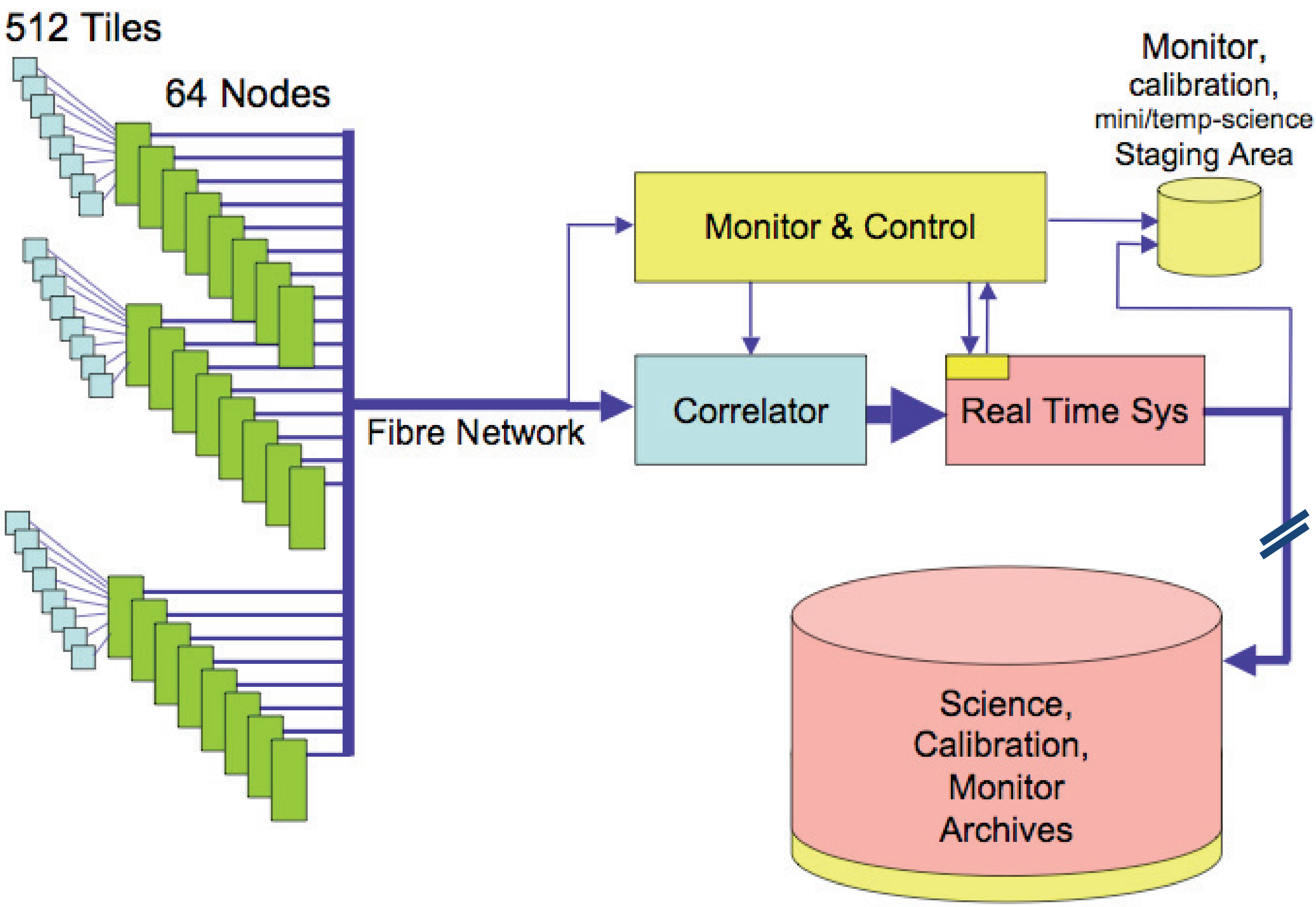
MWA:
OFF-LINE FOREGROUND SUBTRACTION AND PROCESSING

Judd D. Bowman (Caltech)
for the MWA-EOR Collaboration

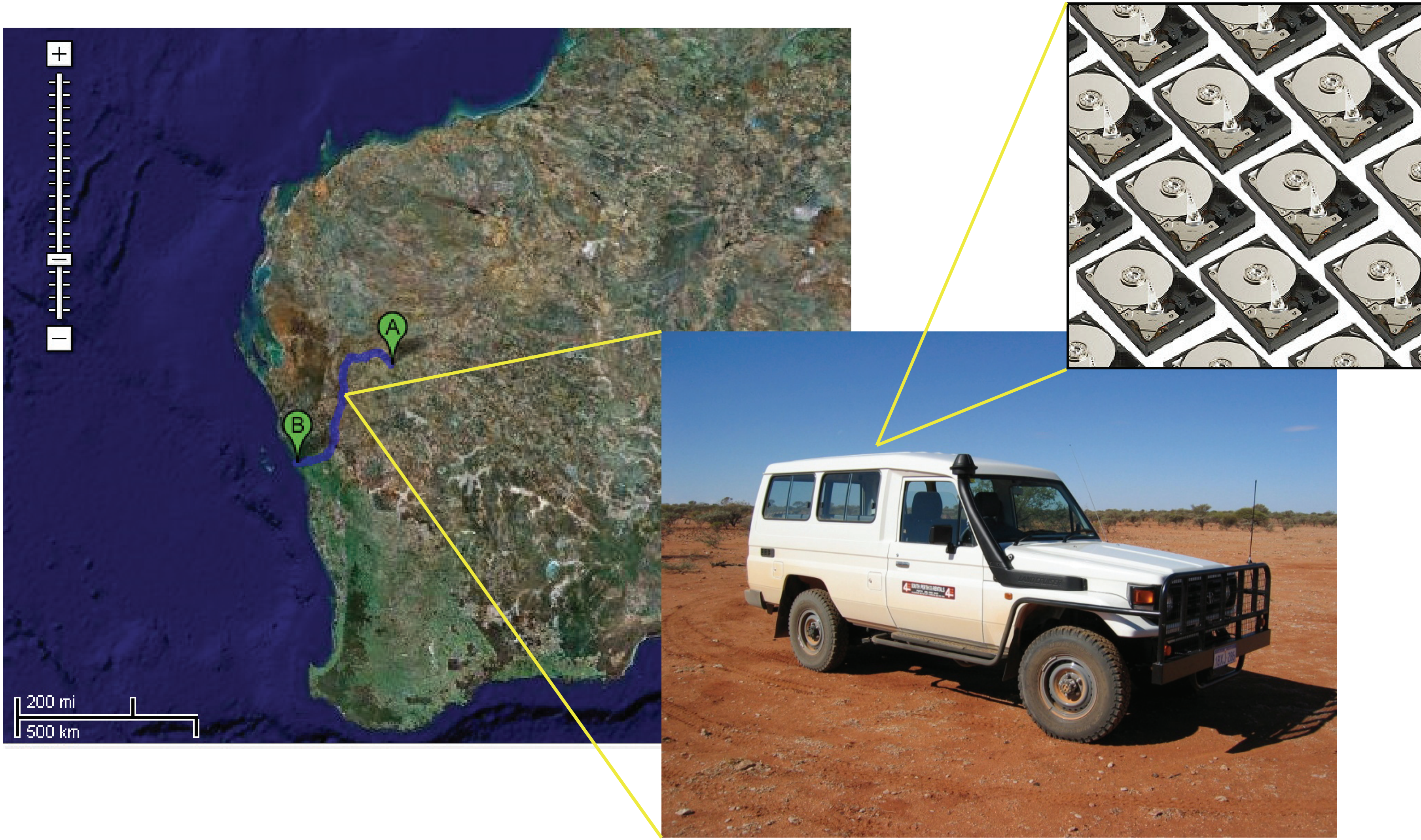
August 10, 2008

MWA: system diagram

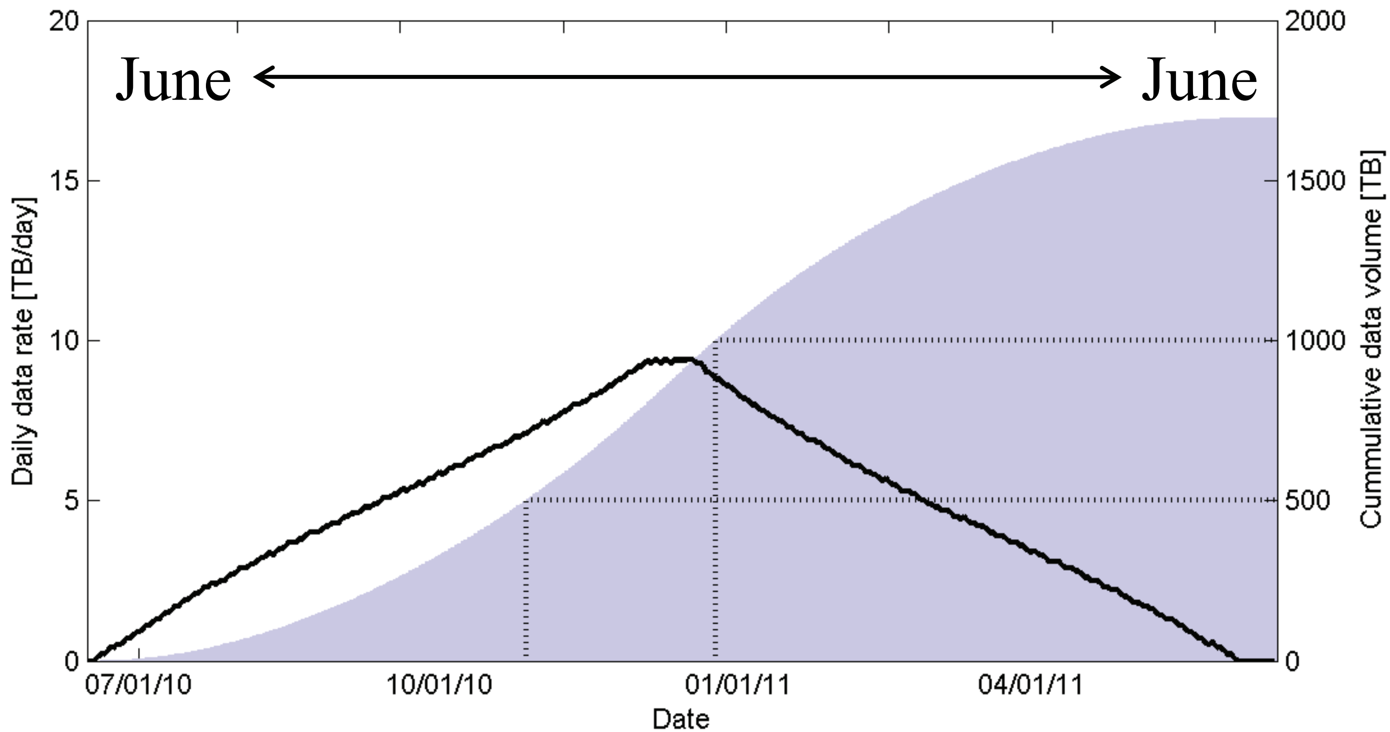


F. Briggs

Off-line… and off-road?!



Estimated EoR data rates



- Maximum data rate ~10 TB/night
- Up to ~1700 hours/year ≈ 1700 TB/year

EoR data archive

- On-site:
  - Integrate in RTS and store in ~10 minute increments
  - ~1 TB/hour or up to 10 TB/night
  - Small repository on site (<100 TB)

- Off-site:
  - Primary archive of ~500 to 1000 TB
  - Cluster of ~50 x 3U computers each with several CPU cores and ~16 1 TB disks
  - Database of primary-beam maps and metadata
  - Reduce off-line data by additional factor of 2-3 to fit full observing year in archive

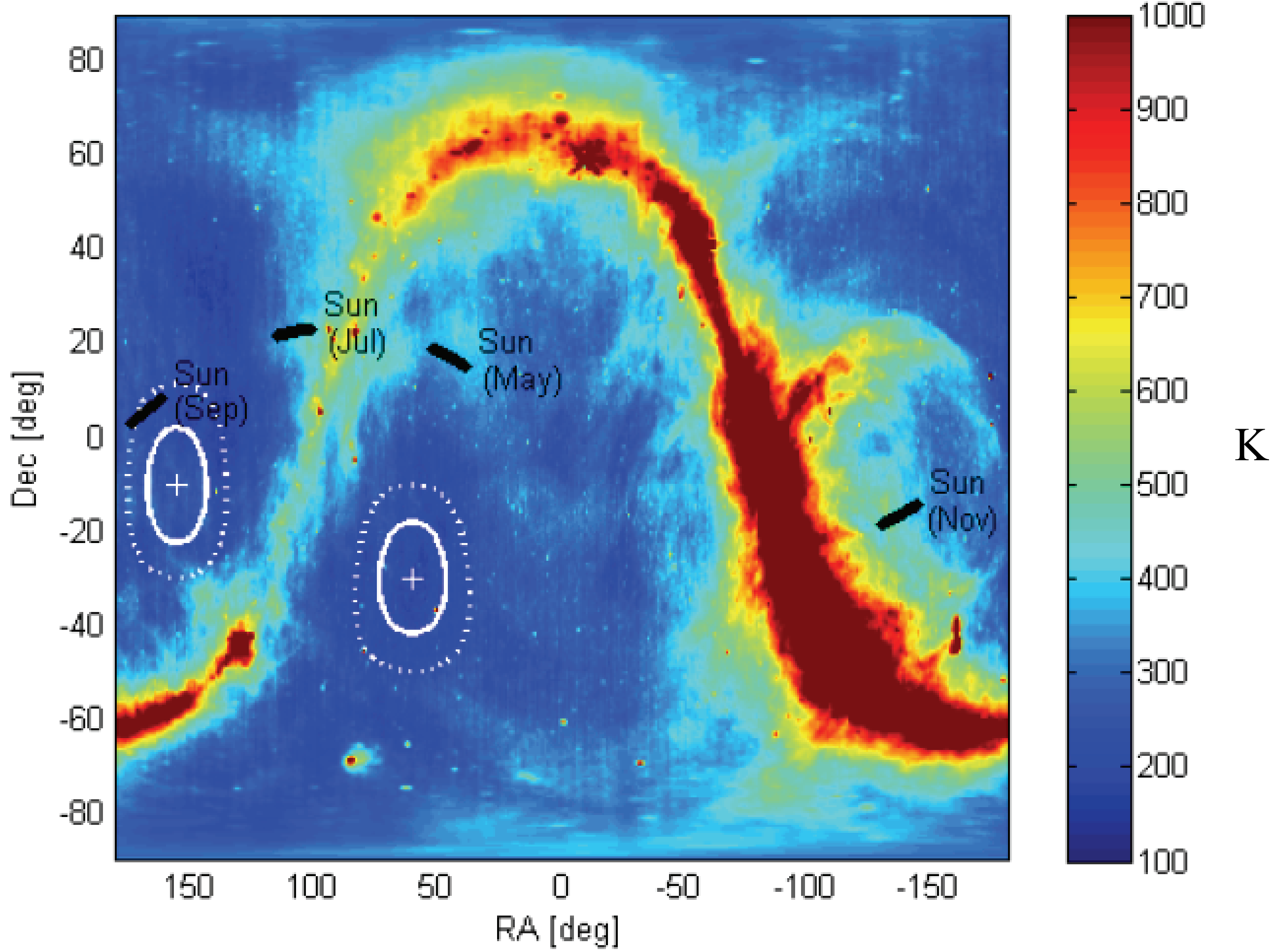
E. Morgan

EoR observations

Primary field:
RA 60.00, Dec -30.00
~1250 hours shared
between 2 bands
(64 MHz) → $6 < z < 9$

Secondary field:
RA 155.00, Dec -10.00,
~450 hours in 1 band
32 MHz → $6 < z < 7$

Data cubes:
30 x 30 deg x 32 MHz
few arcmin x 40 kHz res
4 Stokes
~1000 x 1000 x 768 cells
~10 GB/cube x 4 Stokes

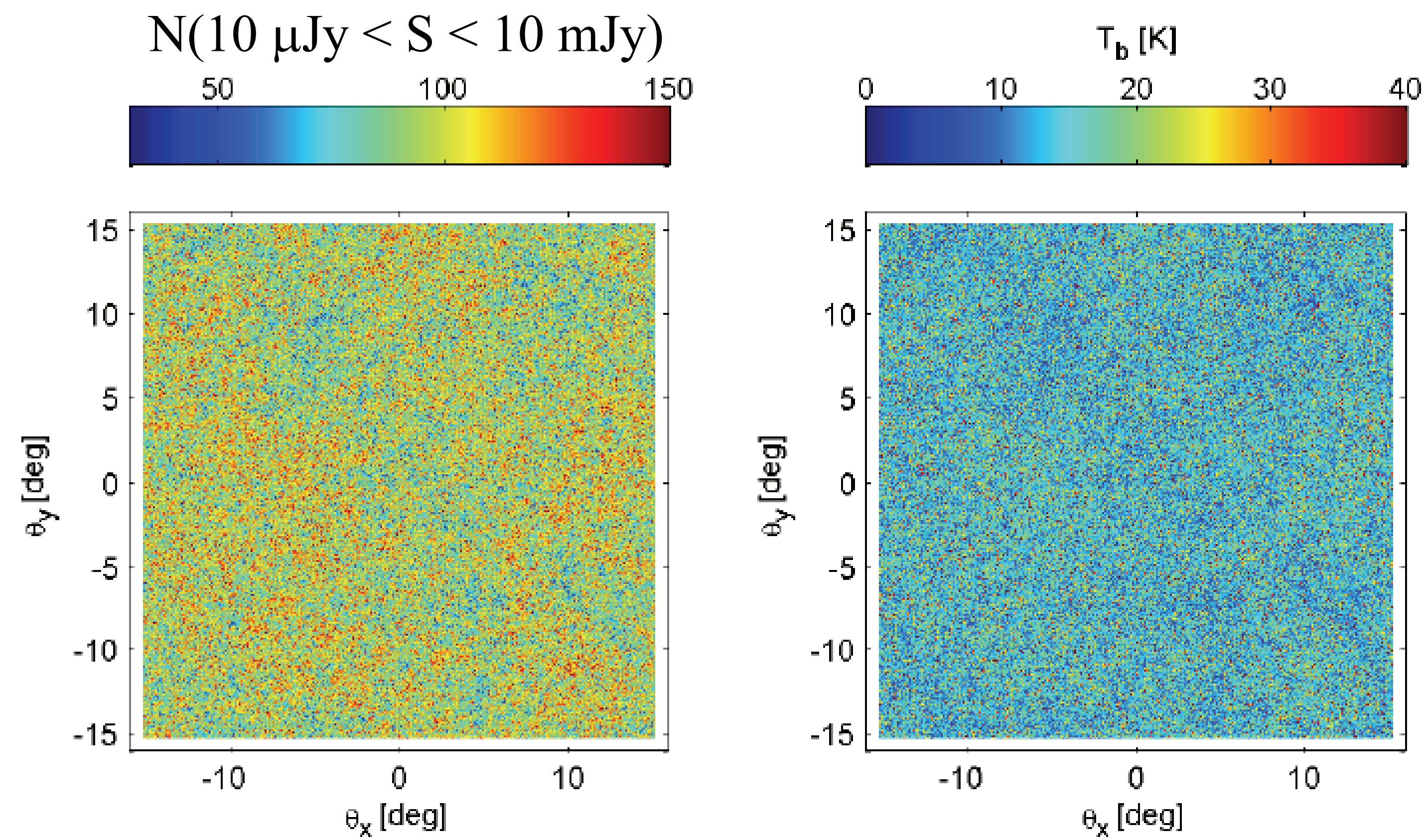


Off-line foreground subtraction: 6-step strategy (to discard 99.99%)

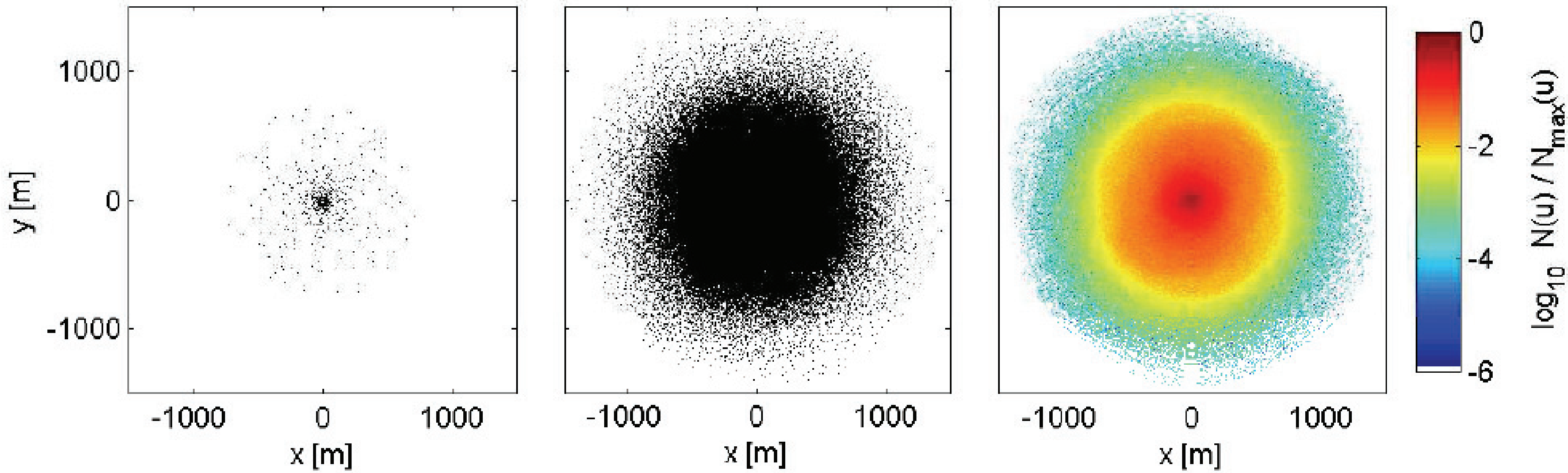
1. Deconvolution/peeling/nulling of bright sources
2. Out-of-beam sky subtraction
3. Polarized leakage
4. Confusion level sources and Galactic emission
5. 3D power spectrum & template fitting
6. Data cross-checks

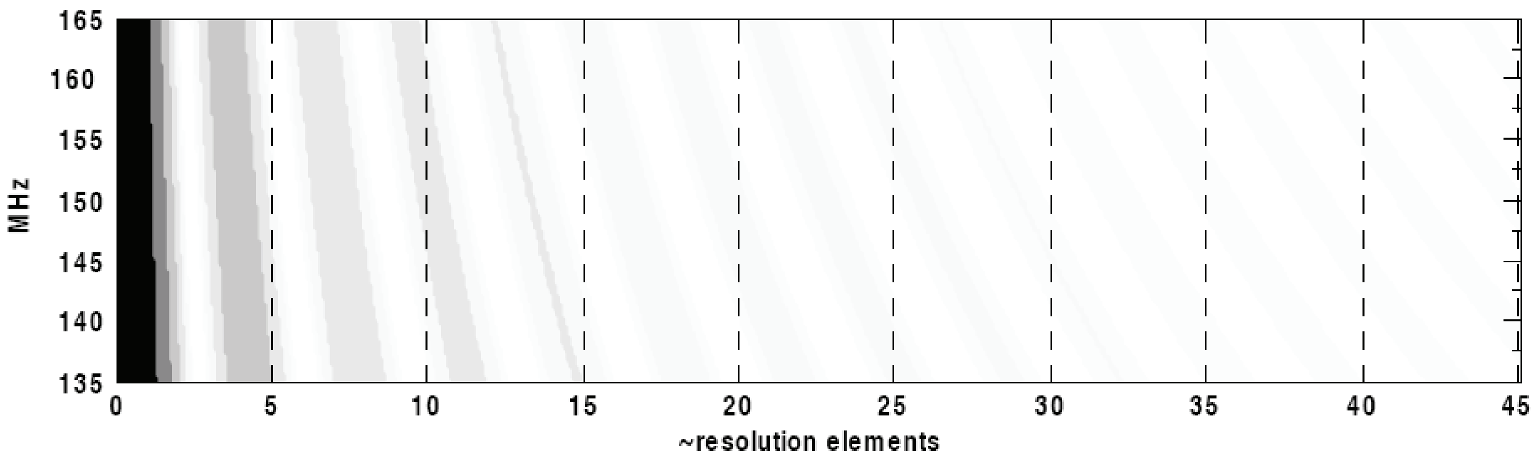
Morales, Bowman, & Hewitt (2006)

4. Confusion-level sources



4. UV coverage





Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

4. Polynomial subtraction

- Create sky model and apply instrument response in uv-plane

$$I'(\mathbf{u}) = \mathbf{B}(\mathbf{u},\mathbf{u})\ I(\mathbf{u}) + n(\mathbf{u})$$

- Calculate dirty map using a weighting function

$$I_U(\theta) = \mathbf{F}(\theta,\mathbf{u})\ \mathbf{U}(\mathbf{u},\mathbf{u})\ I'(\mathbf{u})$$

- Fit 2nd-order polynomial along each pixel (32 MHz) in dirty cube and subtract

$$d\left(\vec{\theta},\nu\right) = a_2(\vec{\theta})\nu^2 + a_1(\vec{\theta})\nu + a_0(\vec{\theta})$$

$$r_U(\theta) = I_U(\theta) - d(\theta)$$

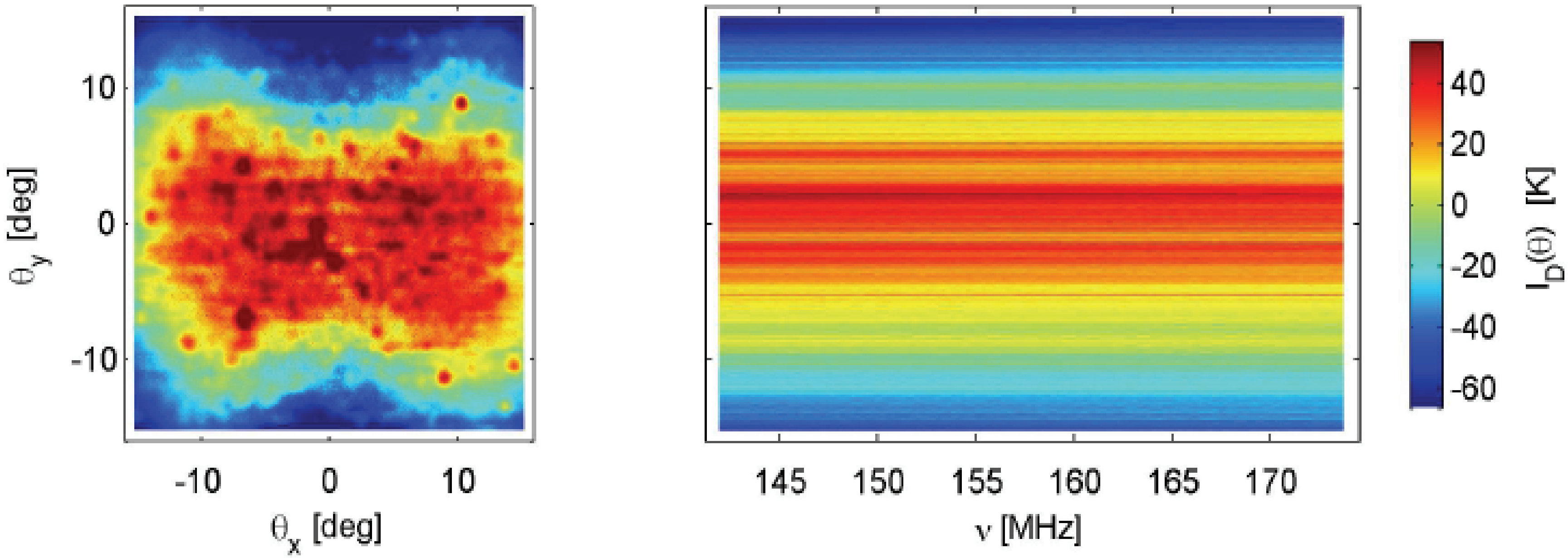
- Remove effects of weighting function

$$r'(\theta) \;=\; I'(\theta) \;-\; d'(\theta)$$

$$d'(\theta) = \mathbf{F}^{-1}(\theta,\mathbf{u})\mathbf{U}^{-1}(\mathbf{u},\mathbf{u})\ \mathbf{F}(\mathbf{u},\theta)\ d(\theta)$$

Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

4. Simulated dirty maps



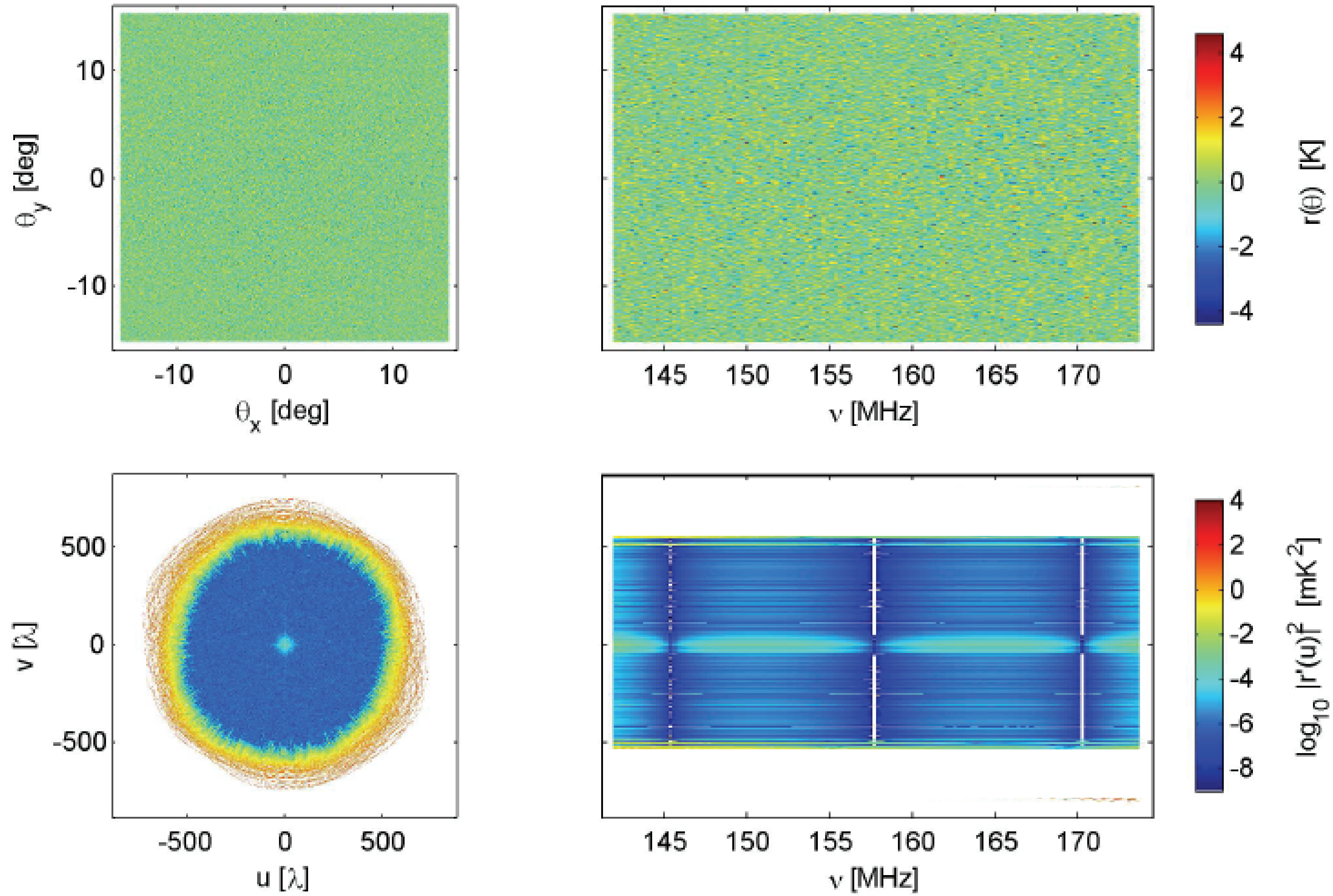
- No ionosphere and no calibration errors
- Bright sources assumed to be removed perfectly (S>10 mJy)

Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

4. Subtraction residuals

residuals in dirty image cube*

residuals in unweighted uvf cube*



* Thermal noise and 21 cm removed artificially for clarity

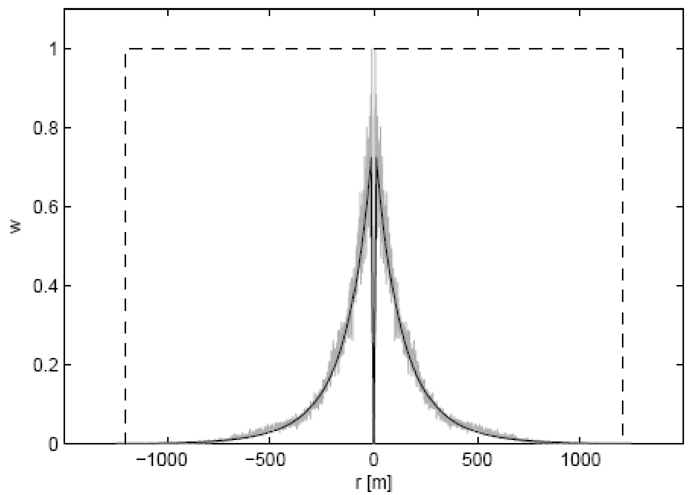
Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

How does confusion-level subtraction depend on…

- Weighting function
- Source peeling cutoff
- Rotation synthesis
- Polynomial fit order
- Tile arrangement
- Other beam manipulations

4. Weighting

weight

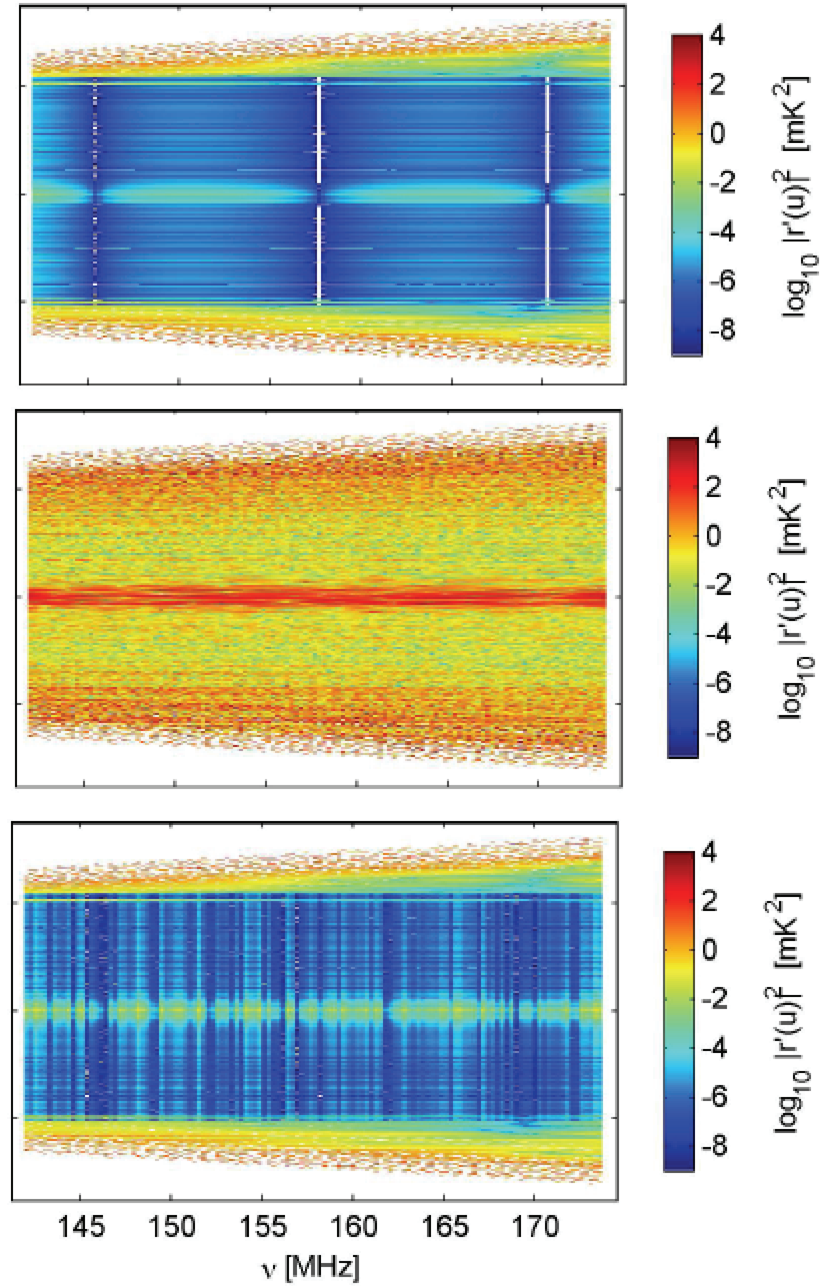


radius [m]

uniform

natural

smoothed natural



- Natural weighting couples structure in uv-distribution to spectral domain
- Tolerable uv-structure:

$$\frac{l_u}{|u|} = \frac{l_\nu}{\nu_0} \qquad l_u \gtrsim |u| \frac{B}{\nu_0} \approx \frac{|u|}{10}$$

Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

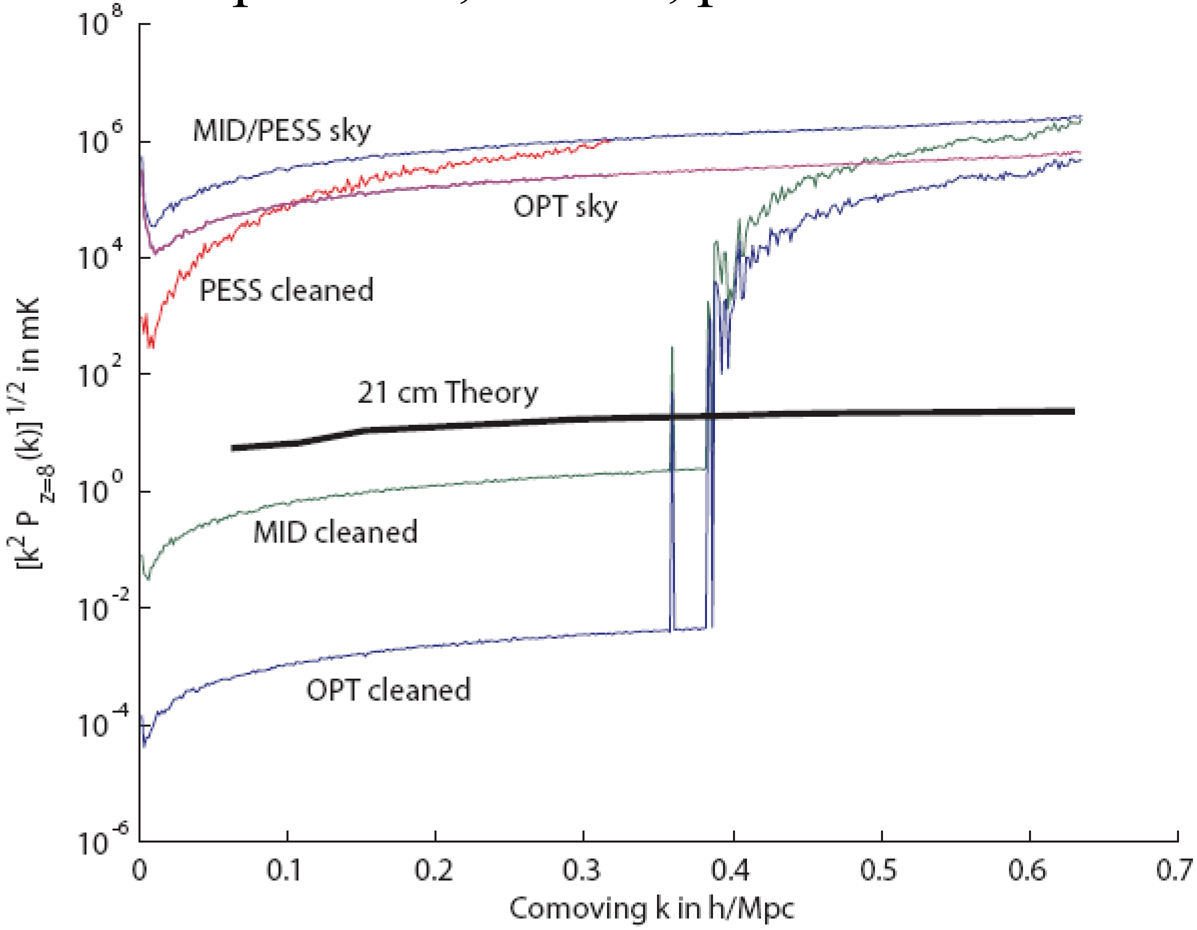
4. Parameter exploration

Assumptions		Low-Performance Extreme	Fiducial Model	High-Performance Extreme
Experimental	Tile Arrangement	$\rho(r) \sim r^{-2}$	$\rho(r) \sim r^{-2}$	Monolithic with tiles separated by 40 m
	Rotation synthesis	None	6 hours, continuous	6 hours, continuous
	Noise level	$\sigma_T \sim 1\,\mathrm{mK}$	Noiseless [5]	Noiseless
Analysis	Primary beam width adjustments	None	None	Adjusted to be frequency-independent
	Bright point source flux cut S_{cut}	100 mJy	10 mJy	0.1 mJy
	Synthesized beam width adjustments	None	None	Resolutions equalized by extra smoothing
	u-v plane weighting	None (natural)	Uniform	Uniform
	Order of polynomial fit	Constant	Quadratic	Quintic
	Range of polynomial fit	80 MHz	2.4 MHz	2.4 MHz

Liu, Tegmark & Zaldarriaga (2008, submitted)

4. Parameter exploration

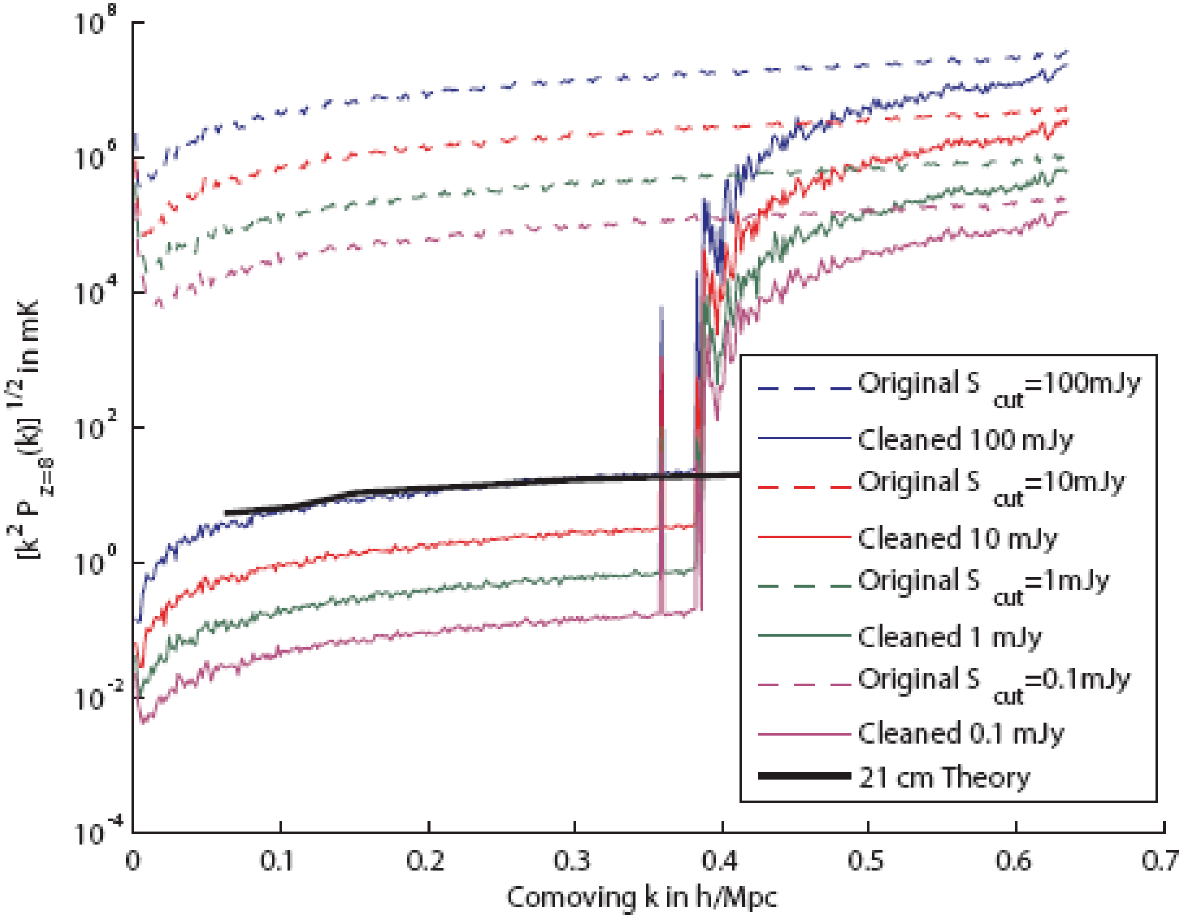
Optimistic, middle, pessimistic case



Liu, Tegmark & Zaldarriaga (2008, submitted)

4. Parameter exploration

Bright source cleaning level



Liu, Tegmark & Zaldarriaga (2008, submitted)

5. Power spectrum

- Convert from gridded visibility data cube to cosmological k-space [Mpc^{-1}]

$$r'(\mathbf{k}) = \mathbf{J}(\mathbf{k}, \{\mathbf{u}, \eta\})\mathbf{F}(\{\mathbf{u}, \eta\}, \{\mathbf{u}, \nu\})r'(\{\mathbf{u}, \nu\})$$

- Want to constrain contributions to power spectrum with different symmetries

$$P_{21}\left(\vec{k}\right) = P_{\mu^0}(k) + \mu^2 P_{\mu^2}(k) + \mu^4 P_{\mu^4}(k) + \ldots,$$

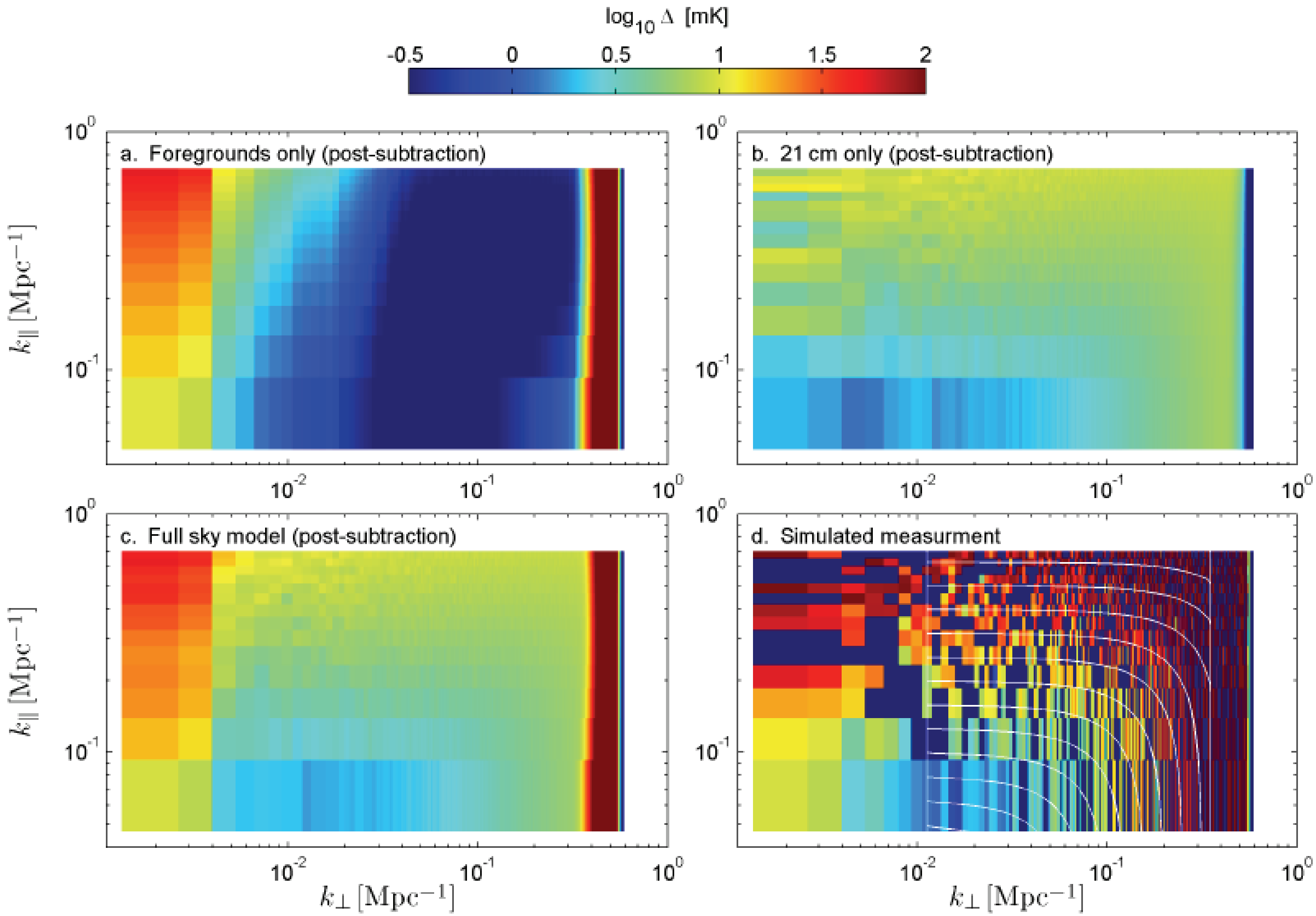
$$\mu \equiv \hat{k} \cdot \hat{n}$$

- Can reduce to 2 dimensions without loss of generality

$$P_{21}(k_{\perp}, k_{||})$$

- Or to 1 dimension (assuming spherical symmetry) to increase SNR

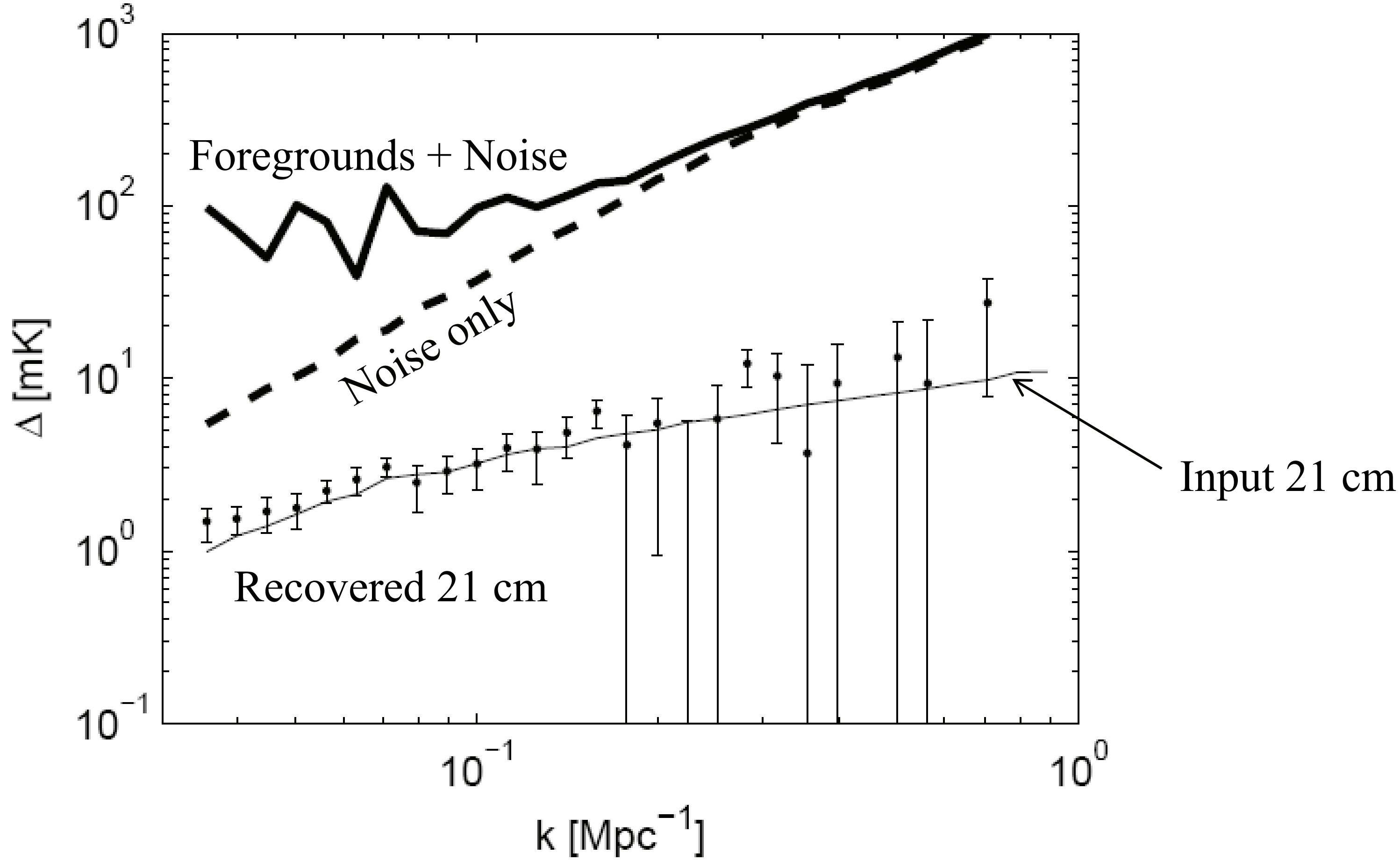
5. Power spectrum 2D



Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

5. Power spectrum 1D

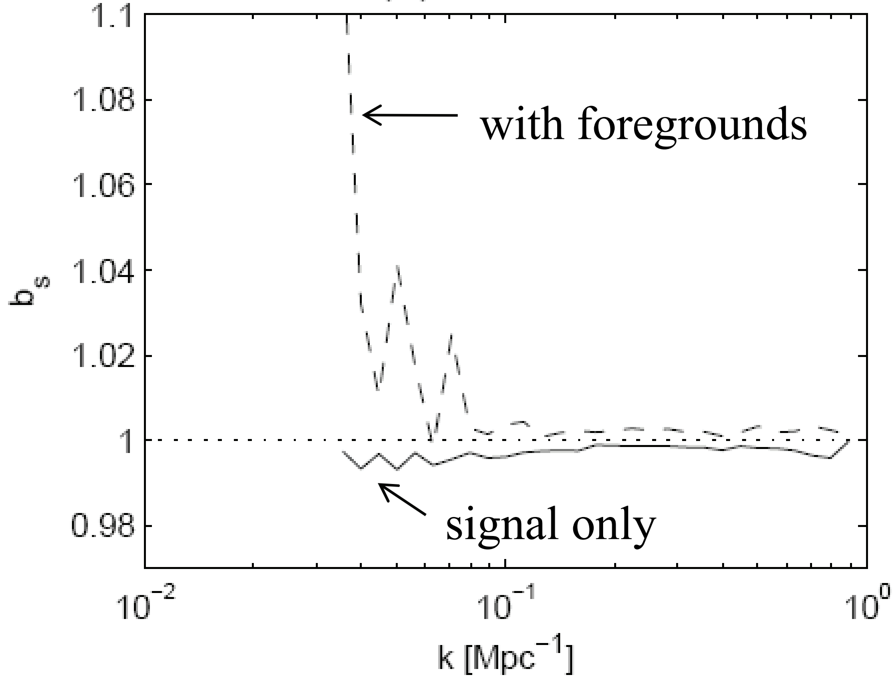
B = 8 MHz   t = 360 hours



Bowman, Morales & Hewitt 2008 (arXiv:0807.3956)

5. PS 1D residuals

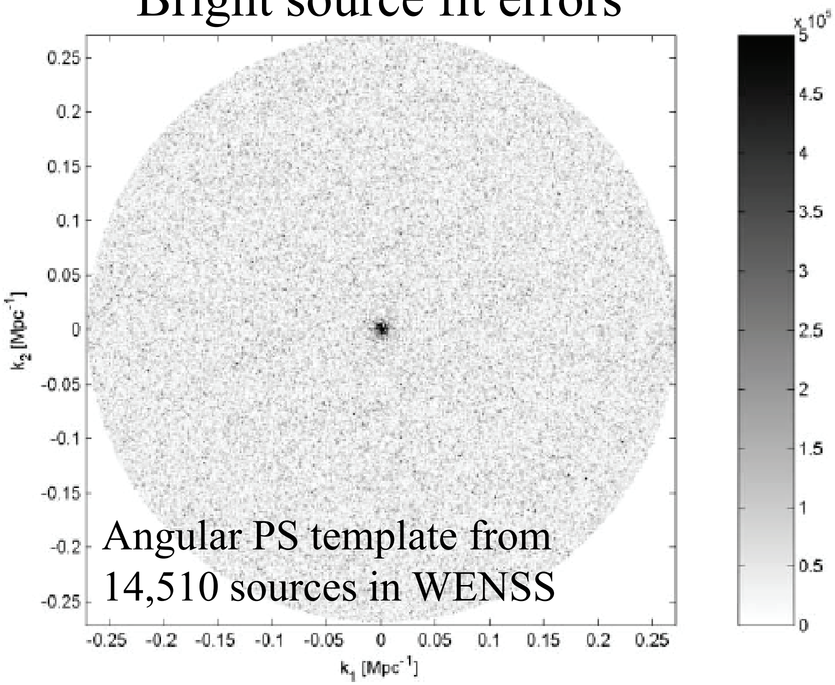
$$P'_{21}\left(\vec{k}\right) = b_s^2(\vec{k})P_{21}(\vec{k})$$



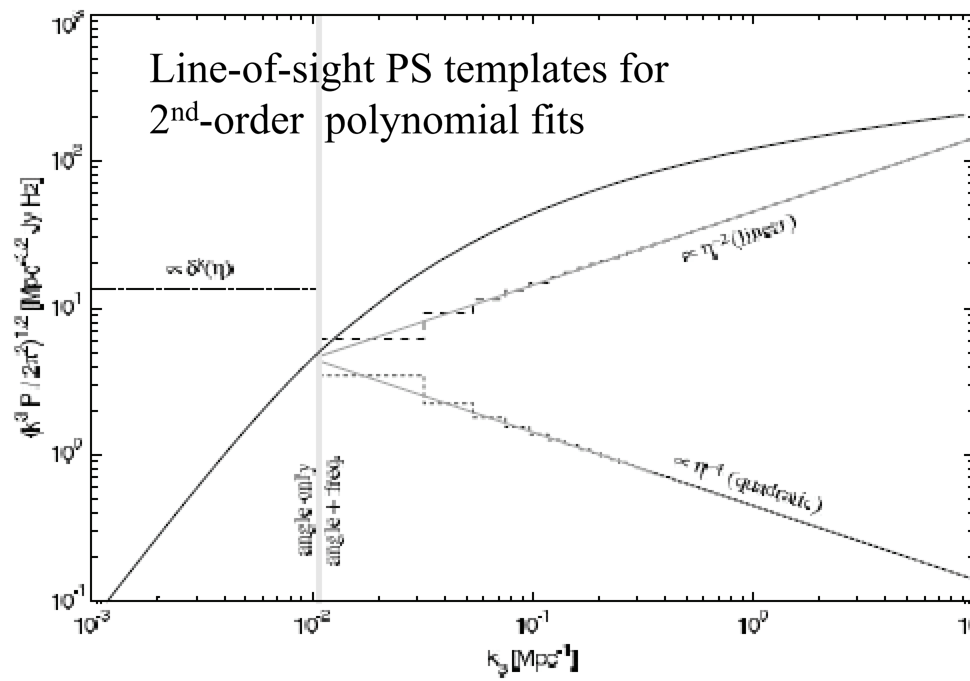
5. PS template fitting

- Each foreground subtraction step will introduce residual errors
- Residuals that cannot be predicted for individual cases (such as mis-fit of polynomial along any given sight-line), can be identified and removed using templates in the power spectrum
- Example templates due to:

Bright source fit errors



Confusion-level fit errors



Morales, Bowman, & Hewitt (2006)

Summary

- MWA-EOR data archive will store ~1 year of "raw" observations
  - Data volume reduced both on-line (by factor of ~100) and off-line (by factor of ~few) through integration and cuts

- Sobering 6-step foreground subtraction plan
  - Framework for confusion-level subtraction and PS template fitting in hand
  - Development of bright source removal and polarization leakage removal underway, extensions of existing techniques (RM synthesis, peeling, RTS)
  - Build library of PS templates as we simulate additional steps of foreground subtraction
  - End-to-end simulations planned, along with tests on 32T and early observations