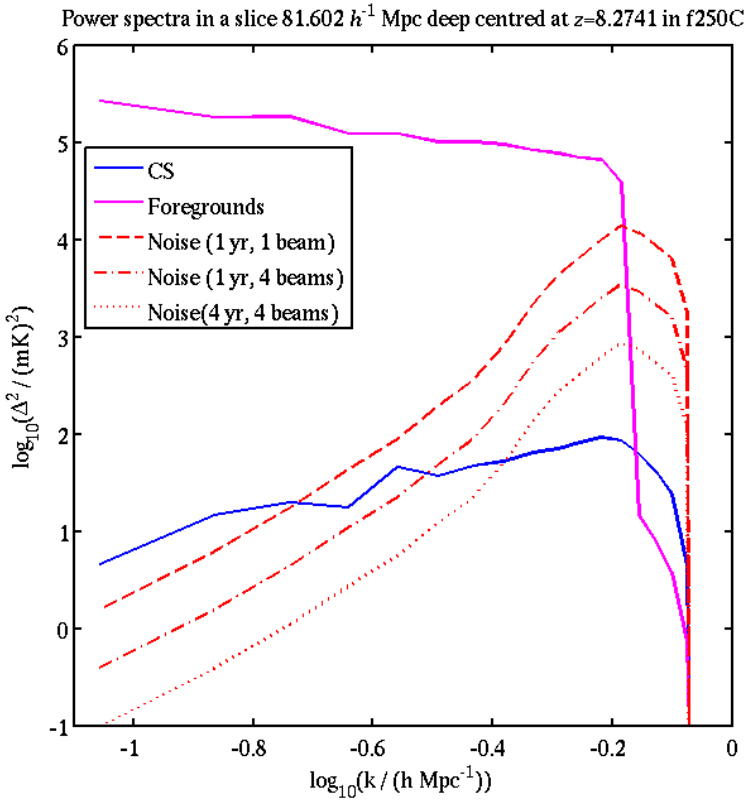


Scale dependence: variance of smoothed maps

- Our 256x256 pixel maps are oversampled.
- Smooth with a 4x4 boxcar filter before fitting foregrounds.
- This retains more of the variance of the cosmological signal.
- Also highlights over-fitting at high redshift.

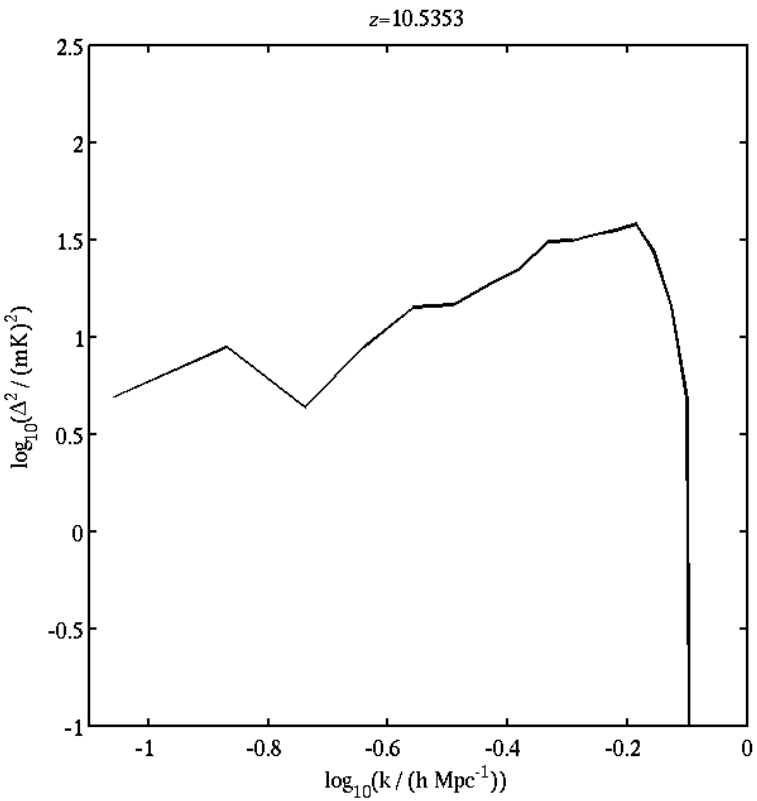


Scale dependence and size of components of the signal



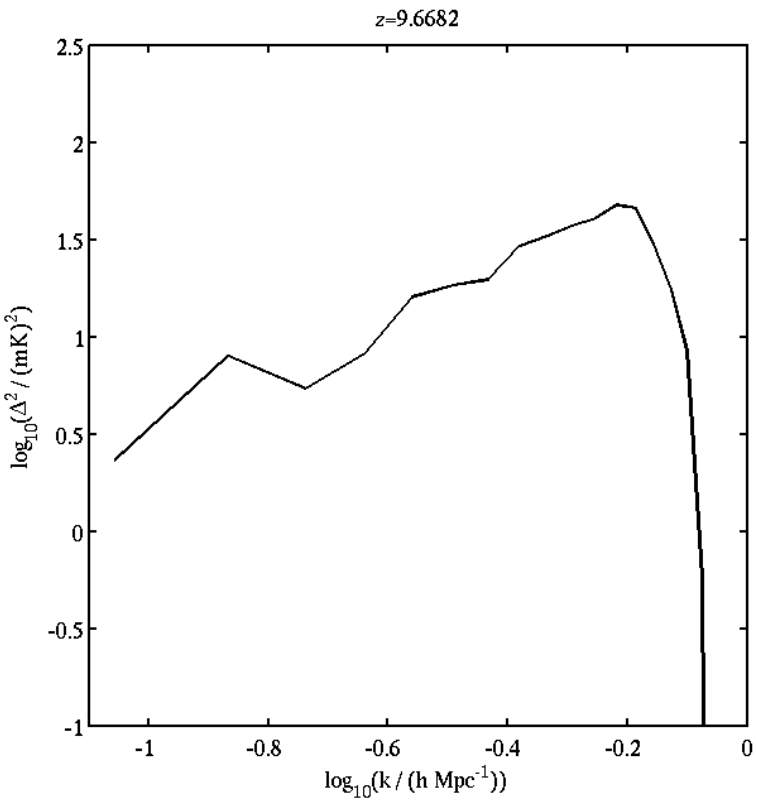
- Noise (receiver noise plus sky noise) dominates on small scales, leading to problems from over-fitting.
- Foregrounds dominate on larger scales, leading to problems from under-fitting.
- All scales contribute to the integrated RMS, but using the whole power spectrum we may be able to pick out the most favourable scales.
- Recovered shape provides a further check?

Example evolution of the signal power spectrum



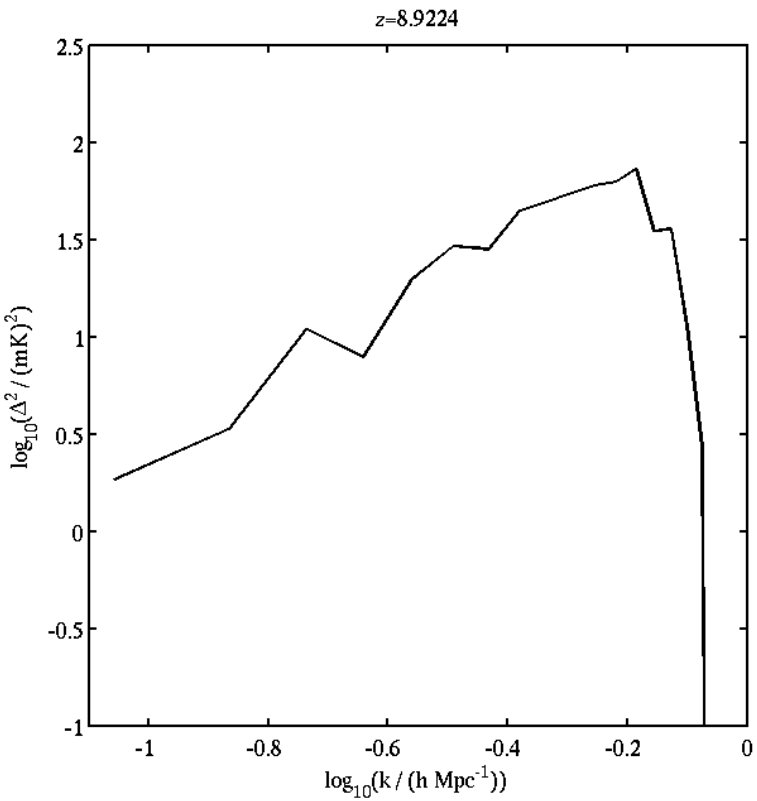
- Spectra here are convolved with the instrumental response (hence high-*k* cut-off).
- Growth of a feature on small scales due to formation of bubbles.
- Feature broadens and moves to larger scales as bubbles grow.
- At low redshift, signal drops because of a low neutral fraction.

Example evolution of the signal power spectrum



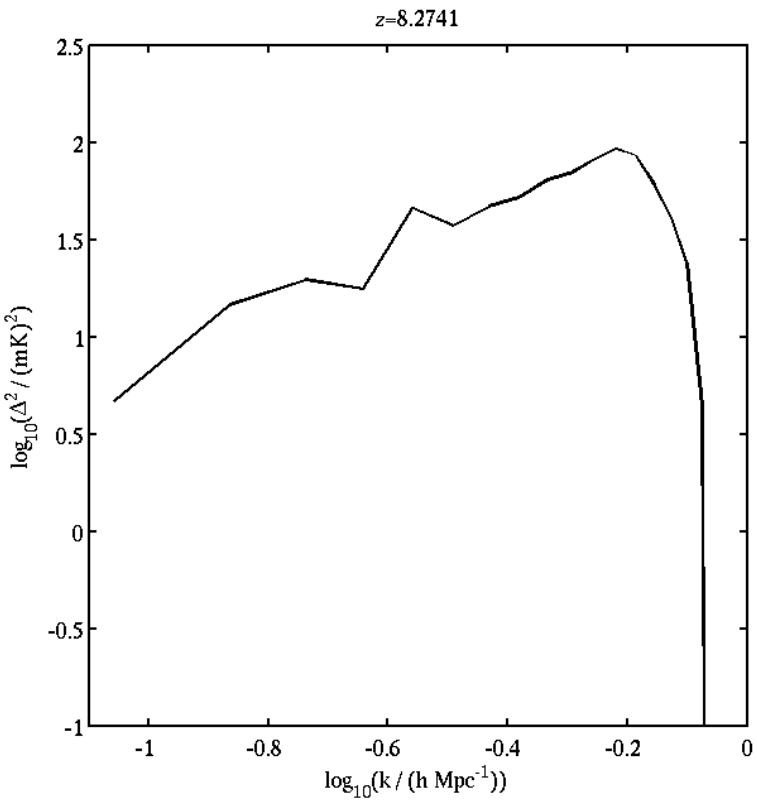
- Spectra here are convolved with the instrumental response (hence high-*k* cut-off).
- Growth of a feature on small scales due to formation of bubbles.
- Feature broadens and moves to larger scales as bubbles grow.
- At low redshift, signal drops because of a low neutral fraction.

Example evolution of the signal power spectrum



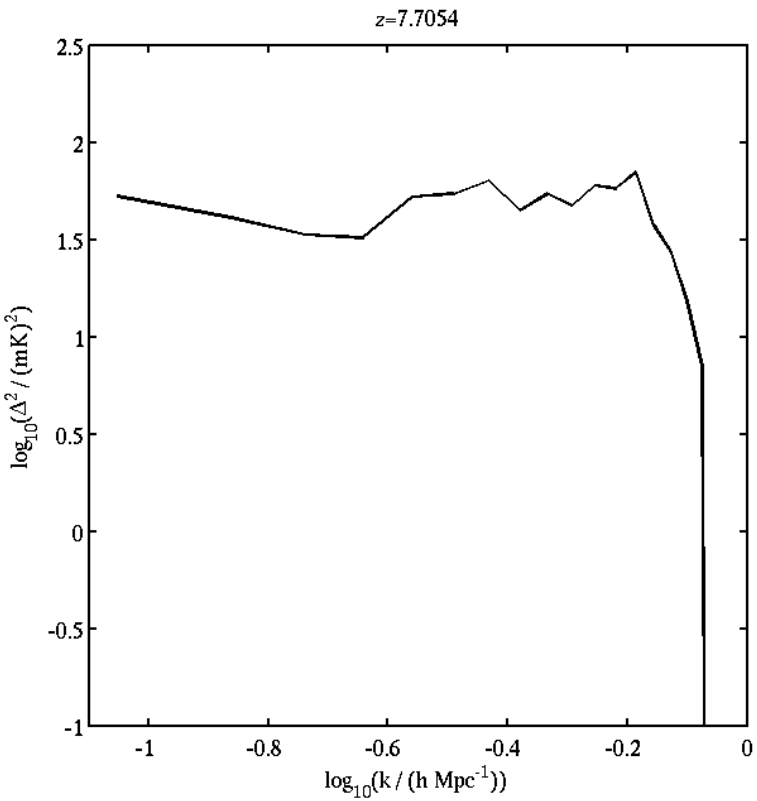
- Spectra here are convolved with the instrumental response (hence high-*k* cut-off).
- Growth of a feature on small scales due to formation of bubbles.
- Feature broadens and moves to larger scales as bubbles grow.
- At low redshift, signal drops because of a low neutral fraction.

Example evolution of the signal power spectrum



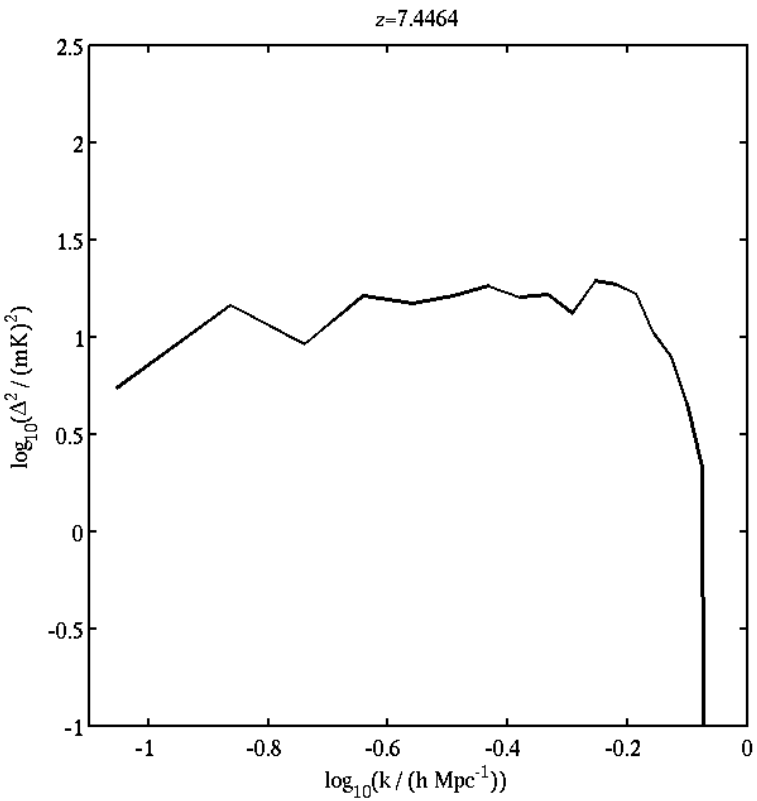
- Spectra here are convolved with the instrumental response (hence high-*k* cut-off).
- Growth of a feature on small scales due to formation of bubbles.
- Feature broadens and moves to larger scales as bubbles grow.
- At low redshift, signal drops because of a low neutral fraction.

Example evolution of the signal power spectrum



- Spectra here are convolved with the instrumental response (hence high-*k* cut-off).
- Growth of a feature on small scales due to formation of bubbles.
- Feature broadens and moves to larger scales as bubbles grow.
- At low redshift, signal drops because of a low neutral fraction.

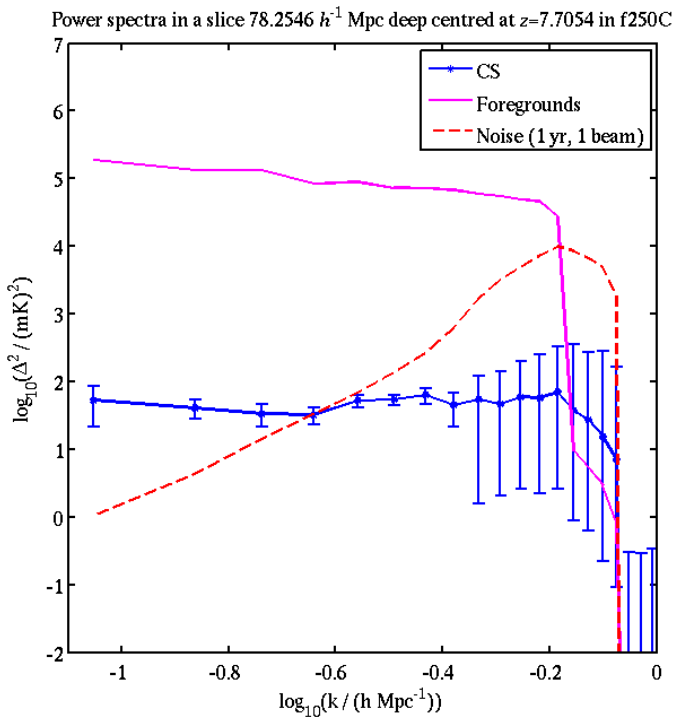
Example evolution of the signal power spectrum



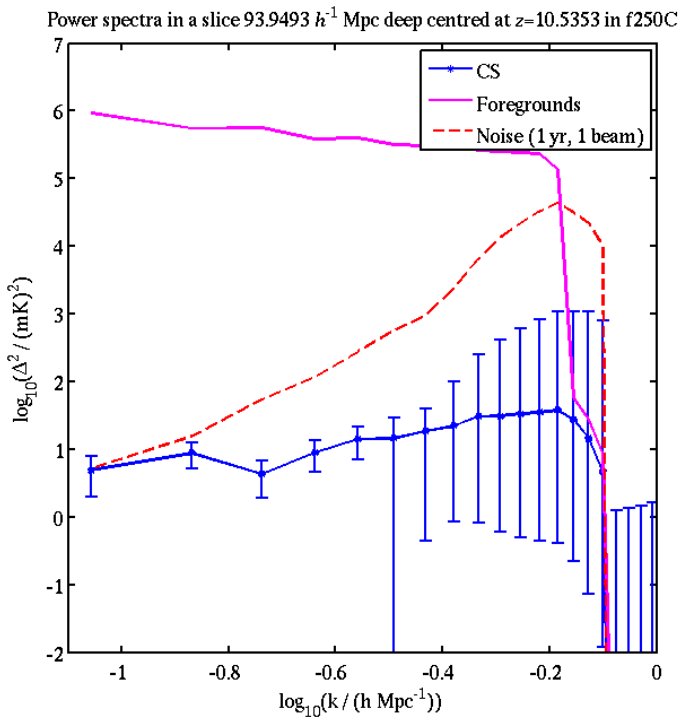
- Spectra here are convolved with the instrumental response (hence high-*k* cut-off).
- Growth of a feature on small scales due to formation of bubbles.
- Feature broadens and moves to larger scales as bubbles grow.
- At low redshift, signal drops because of a low neutral fraction.

Power spectra with perfect foreground subtraction (1 year, 1 beam)

Low redshift

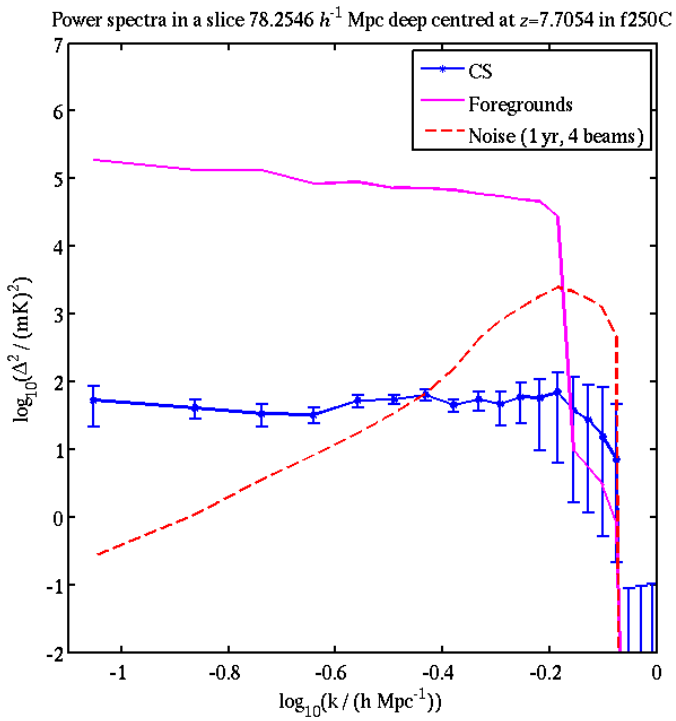


High redshift

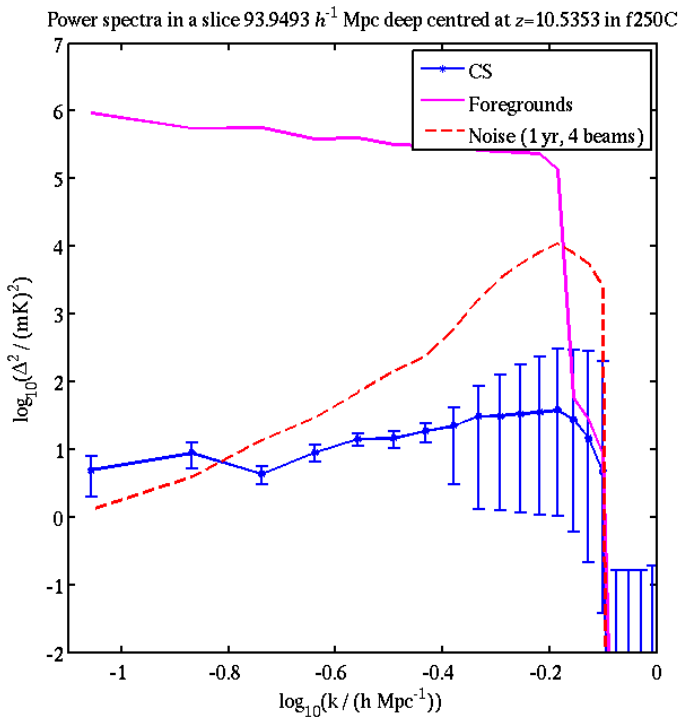


Power spectra with perfect foreground subtraction (1 year, 4 beams)

Low redshift

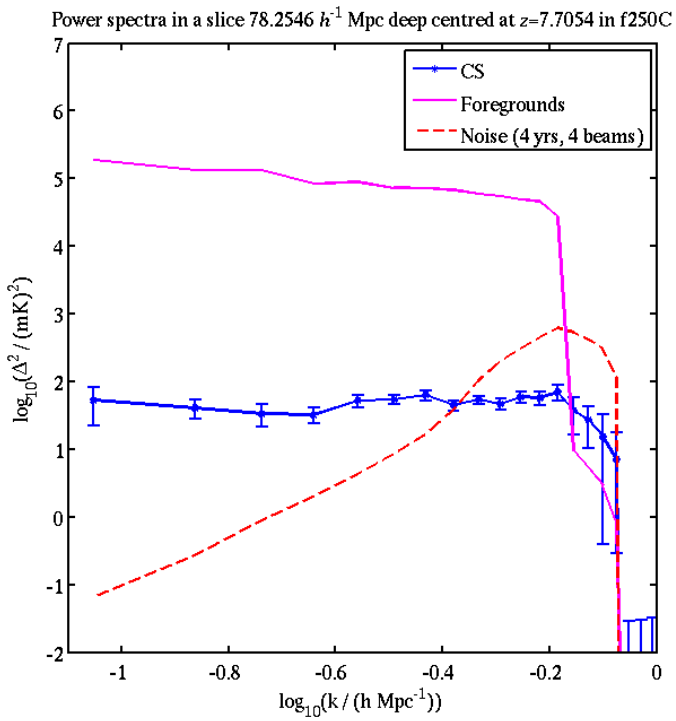


High redshift

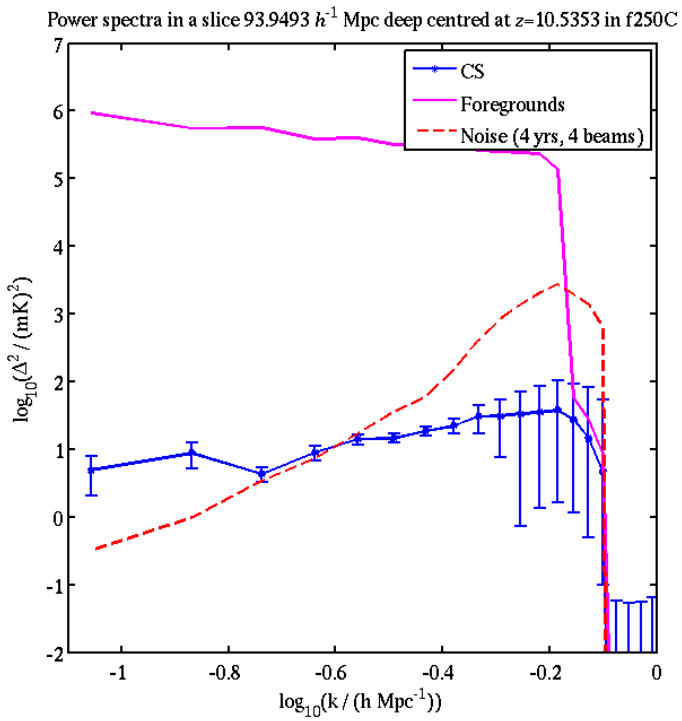


Power spectra with perfect foreground subtraction (4 years, 4 beams)

Low redshift

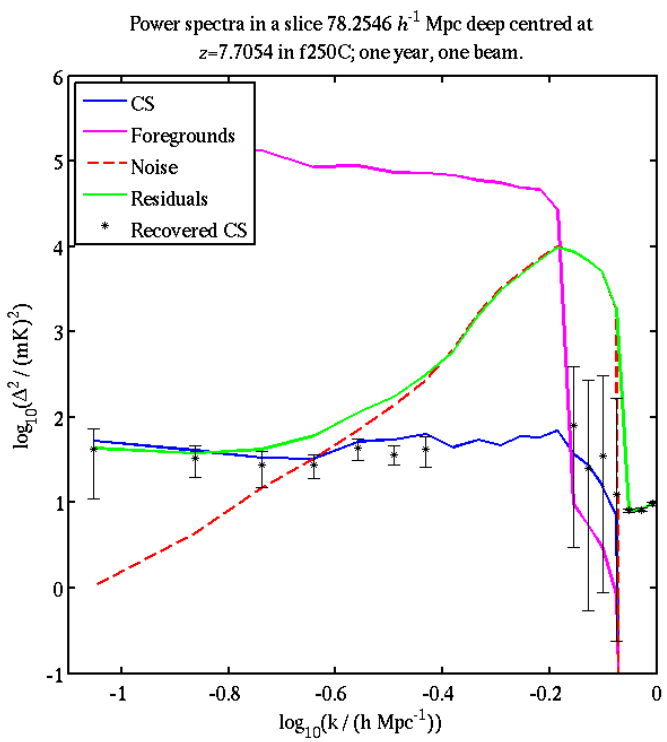


High redshift

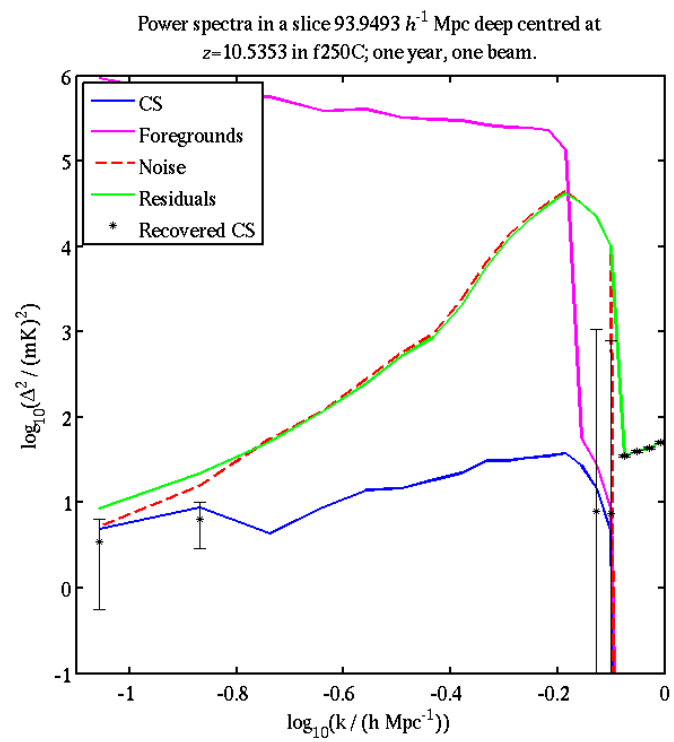


Results using Wp smoothing for foreground subtraction (1 yr, 1 beam)

Low redshift

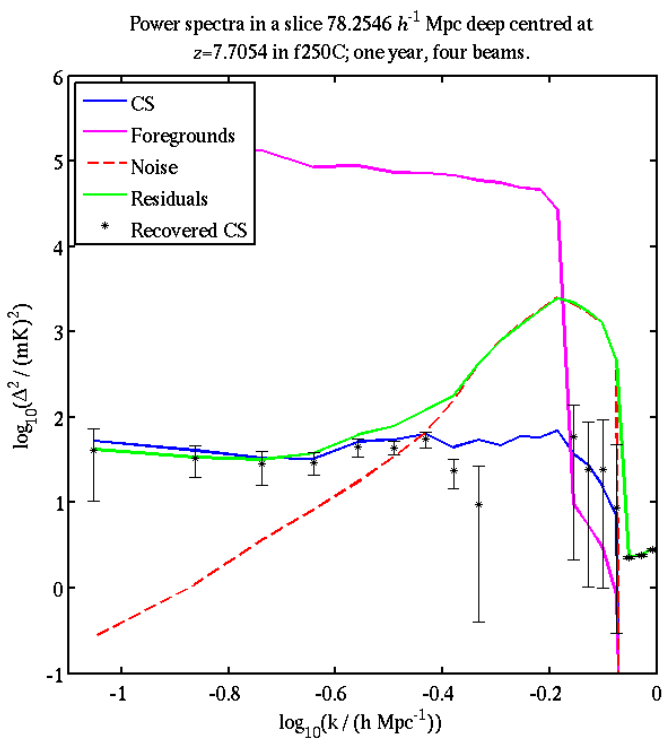


High redshift

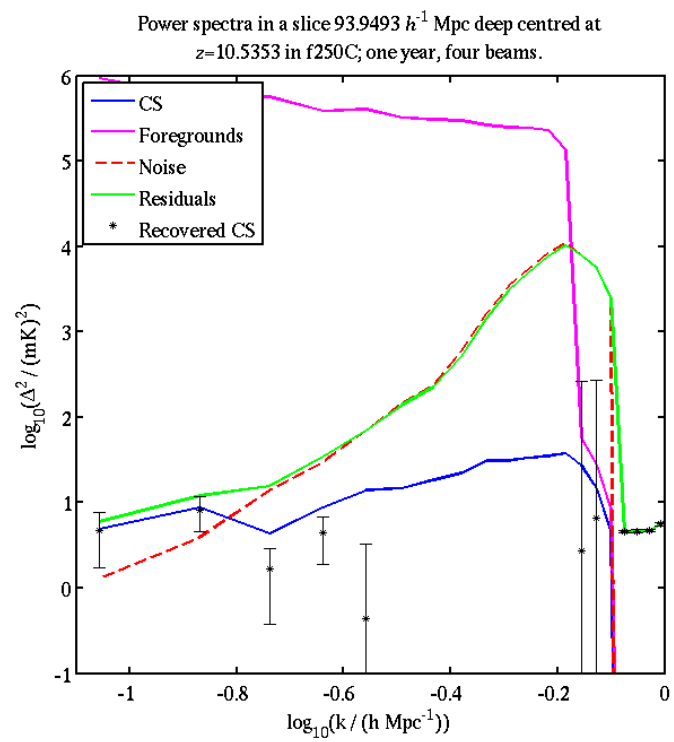


Results using Wp smoothing for foreground subtraction (1 yr, 4 beams)

Low redshift

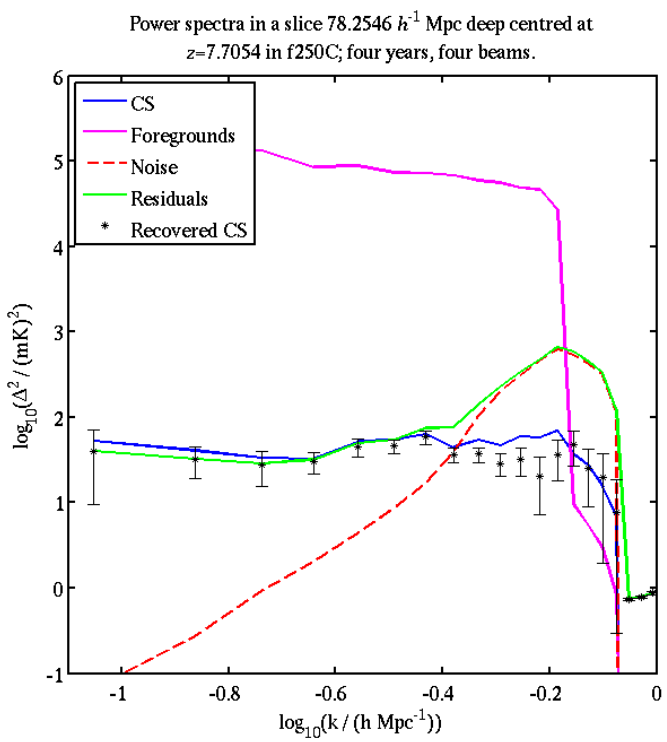


High redshift

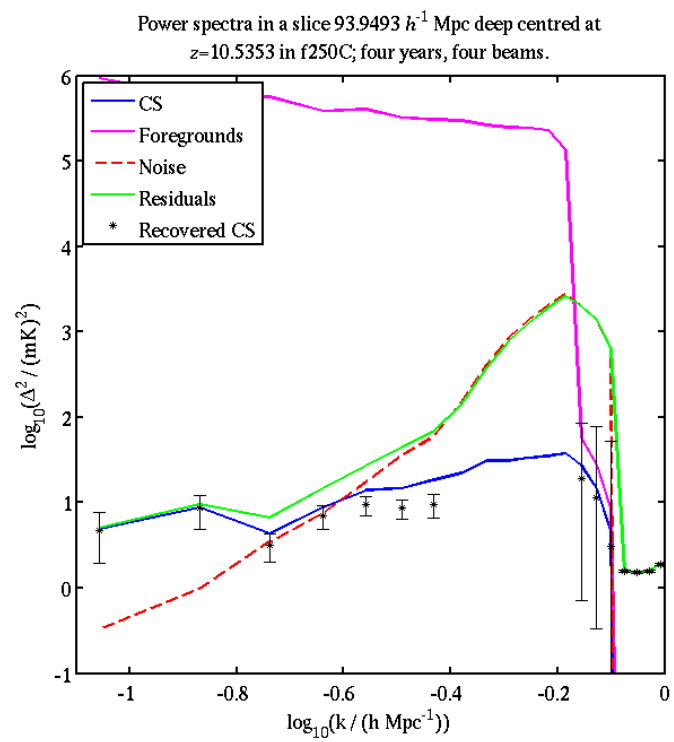


Results using Wp smoothing for foreground subtraction (4 yr, 4 beams)

Low redshift

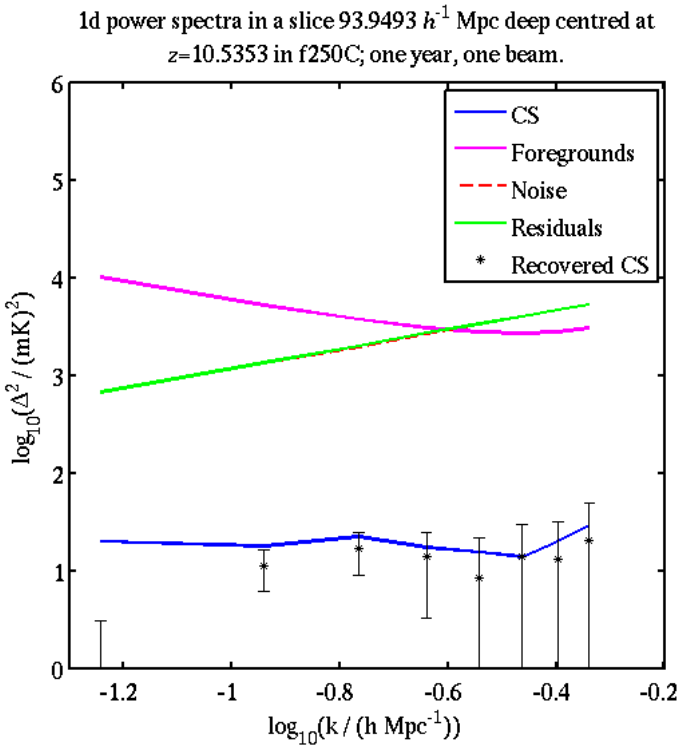


High redshift

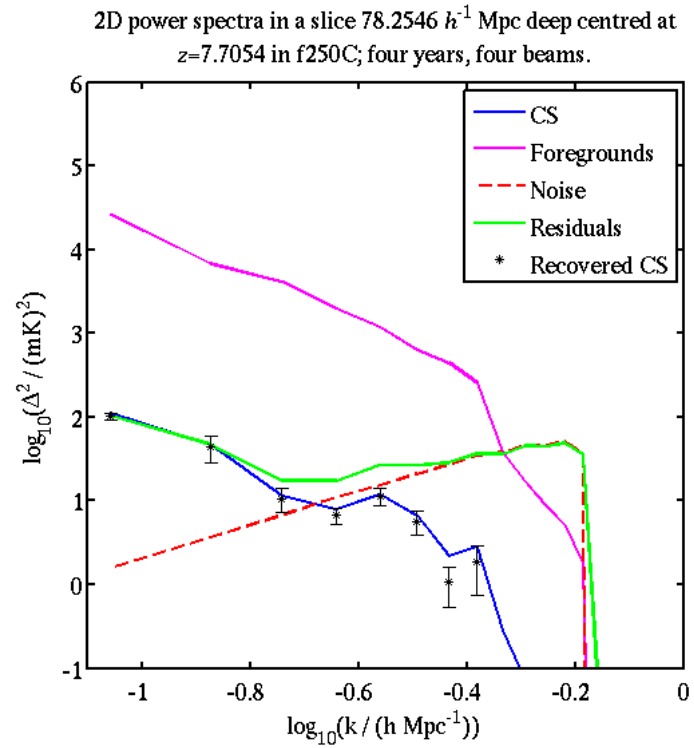


1D and 2D power spectra

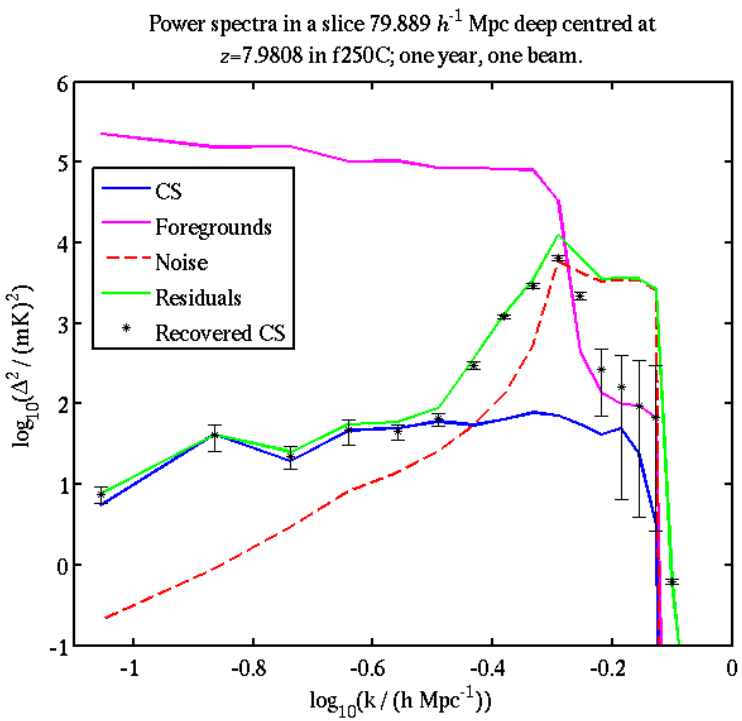
1D power spectrum – high *z*, 1 yr, 1 beam



2D power spectrum – low z, 4 yrs, 4 beams



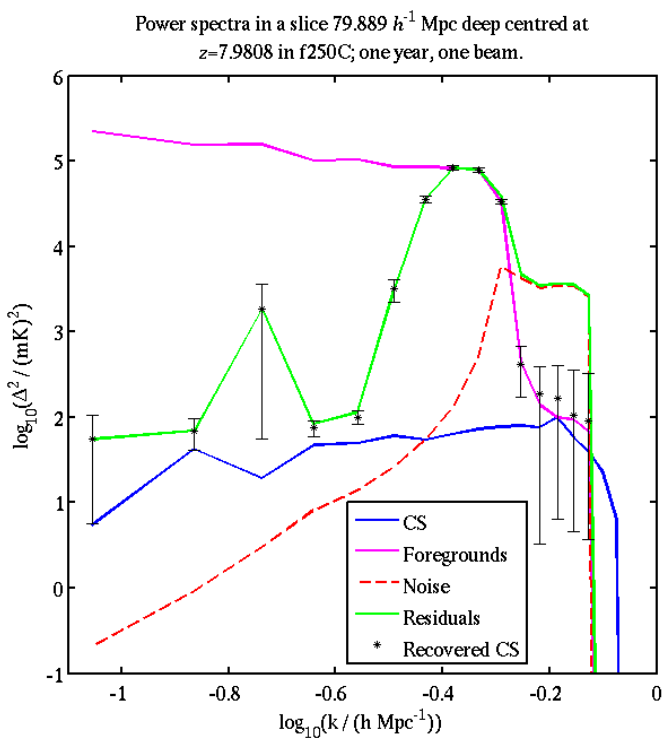
Redshift-dependent *uv* coverage



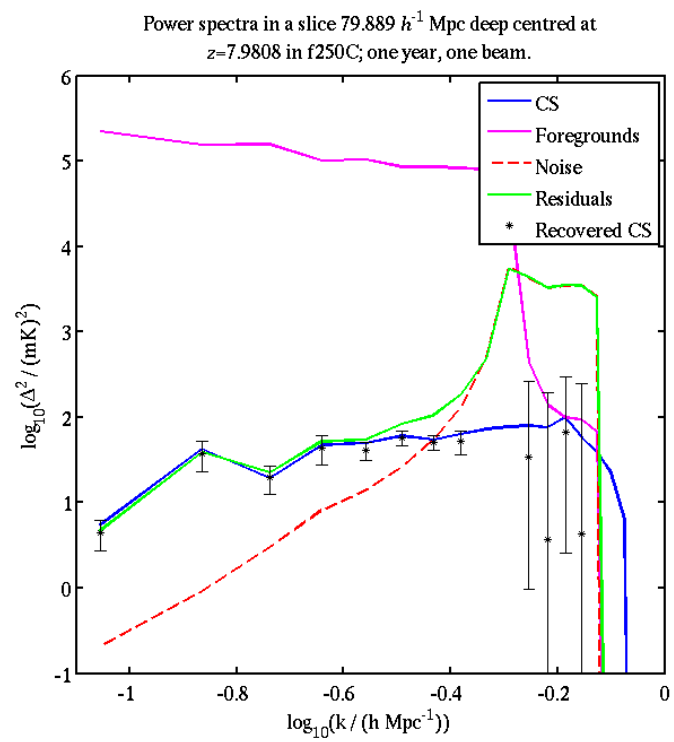
- The *uv* coverage changes as a function of redshift: high v increases the maximum *k*, moves holes across the *uv* plane and contracts 'frizz' from the PSF across unresolved point sources in the image plane.
- The latter effect introduces spurious small-scale power if we fit the foregrounds in the image plane.
- The most drastic solution is to throw away data until the *uv* coverage is the same at all frequencies.

Fitting in the *uv* plane

Starting from images



Starting from Fourier space



Future work

- How accurately can we really estimate the power spectrum of the noise? Probably need to use the data before they're binned in time and frequency to give a final data cube.
- Can P(k,μ) be extracted? How is the 'separation of powers' (of μ) affected by the fitting and extraction process?
- Continue to integrate improved models of the instrument, foregrounds, noise and a variety of signal models incorporating larger scales.
- Effect of other error sources: ionosphere, polarization calibration, point source subtraction errors…
- Incorporate signal correlations and a systematic way of choosing the amount of smoothing in the foreground fitting.
- Can we gain from mismatched spatial and frequency resolution?
- Effect of power spectrum errors on recovery of evolution of skewness.
- Full error analysis including cosmic variance; effect of multiple windows.

Comments and conclusions

- 21cm emission from high redshift promises to be a rich source of physical information…
- …but many challenges are posed by foregrounds, noise, the ionosphere, instrumental effects, RFI, etc.
- For the LOFAR EoR project we are developing a pipeline to simulate all aspects of the measured signal, so that our signal extraction and analysis techniques can be tested.
- If suitable statistics are used, the levels of foregrounds and noise in themselves do not constitute a deal-breaker for e.g. power spectrum subtraction…
- …but the fitting process introduces biases: can they be corrected for using simulation results?
- As the integration time increases, we can expect continued qualitative improvements in what can be inferred from the data.
- The effects of variable *uv* coverage seem to be manageable with a carefully chosen fitting scheme.