HI intensity mapping with MeerKAT and the SKA

Mário G. Santos, University of the Western Cape

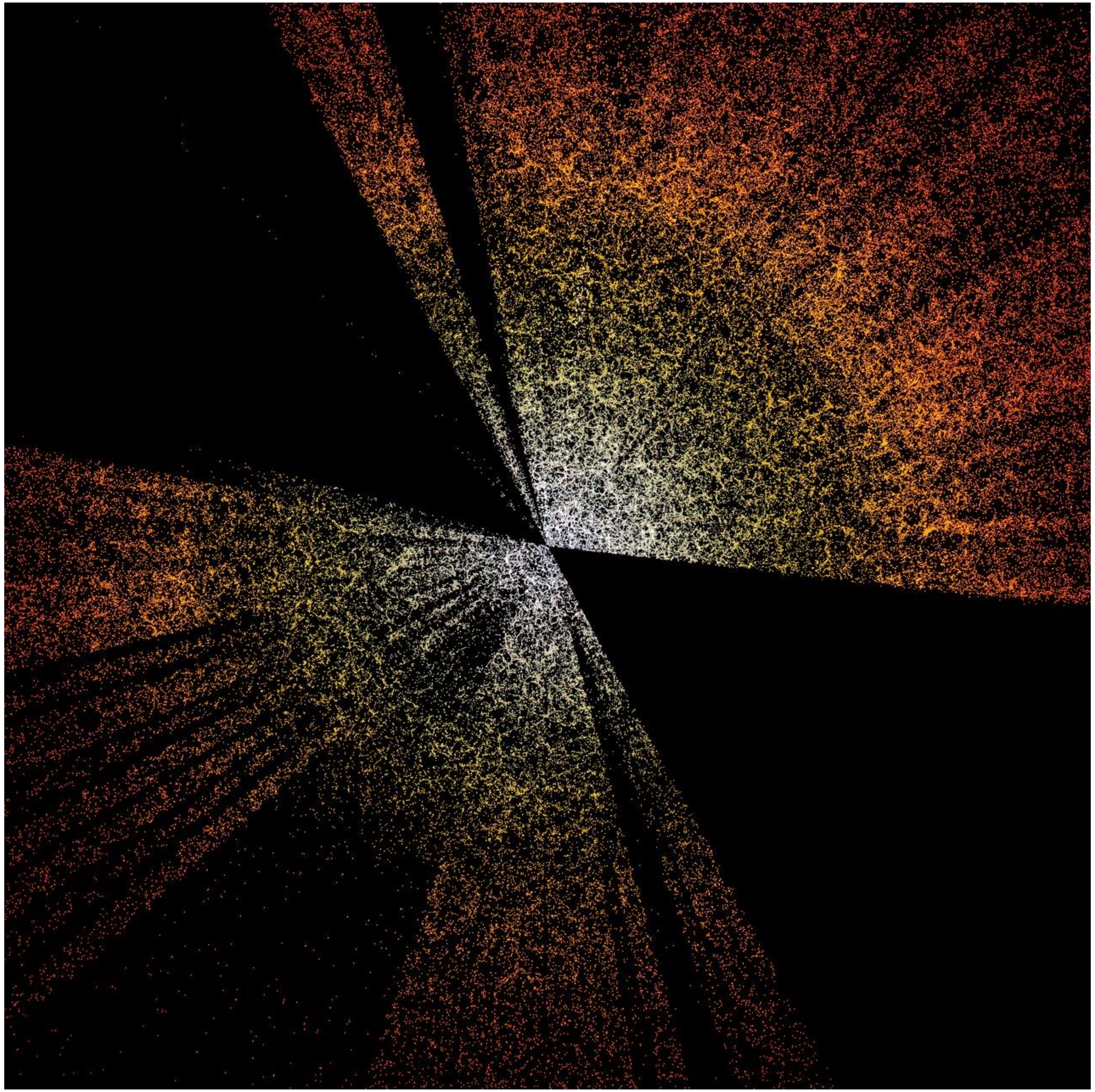
DSU, Kigali

July 10, 2023

CRC
UNIVERSITY of the WESTERN CAPE
UWC ASTROPHYSICS
NRF National Research Foundation
SARAO South African Radio Astronomy Observatory

The future is 4D!

- Current questions: dark energy? dark matter? modified gravity? Primordial fluctuations?
- We need 3D measurements of large scale structure across time to probe fundamental questions in Cosmology
- Usually this is done through galaxy redshift surveys: Euclid, DESI, Roman/WFIRST…
- For the radio: HI 21cm line from each galaxy
- Very “expensive”!

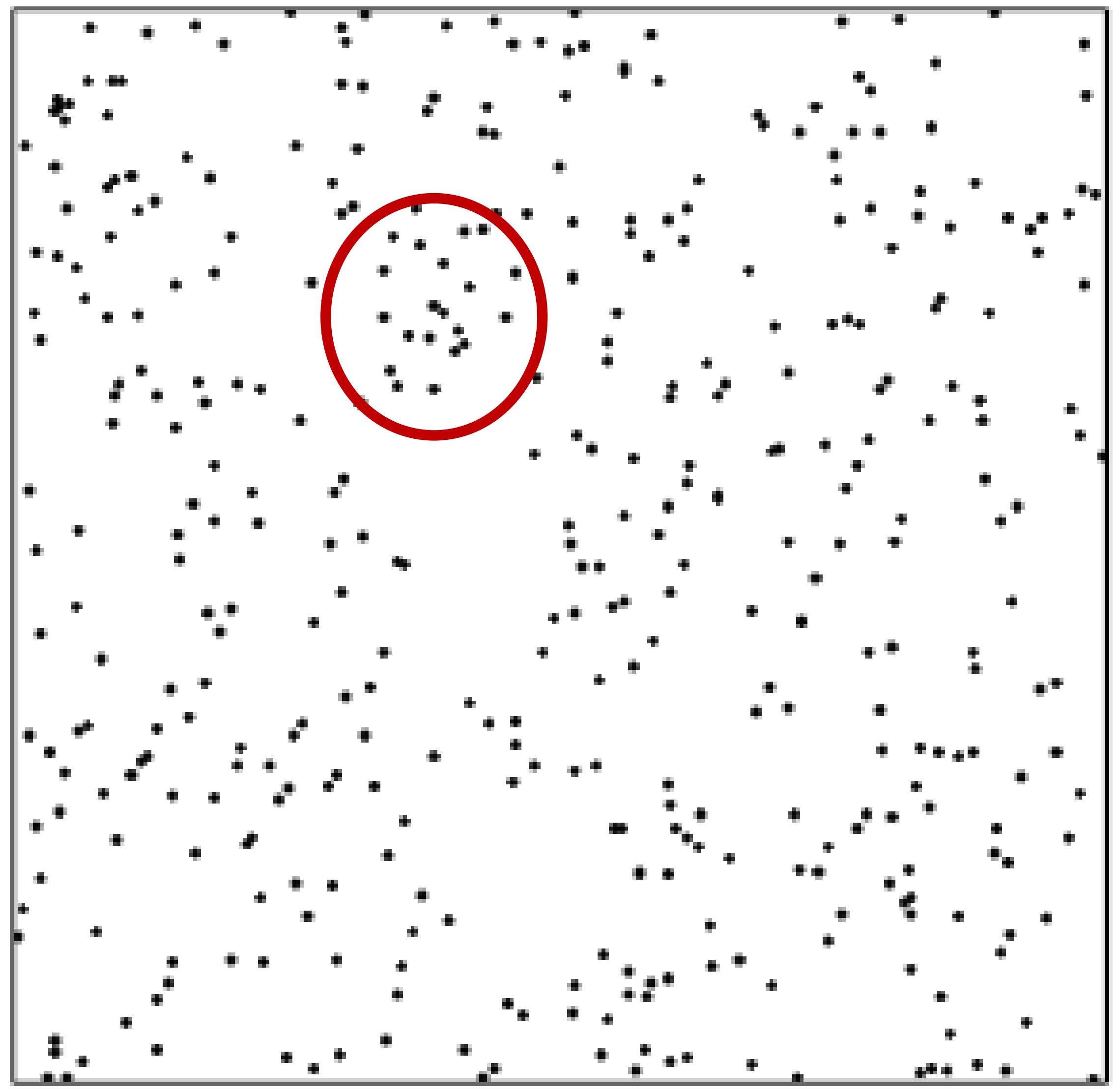


(SDSS/DESI)

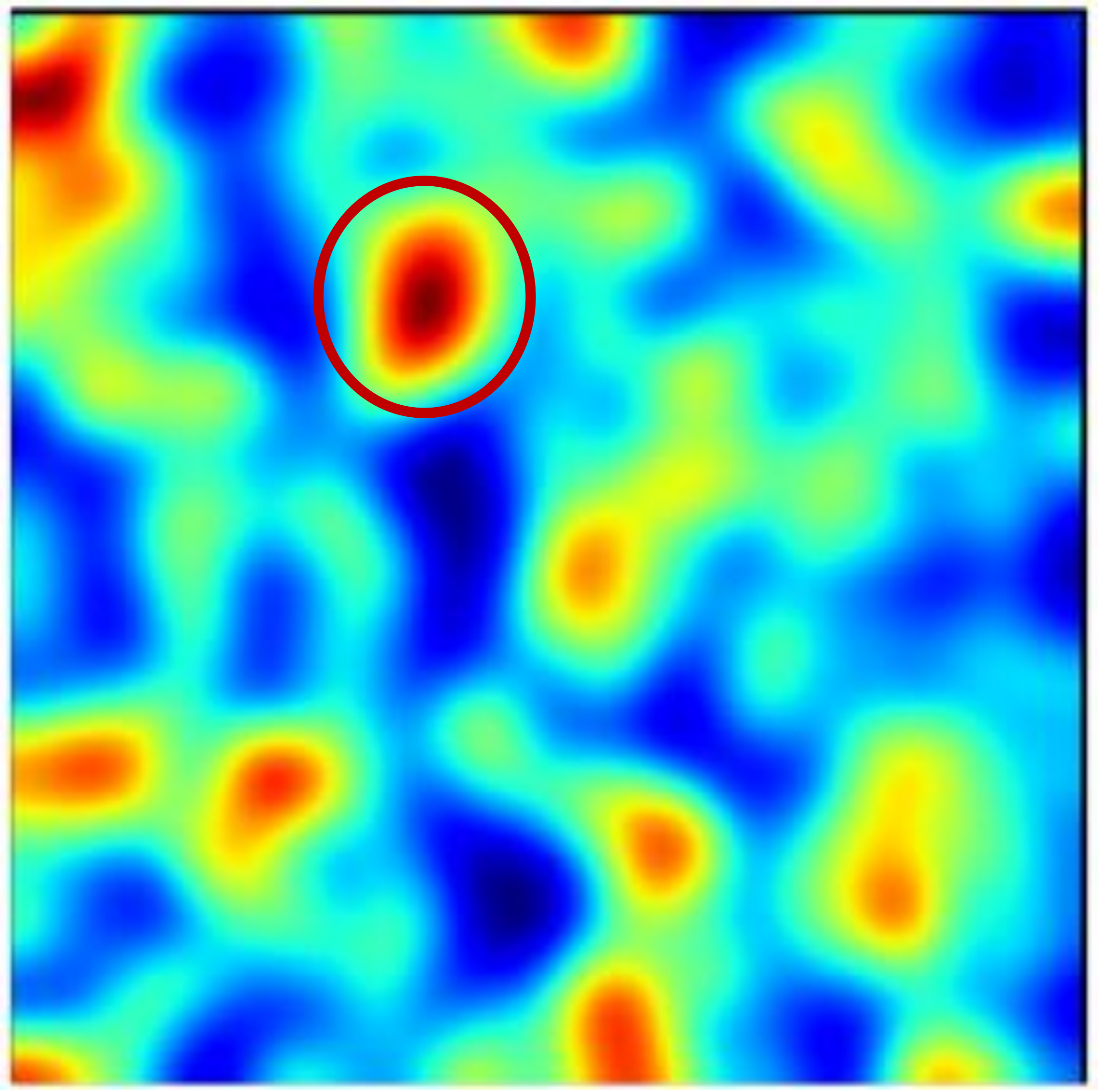
Probing LSS with intensity mapping

- For Cosmology, scales of interest are well beyond galaxy scales (Baryon Acoustic Oscillations ~ 150 Mpc)
- Intensity mapping is very fast → no threshold cutoff
- Provides high frequency/redshift resolution (in the radio…)
- Pixel will have joint emission from multiple galaxies
- Signal ~ 200 uK at z~1

Note: only way to probe the IGM HI

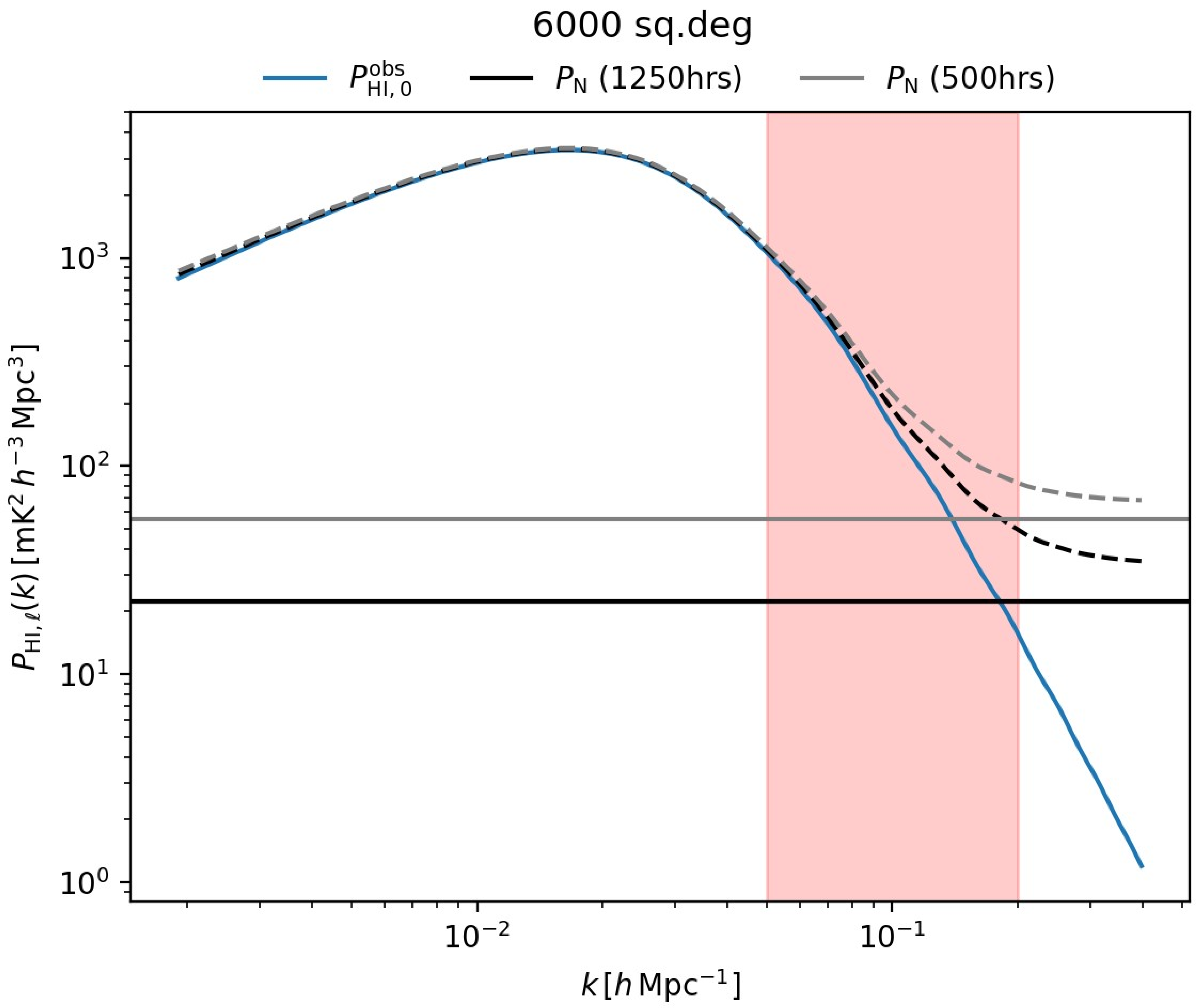


galaxies



Intensity map

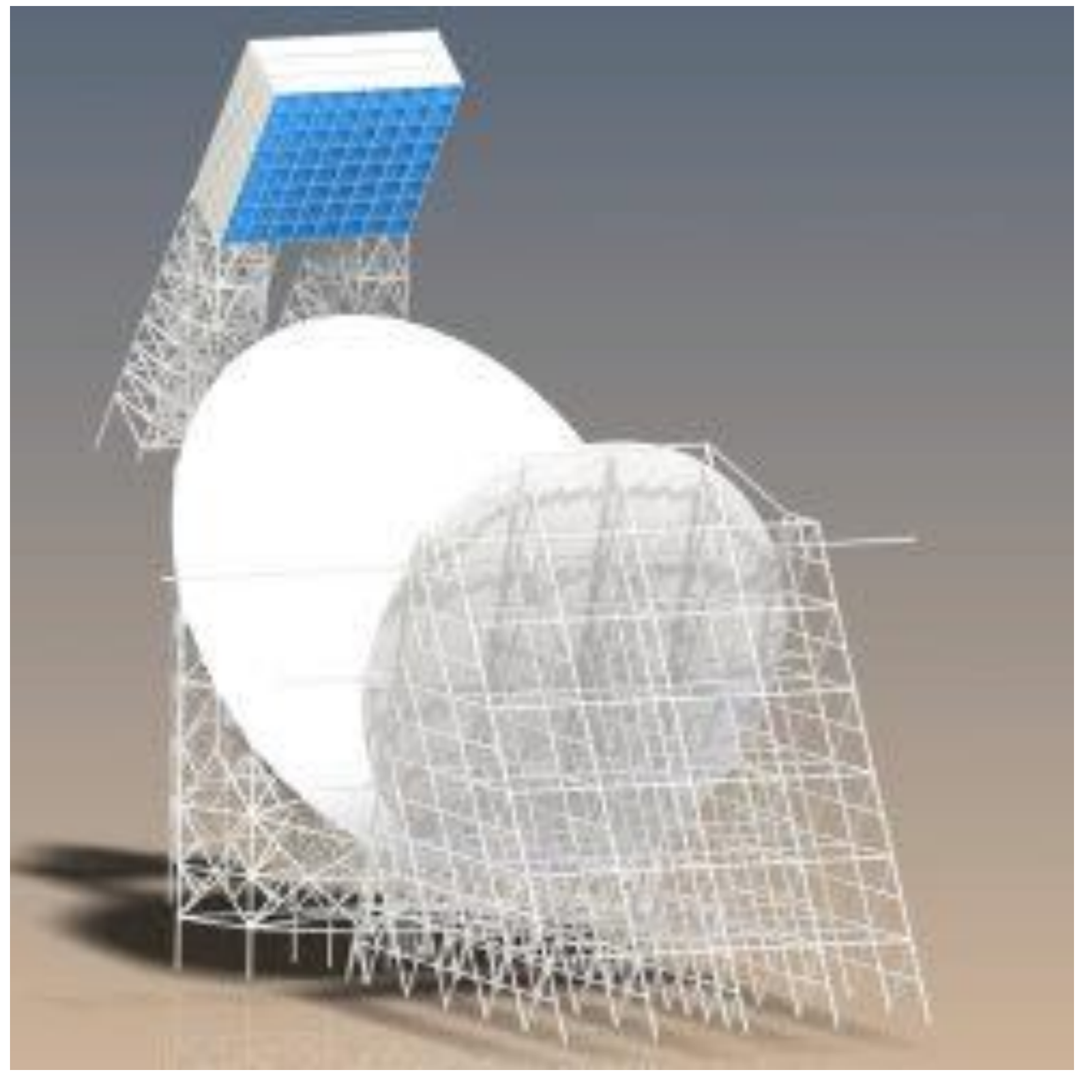
HI IM makes it “easy” to probe the power spectrum



- An example with the MeerKAT telescope: 500 hours is enough to detect the baryon acoustic oscillations. Noise is well below the signal on large scales

Dedicated experiments…

- HIRAX (South Africa)



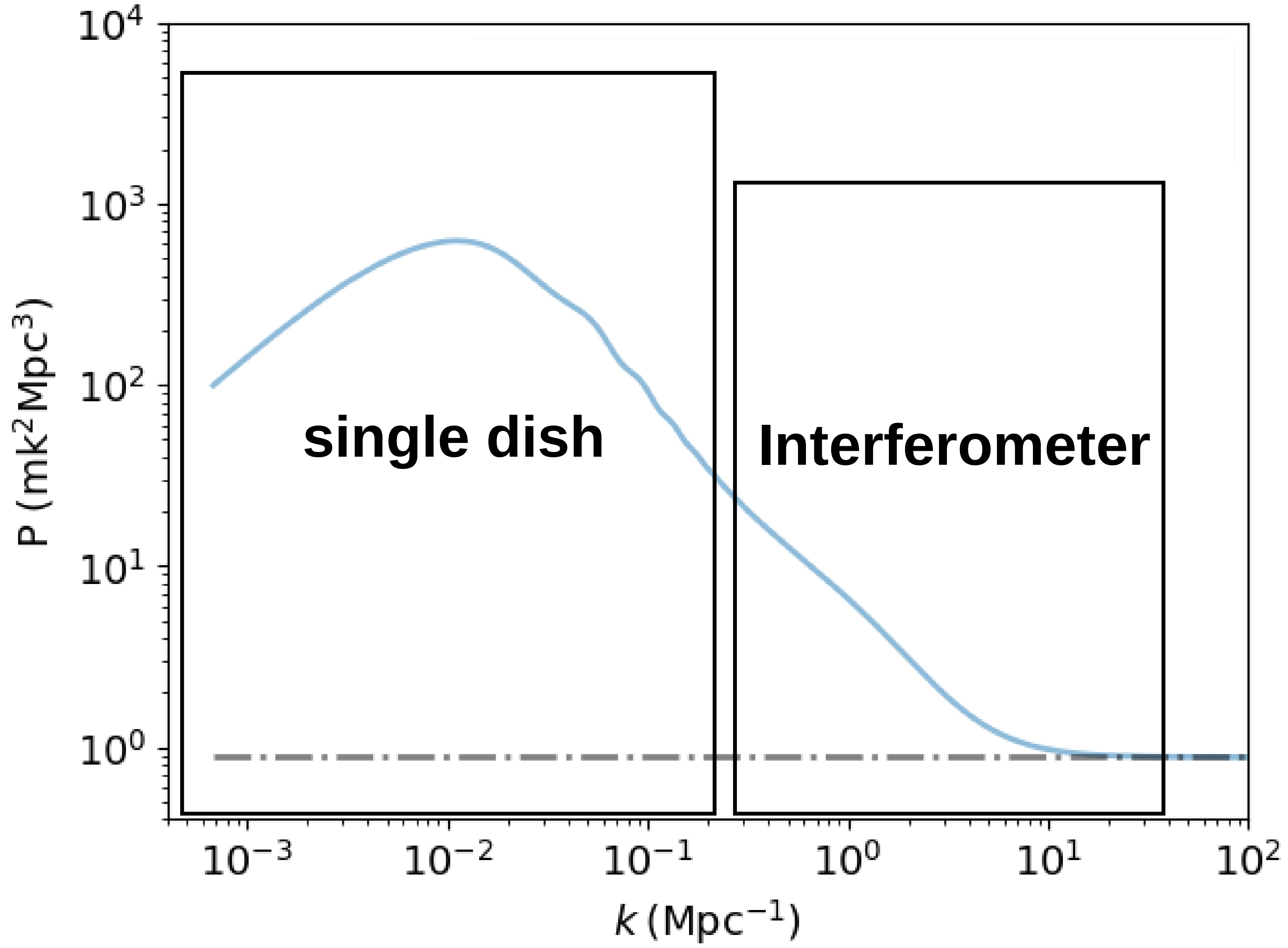
- BINGO? (Brazil)

- Tianlai (China)

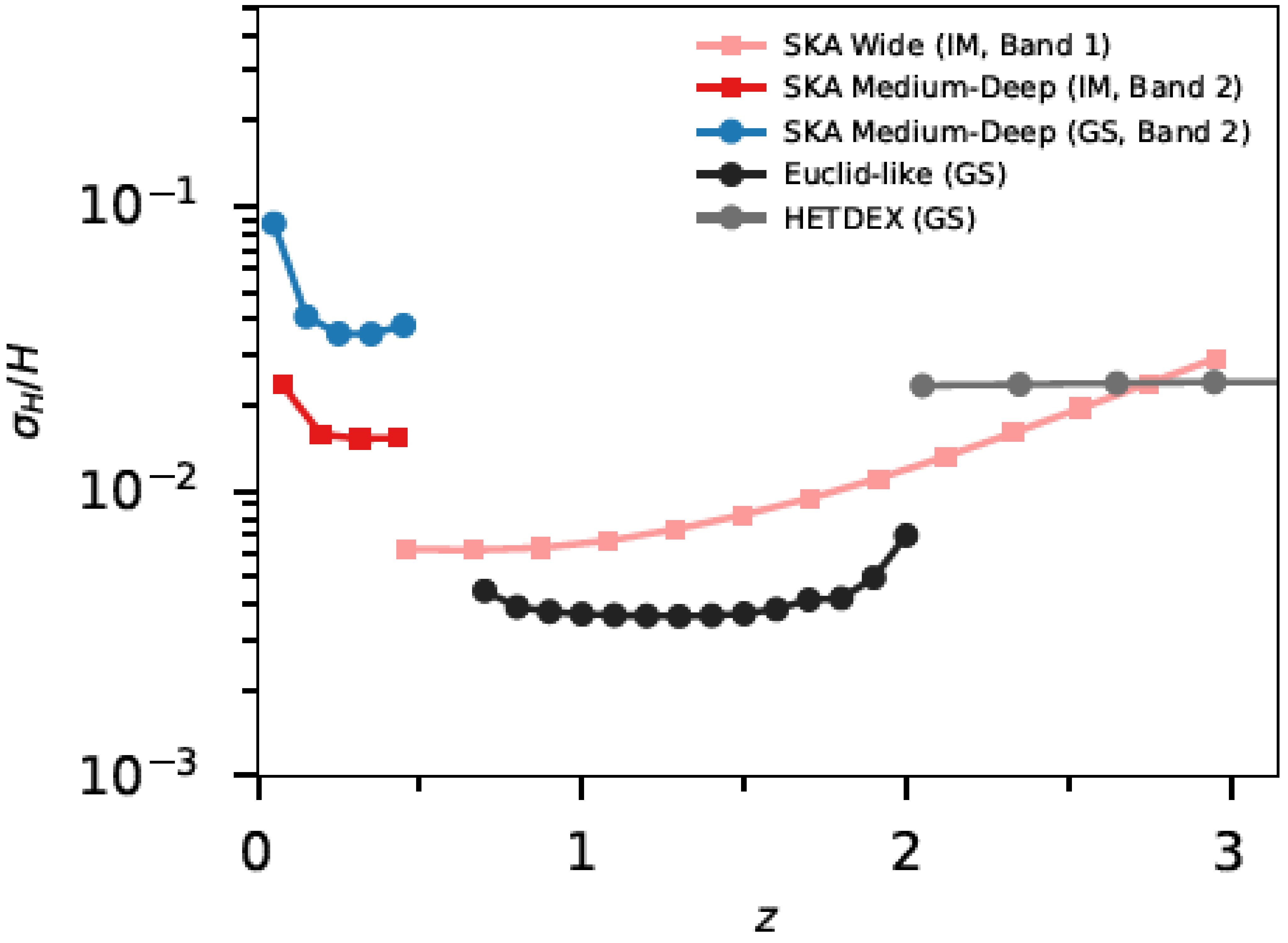
- CHIME (Canada)

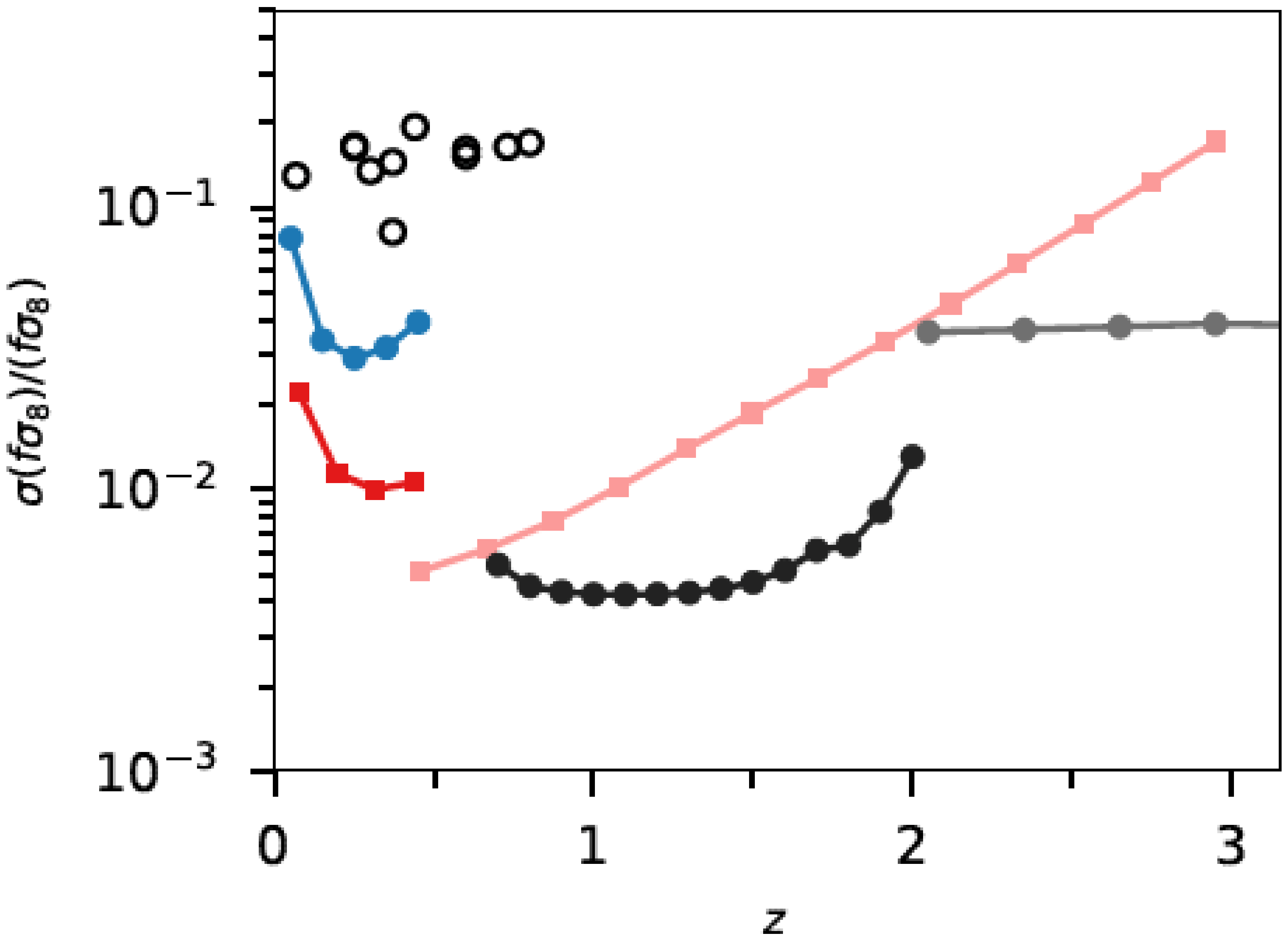
SKA1?

- Need SKA1-MID for z < 3 but baselines not small enough to probe BAO scales and above…
- Plan: use the array in "single dish mode"
- SKA1-MID single dish HI intensity mapping survey will turn SKA into a state of the art cosmology machine
- Only way to really go after the unexplored very large scales
- See: arXiv:1305.6928, arXiv:1405.1452, arXiv:1501.03989, arXiv: 1509.07562, arXiv:1811.02743



"Standard" Cosmology with SKA1-MID

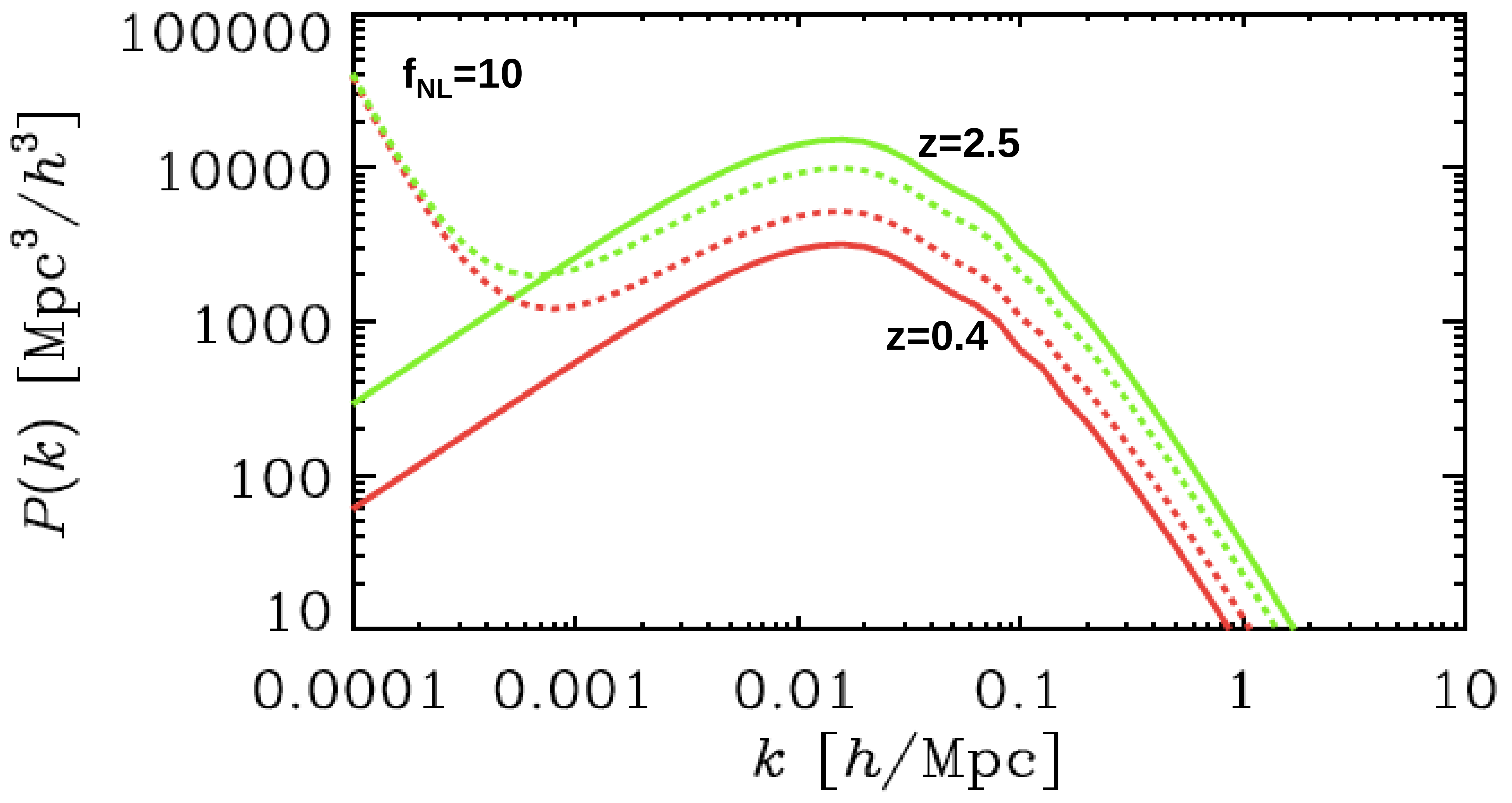




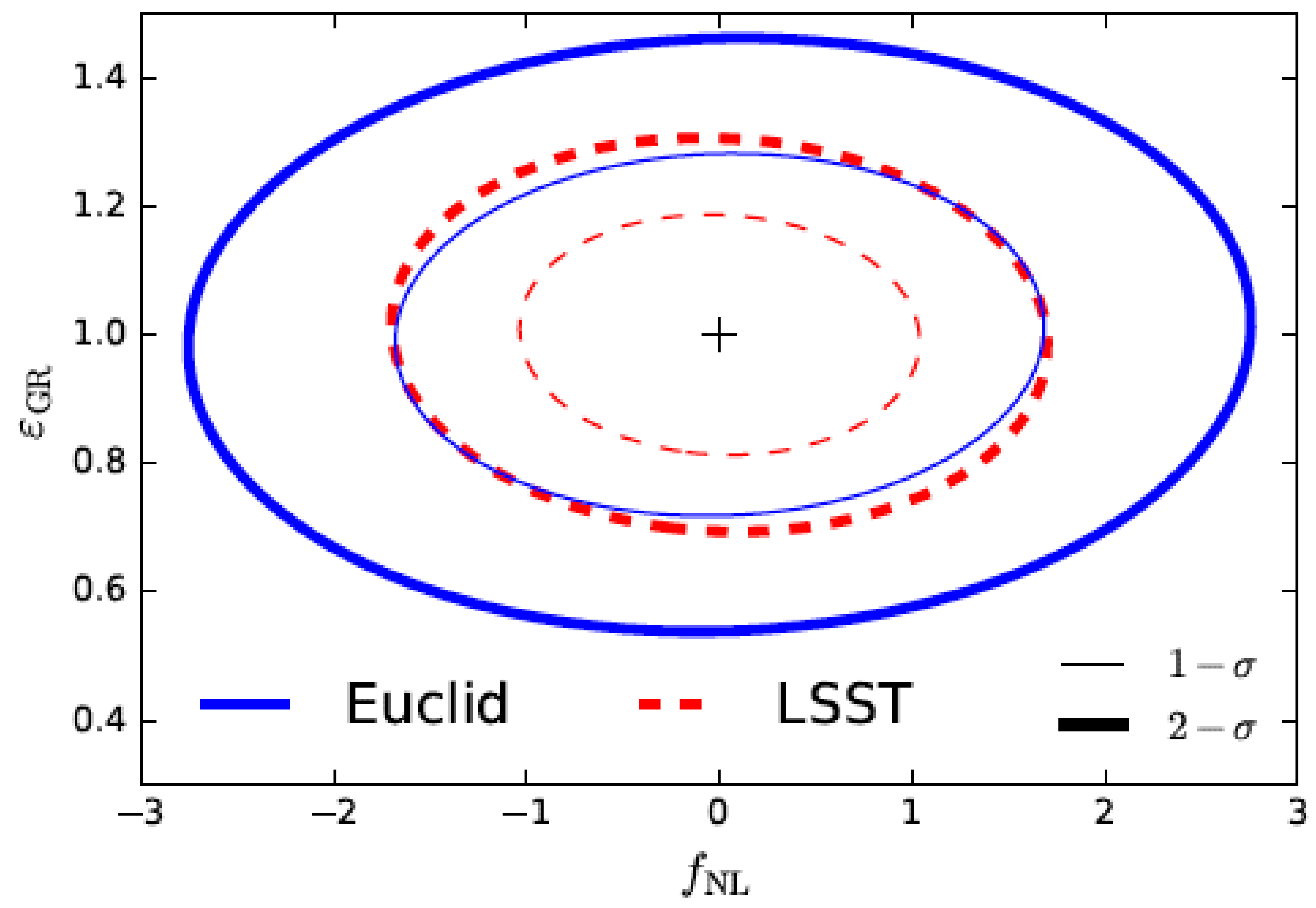
- Error forecast for the Hubble rate
- Error forecast for the growth rate

- SKA1 Cosmology "red book": arXiv:1811.02743

Constraints on large scale effects with SKA1-MID and multi-tracers



Camera et al., PRL, 2013



SKA1 Cosmology “red book”: arXiv:1811.02743

- The information is in the bias with respect to the dark matter field -> use multi-tracers to beat cosmic variance
- Combining an HI intensity mapping survey using SKA1-MID Band 1 with LSST will detect fNL ~ 1 as well as GR corrections
- A nice way to “fight” systematics
- See also: Alonso and Ferreira, PRD, 2015; Alonso et al. ApJ 2015; Fonseca et al., ApJ Letters, 2015; A Witzemann, et al., MNRAS, 2019; Matarrese and Verde, Astrophys.J. 2008; Dalal et al., PRD 2008; Squarotti et al., arXiv:2307.00058v1; Karagiannis et al., arXiv:2305.04028v1; Jolicoeur, arXiv:2301.02406v3…

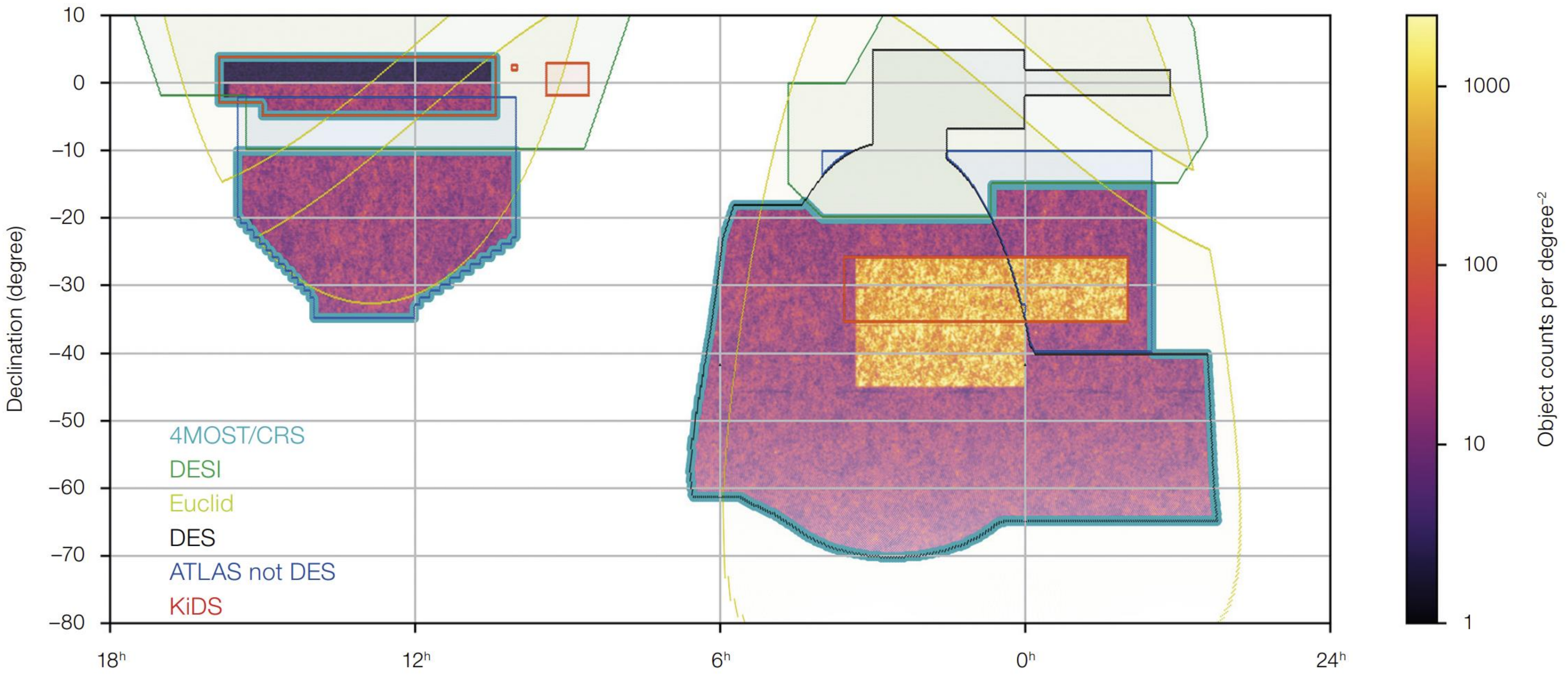
MeerKAT?

- 64, 13.5 m dishes – 2018
- Maximum baseline: 8 Km - soon ~ 20Km
- Frequencies: 580 MHz – 3500 MHz ($0 < z < 1.5$)
- It's in the South!
- Part of SKA1-MID in the future

The present: an SKA cosmology survey precursor with MeerKAT

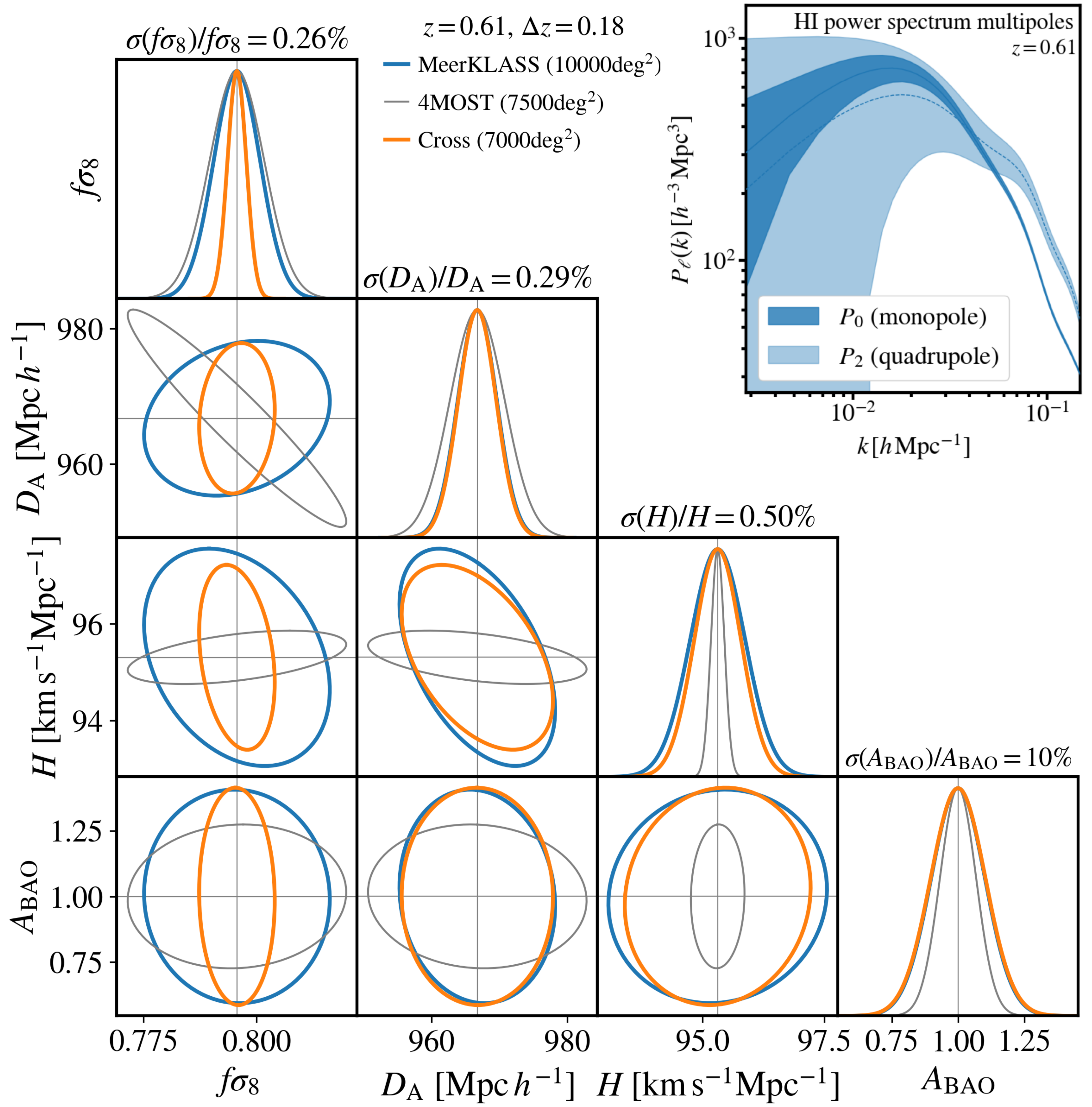
- ***MeerKLASS: MeerKAT Large Area Synoptic Survey:*** http://arxiv.org/abs/1709.06099
- Aim: Cosmology (HI intensity mapping) but commensal with lots of other science (continuum survey)
- Focus on sky patches with multi-wavelength data for cross-correlation (DESI, 4MOST, Euclid, Rubi/LSST, DES)

- L-band: 900-1670 MHz (z<0.58) ~ 100 hours observed
- **UHF band:**
  - 580 MHz-1015 MHz (0.40 < z < 1.45)
  - ~ 120 hours observed
  - Goal: 2,500 hours over 10,000 deg^2 (25 uJy rms in continuum) within next 5 years



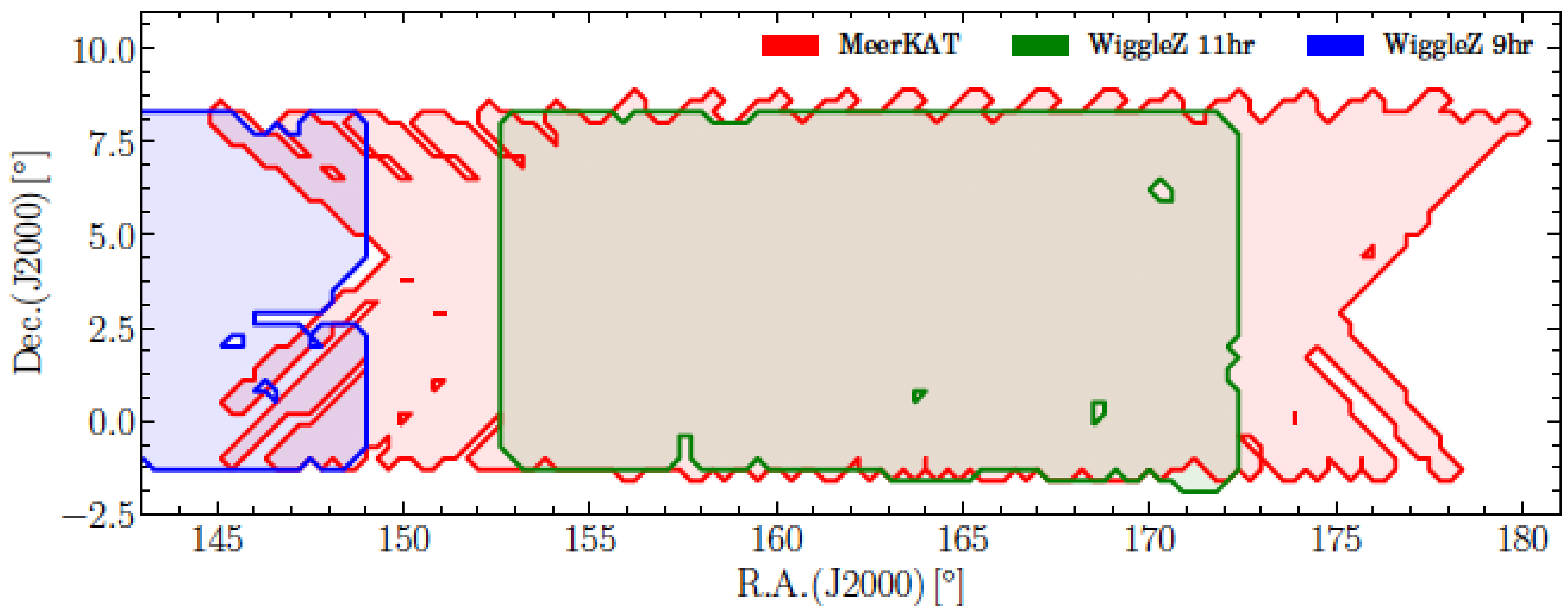
MeerKLASS: Cosmology

- Measurement of Baryon Acoustic Oscillations (BAO), Hubble rate and redshift space distortions
- Measure the HI content of the Universe at 0.4 < z < 1.4 (UHF-band)
- Cross-correlations with galaxy surveys
- Constraints of primordial non-Gaussianity (f_{NL}) by measuring large scale correlations and multi-tracers (Fonseca et al., arXiv1611.01322)
  - xDESI ~ 4.3
  - x4MOST ~ 3.5
  - xEuclid ~ 1.5
  - xDES ~ 3.5
  - xRuby/LSST ~ 1.8
  - (compare to CMB ~ 5 and eBOSS ~ 20)

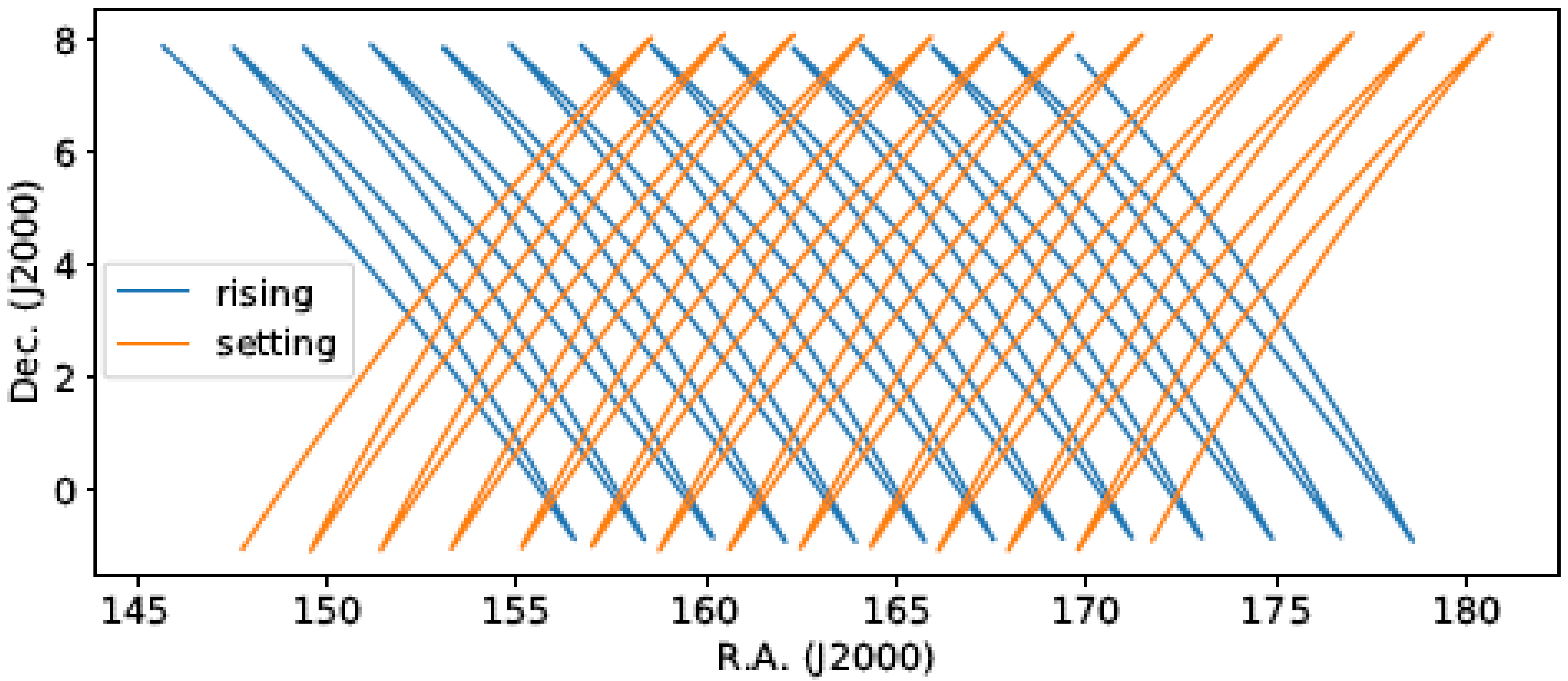


MeerKAT: 1,300 hours. 60 dishes

First results with a MeerKAT single dish pilot survey
(Wang et al., MNRAS, arxiv:2011.13789)



Footprint

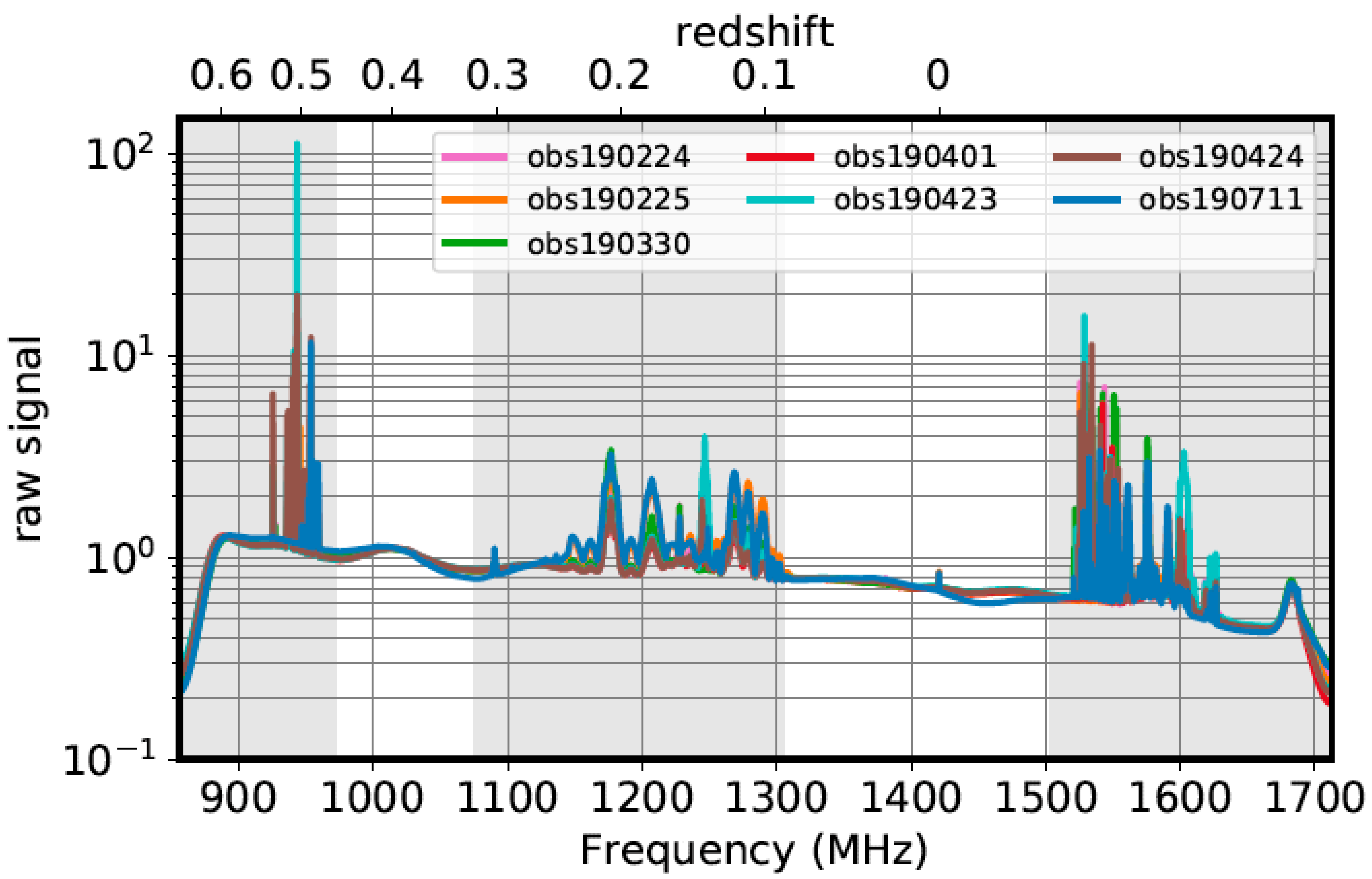


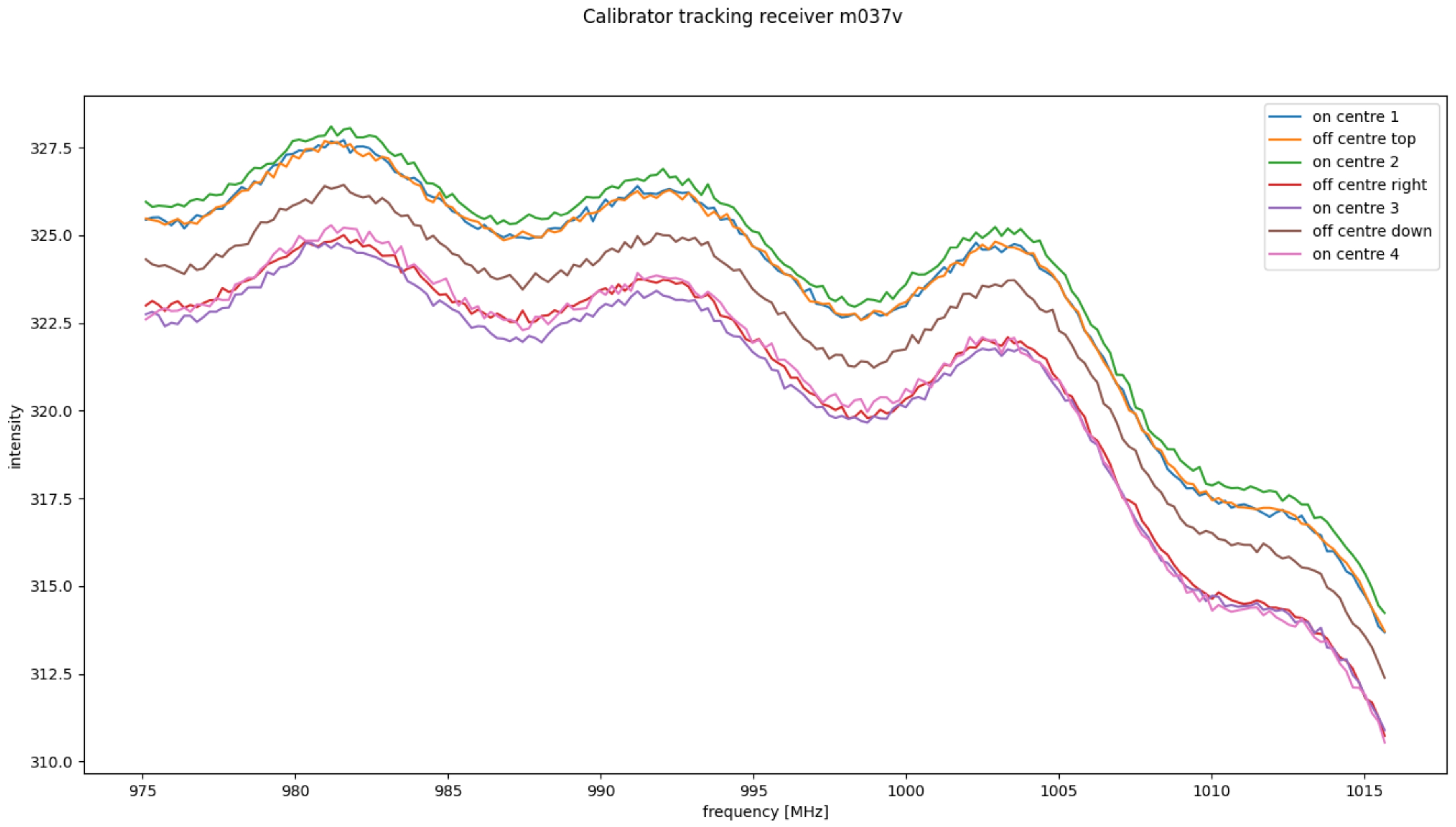
Scanning pattern

- ~ 15 hours
- ~ 60 dishes used (~ 600 hours combined)
- ~ 200 deg^2 over the WiggleZ 11h field
- L-Band: 900 MHz - 1700 MHz ($z < 0.5$)

- Resolution: 2 sec/0.2 MHz
- Scans at constant elevation (> 40 deg)
- Speed: 5 arcmin/sec
- ~ 200 sec per scan line, 1.5 hours per block

Contamination (sky foregrounds+ground+RFI+instrumental)

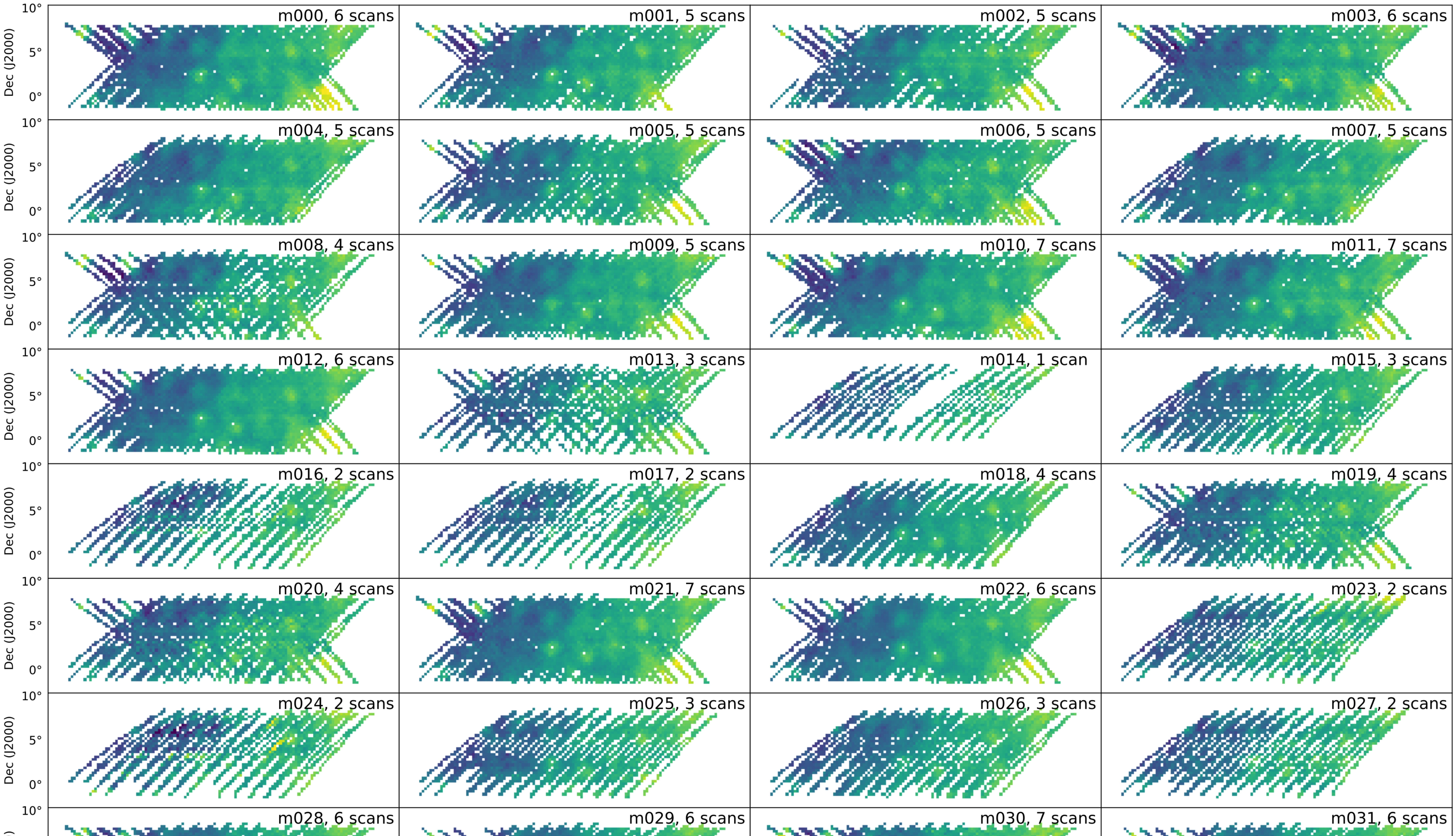




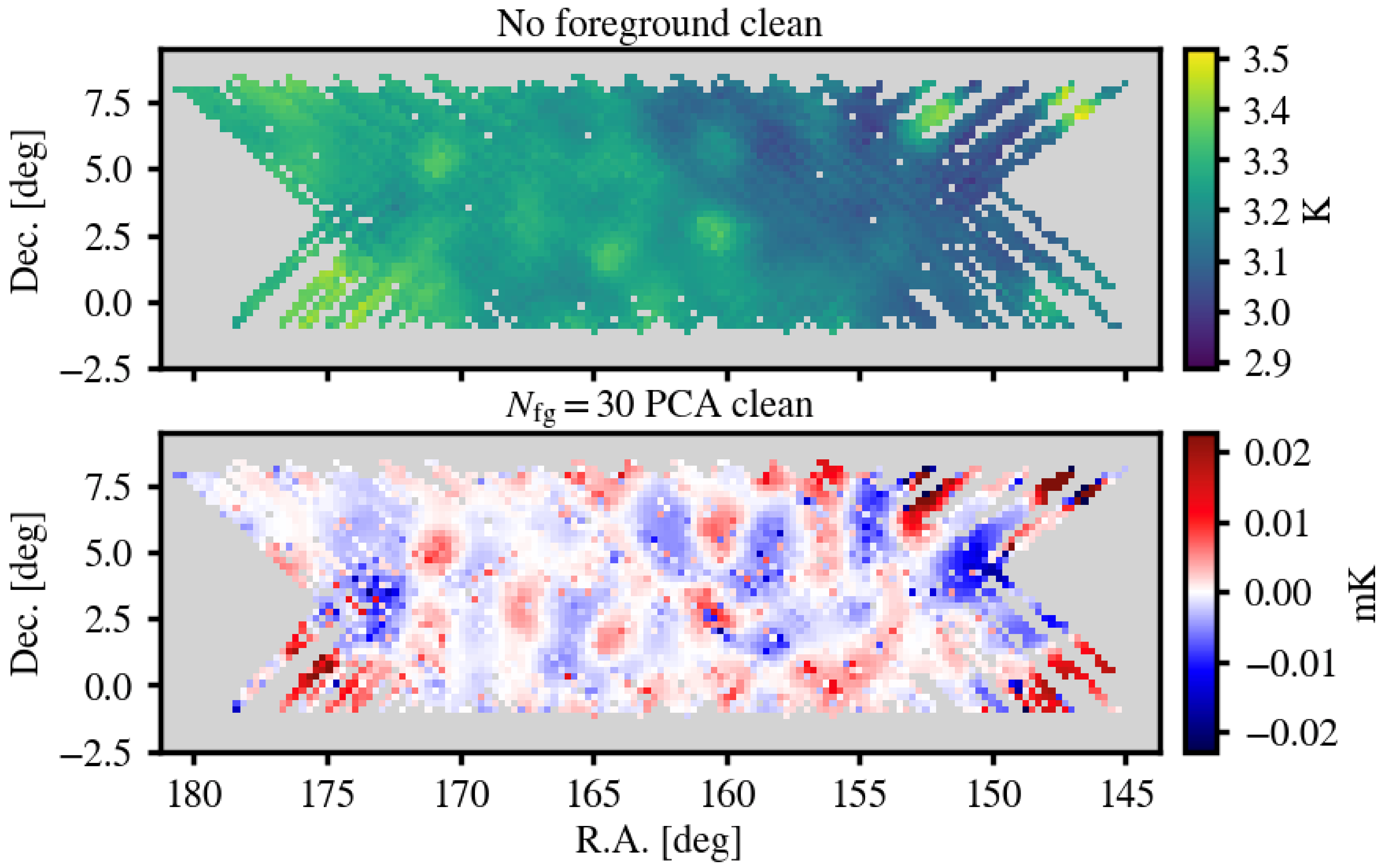
Looking for fluctuations ~ 1/10^5

- Satellites are a big concern, in particular with single dish data and in particular from the beam sidelobes
- RFI free regions in L band: **0.32 < z < 0.46**
- Calibration/modelling is crucial
- Methods for foreground cleaning are crucial (PCA, GMCA, Gaussian Processes,Machine Learning…)
- Also important to improve signal extraction methods (power spectrum)

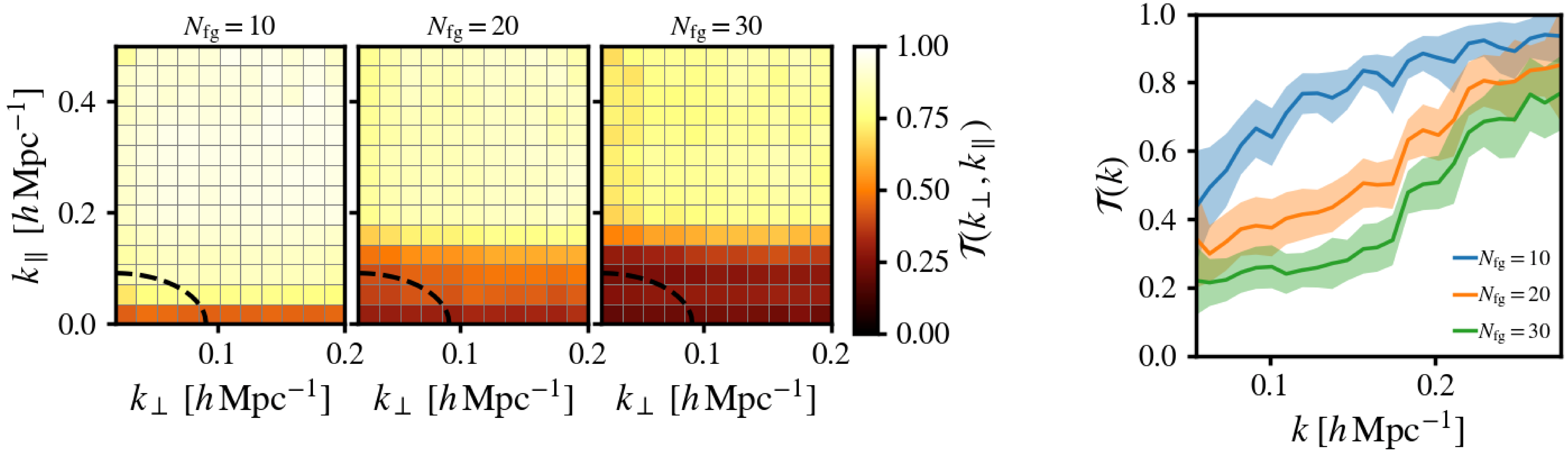
Temperature maps at 1023 MHz – we can cross-correlate between dishes!



Foreground cleaned maps



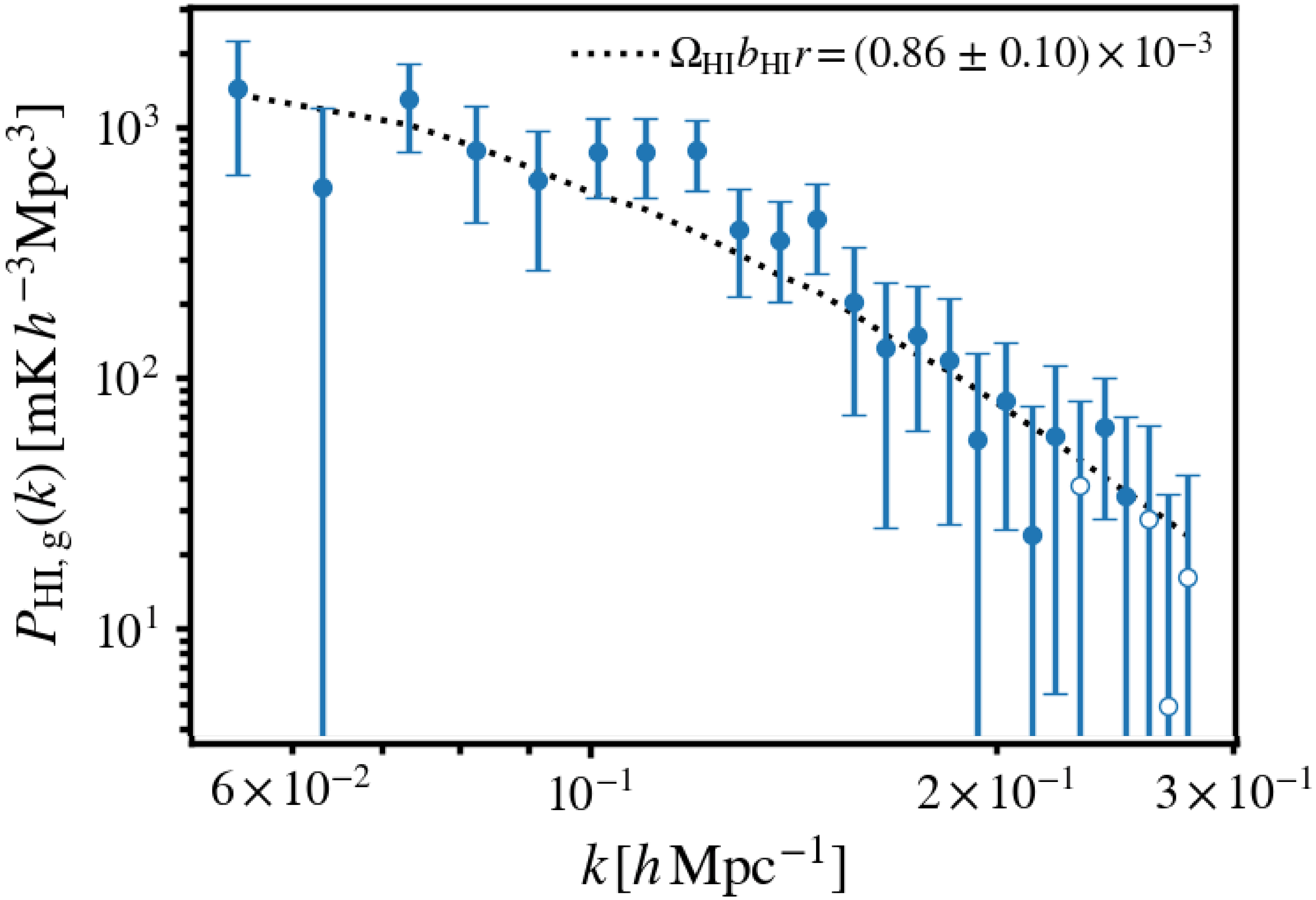
Transfer function / signal loss



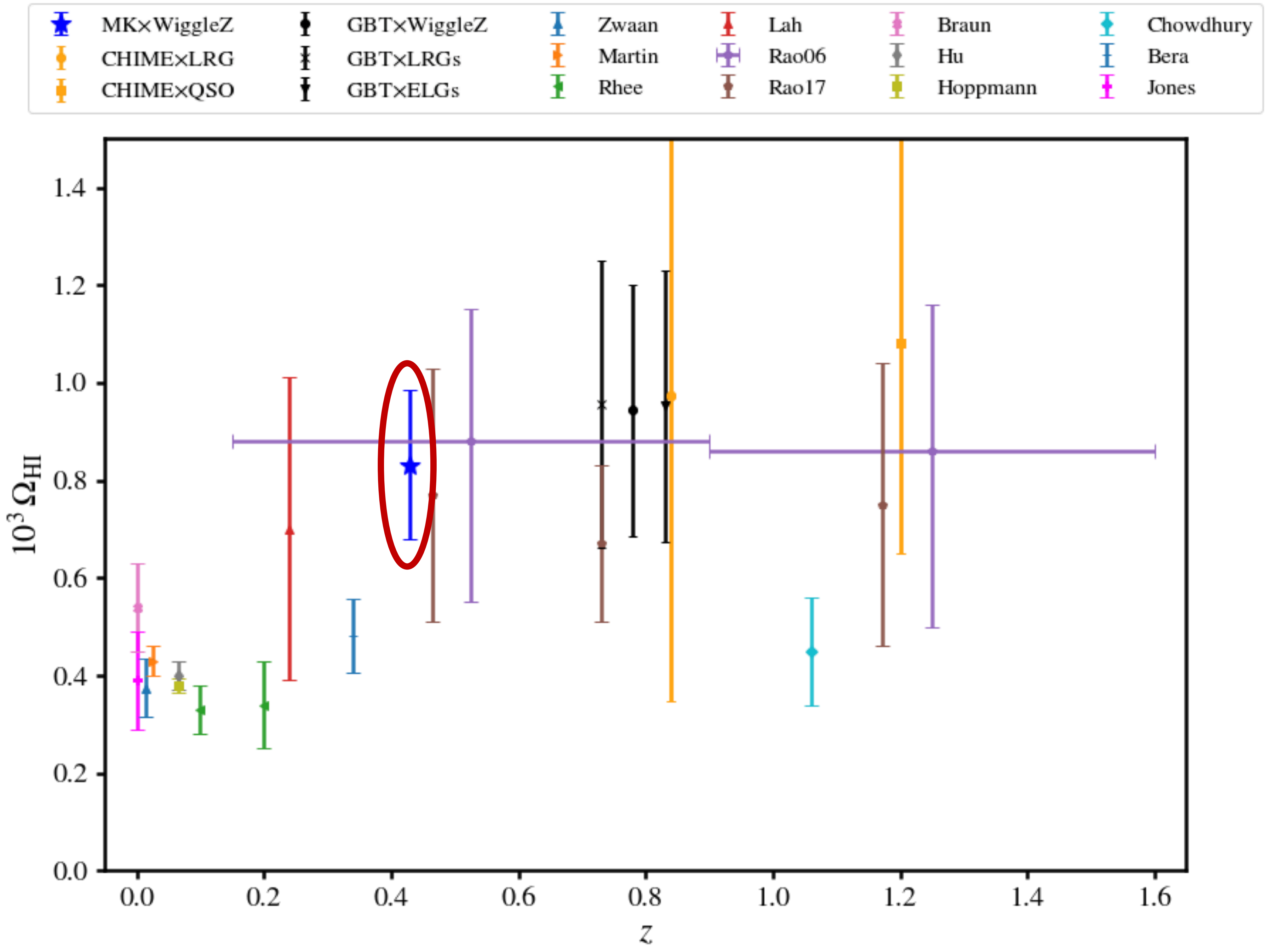
- Signal loss due to foreground cleaning affects all scales but mostly small k_par
- Transfer function crucial to unbias the power spectrum estimator
- Calculated through signal injection
- Need to improve calibration to reduce foreground cleaning!

See Cunnington, et al., MNRAS, arXiv:2302.02683

First cosmological results with MeerKAT: Detection of the cross-correlation power spectrum with WiggleZ galaxies



0.400 < z < 0.459

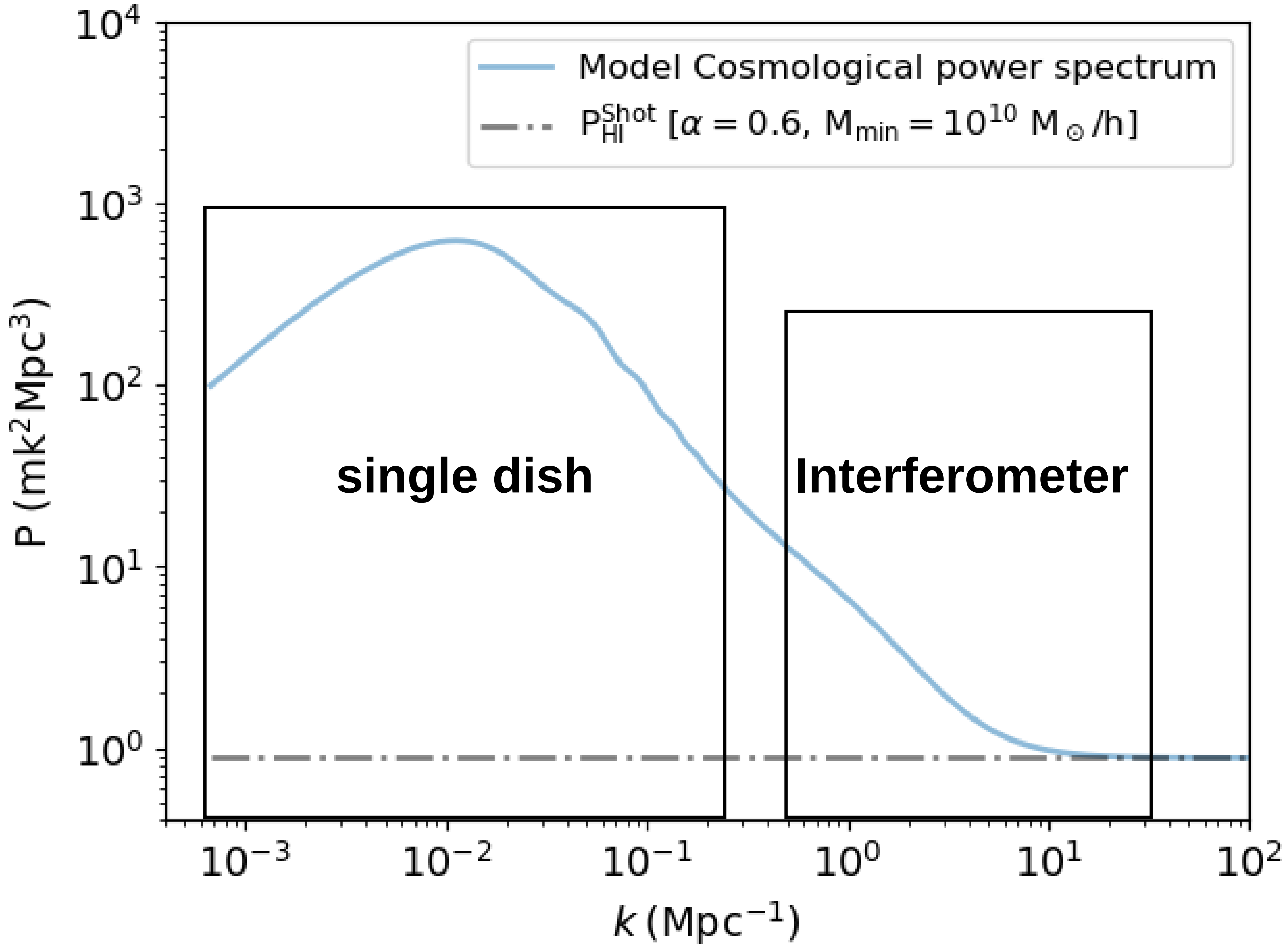


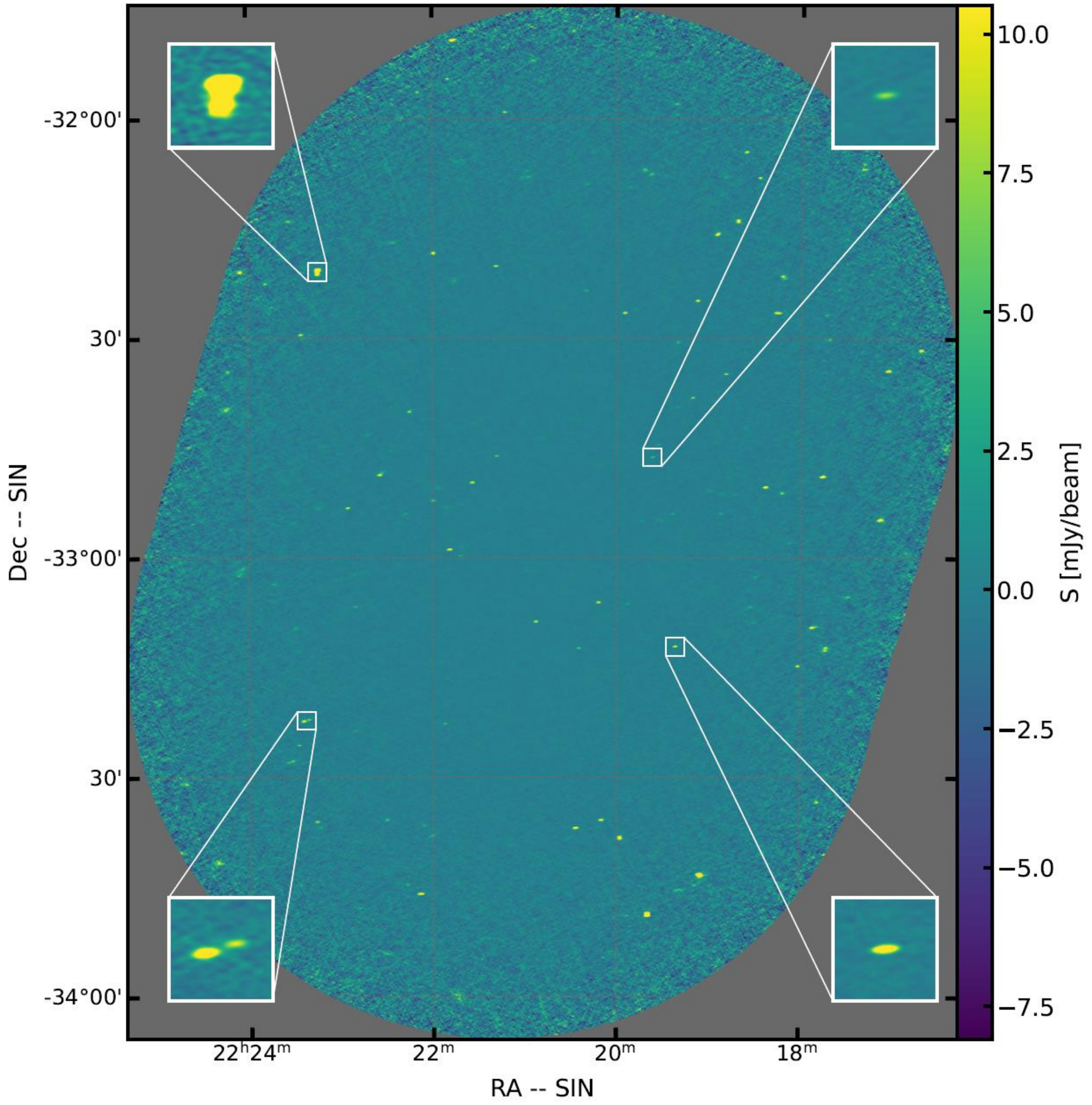
Cunnington, Li, Santos, et al., MNRAS 2023

What about the interferometer data?

- HI intensity mapping can still measure quasi-linear cosmological scales (k ~ 1 Mpc^{-1} and above)
- Great for comparison to full HI simulations
- Can be used for other science (continuum…)

z ~ 0.3

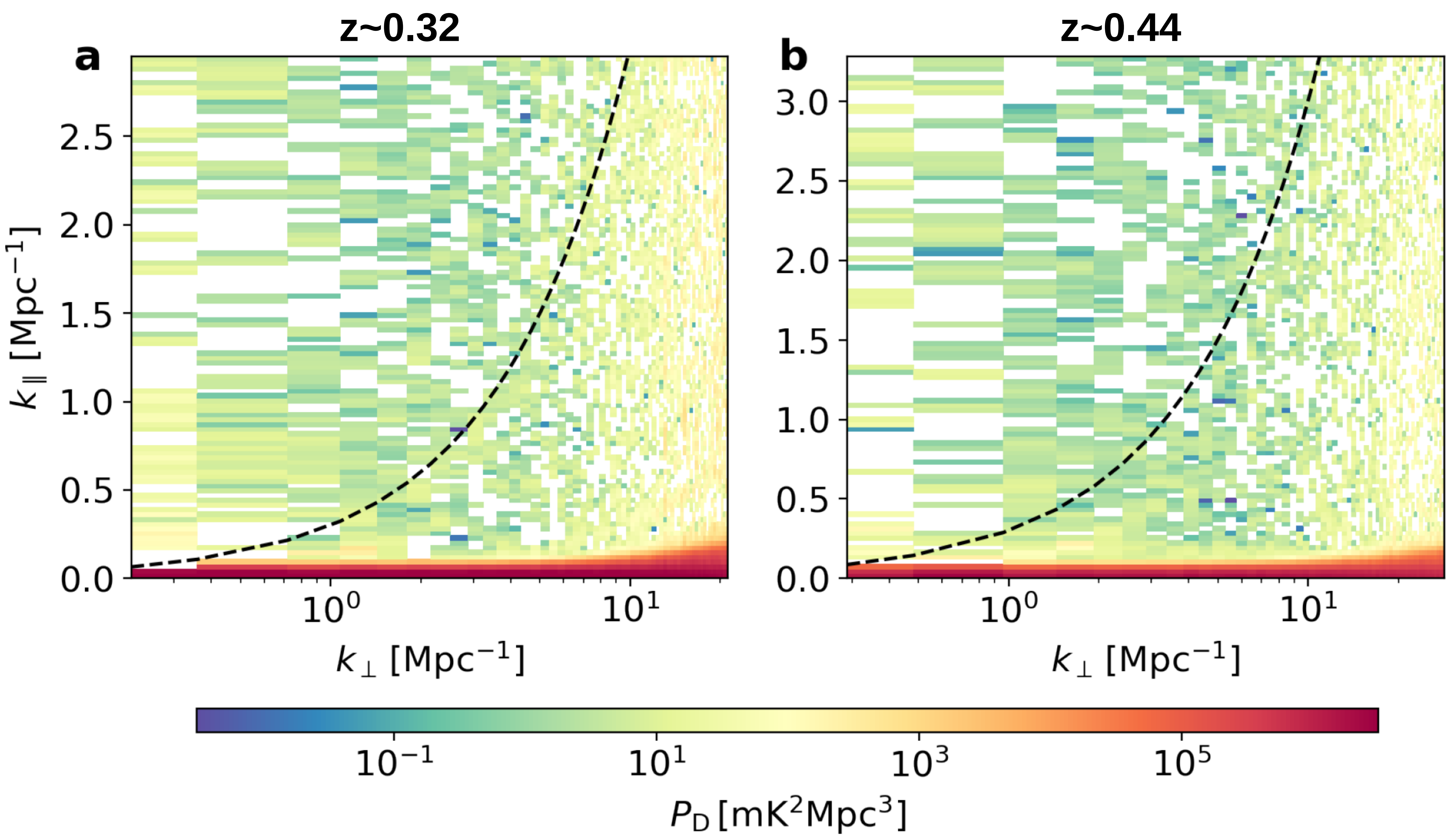




Kristof Rozgonyi

HI IM using the MeerKAT interferometer

- Field: ~ 2 deg^2, 96 hours
- 2d power spectrum (cross-correlate different times)
- Smooth components in frequency will be at low $k_{\parallel}$
- Pick a window with low continuum contamination ("foreground avoidance")
- Great way to test the halo model and compare to HI simulations!

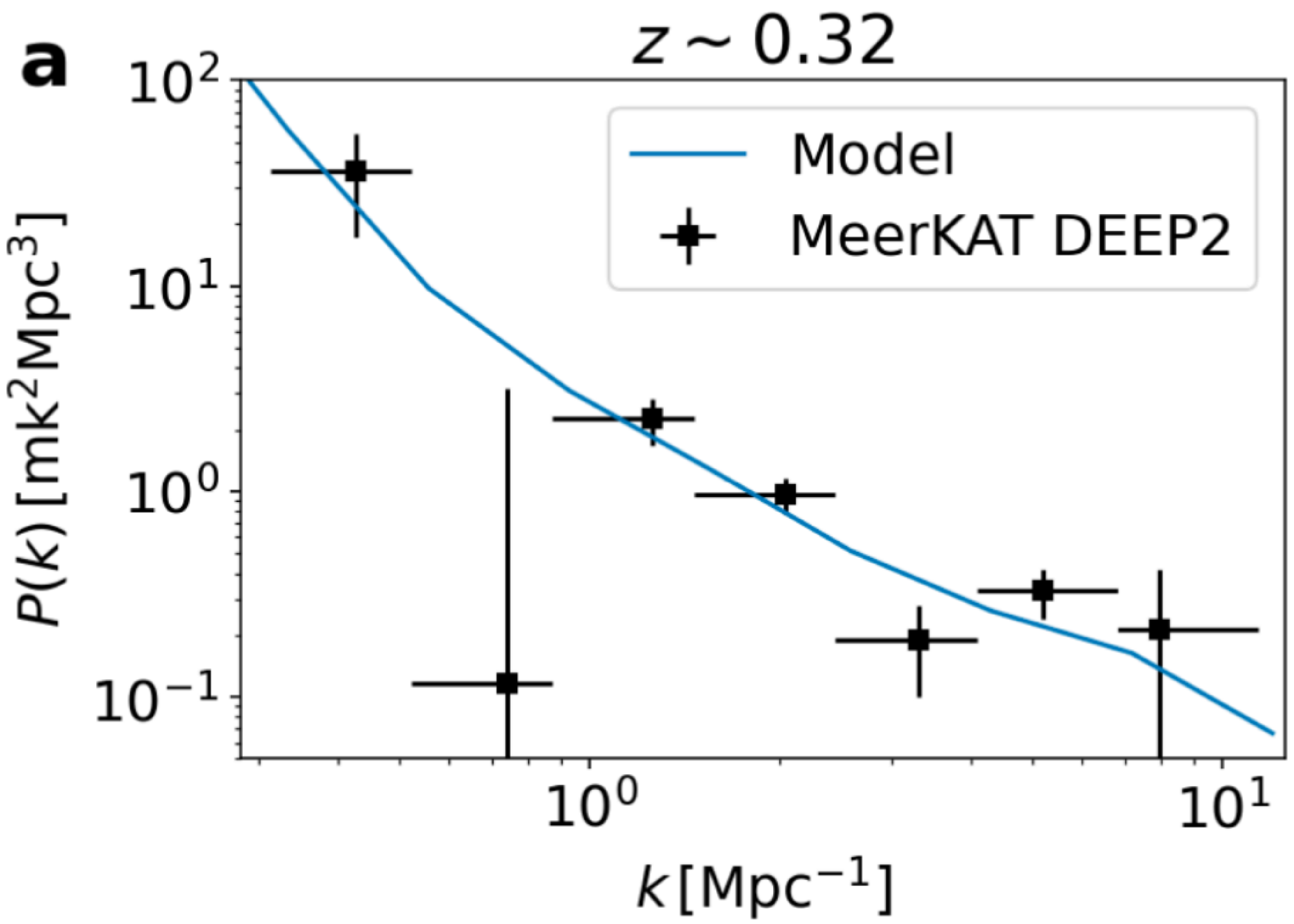


$$P_{\rm D}(k_\perp, k_\parallel) = P_{\rm 2h}(k_\perp, k_\parallel) + P_{\rm 1h}(k_\perp, k_\parallel) + P_{\rm SN}(k_\parallel)$$

$$P_{\rm SN}(k_\parallel) = \frac{P^0_{\rm SN}}{1 + (k_\parallel \sigma_p)^2/2}$$

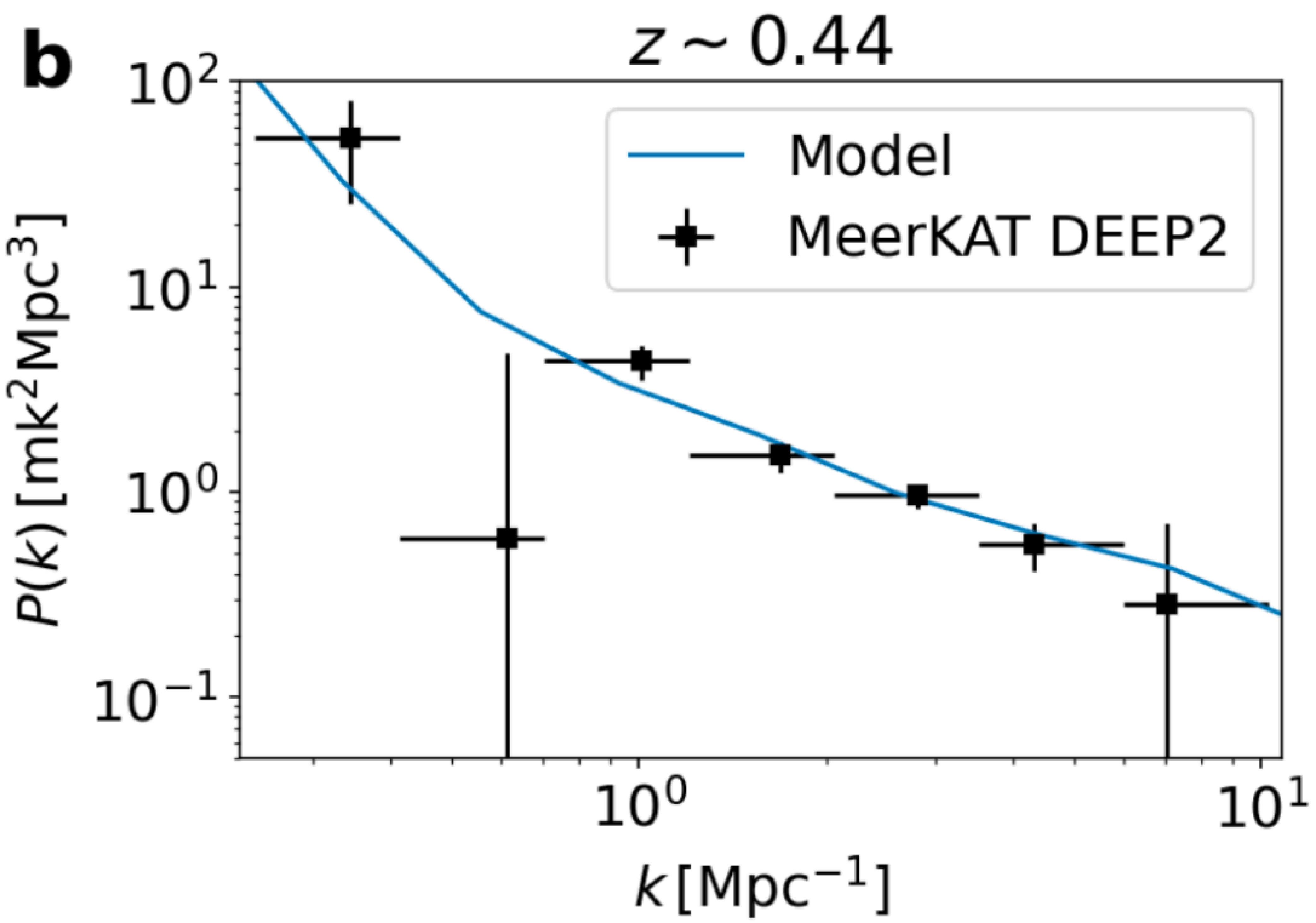
First direct detection of the HI power spectrum on Mpc scales

$\sigma_{\mathrm{HI}}\,(1\mathrm{Mpc}) = 0.44 \pm 0.04\ \mathrm{mK}$



8σ

$\sigma_{\mathrm{HI}}\,(1\mathrm{Mpc}) = 0.63 \pm 0.03\ \mathrm{mK}$



11.5σ

Sourabh Paul, Mario G. Santos ,Zhaoting Chen and Laura Wolz, https://arxiv.org/abs/2301.11943

Summary

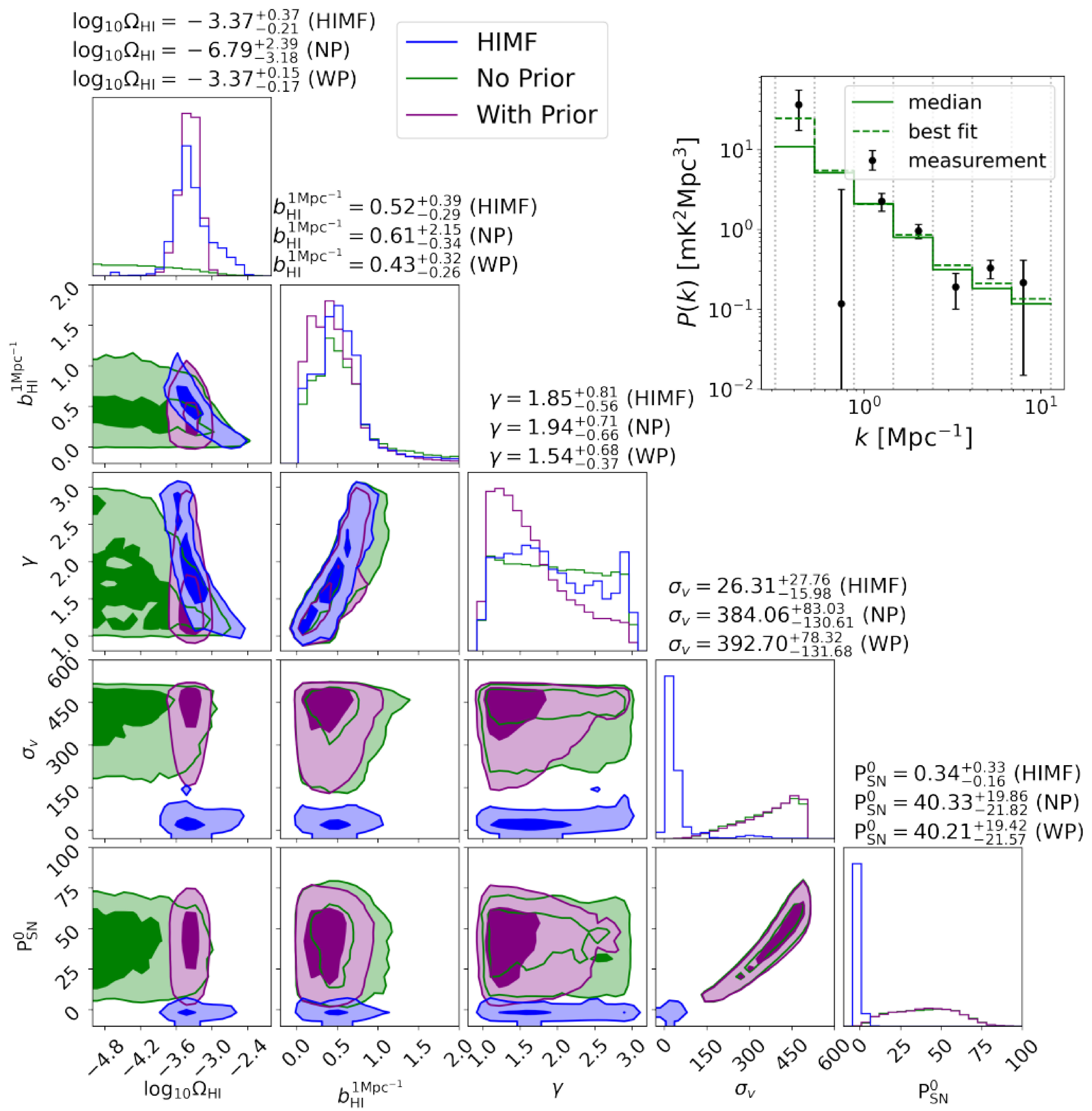
- HI intensity mapping with MeerKAT/SKA in single dish mode will deliver state of the art cosmological constraints: BAO in HI – dark energy, RSDs – modified gravity, primordial non-Gaussianity…
- Multi-wavelength cross correlations adds more than the sum of the parts
- MeerKAT interferometer data adds an exciting observational window to HI IM on small scales
- We have HI IM detections using the MeerKAT single dishes in cross with optical galaxies and auto detections with the interferometer – time to start fitting those theories!
- Dealing with low level systematics is the main challenge
- Ongoing observations and data processing with MeerKAT UHF single dish (BAO and past equality peak) – plan is to observe 2,500 hours over 10,000 deg^2 by 2028
- Ongoing data processing with MeerKAT interferometer data (small scales)
- Expect more results towards end of the year

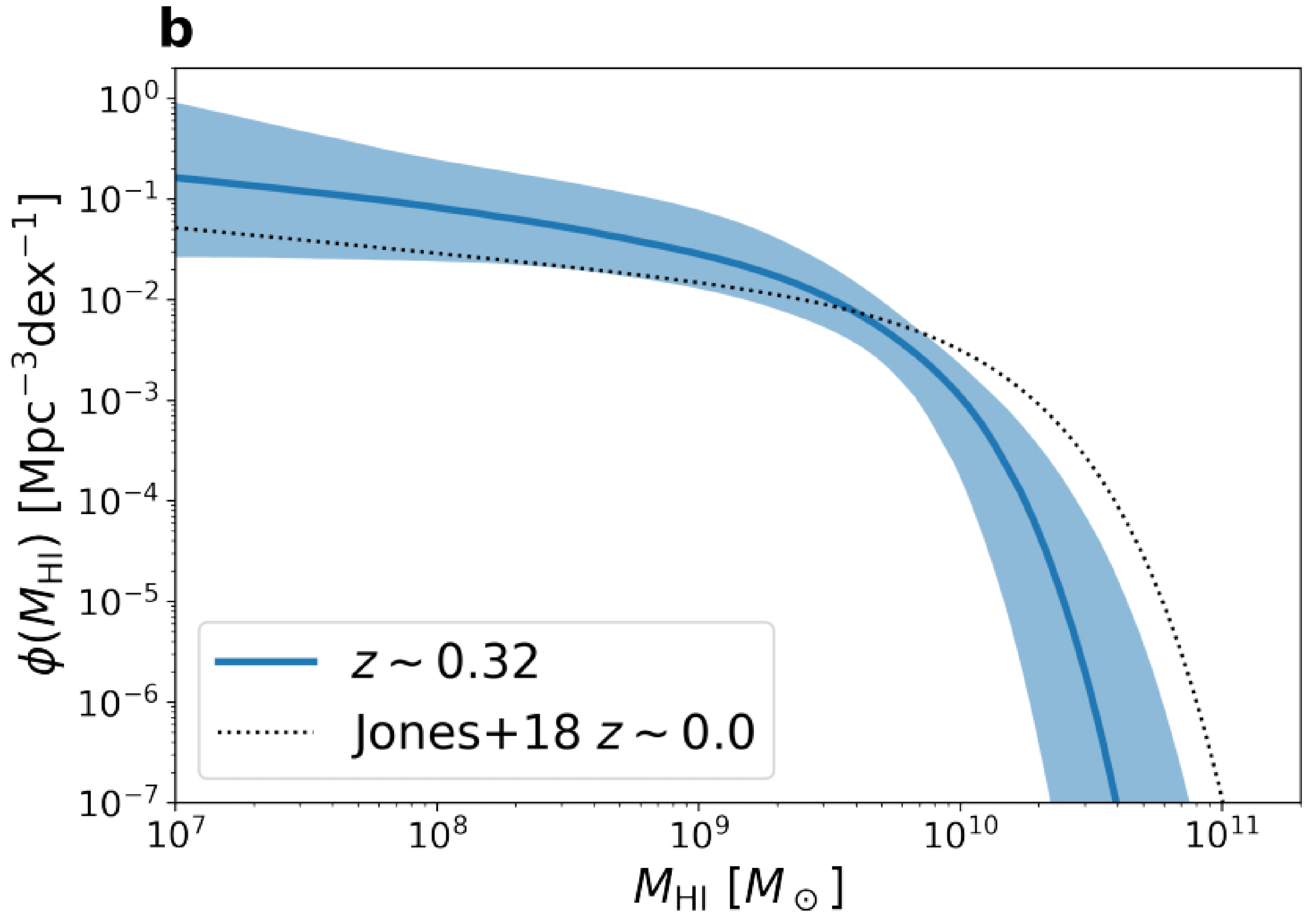
Bonus slides

Experiments

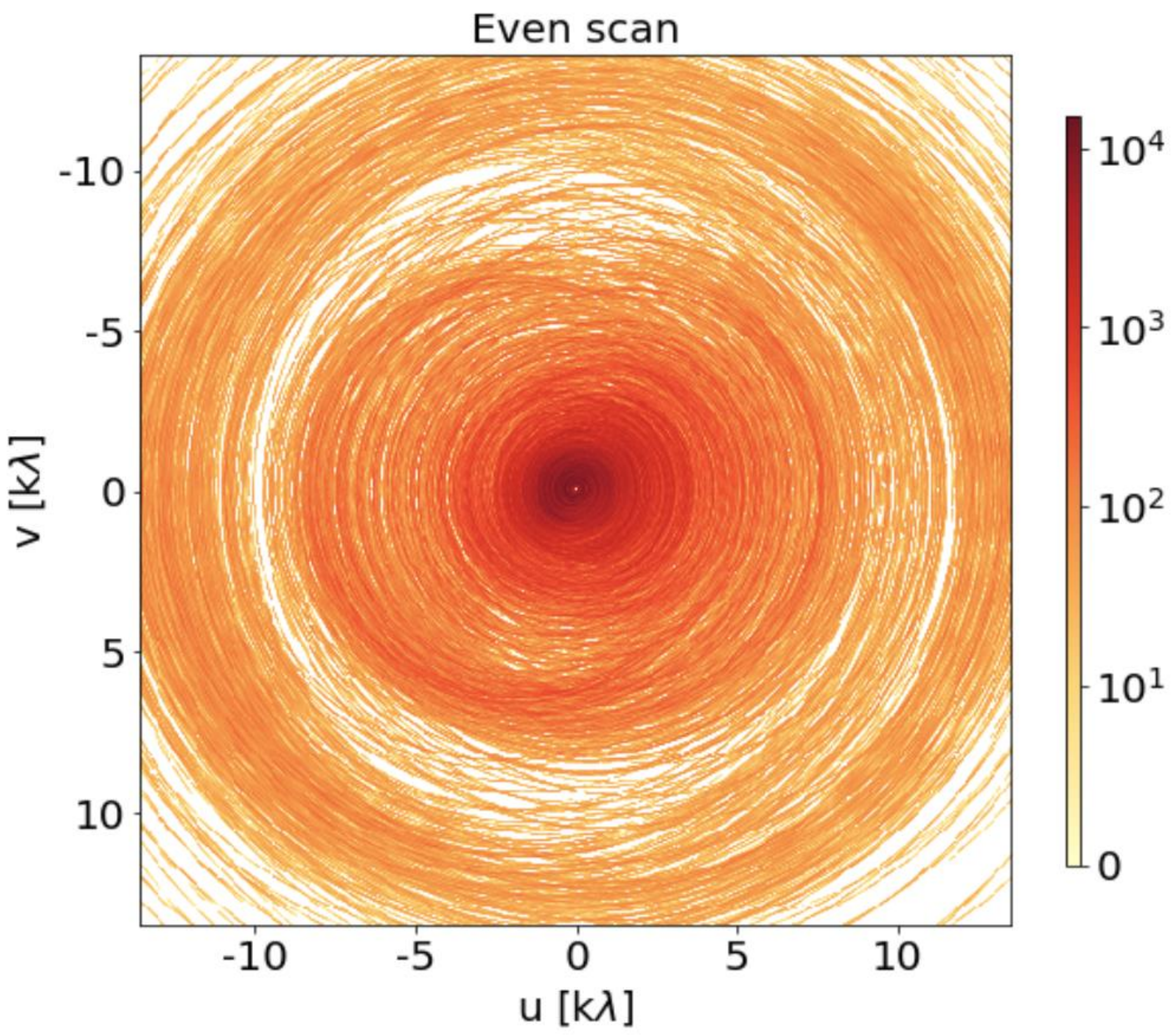
- “Tailored” interferometers: designed to have lots of measurements on the relevant angular scales (BAO scales). Highly redundant baselines
  - CHIME: Cylinder. Been taking data for a few years
  - HIRAX: ~ 1000 6m dishes placed in a close-packed redundant configuration (funding for 256 elements so far)
  - Tianlai: mix of dishes and cylinders. Been taking data for a few years
  - CHORD: 512 6m dishes + outrigger stations (seed funding?)
  - PUMA: 32,000 6m dishes in a packed array (not funded yet)
  - 2000-element Deep Synoptic Array (DSA-2000): More like an HI galaxy survey. 120 elements funded
- Tailored dishes:
  - BINGO: ~40m dish (primary) + ~ 35m (secondary), 28 feed horns (partially funded?)
- Multi-purpose instruments:
  - GBT
  - Parkes
  - FAST (in operation)
  - MeerKAT (both single dish and interferometer): 64, 13.5m dishes. Already taking data
  - SKA1-MID (both single dish and interferometer): MeerKAT + 133 15m dishes. Funded (I think). 2029?

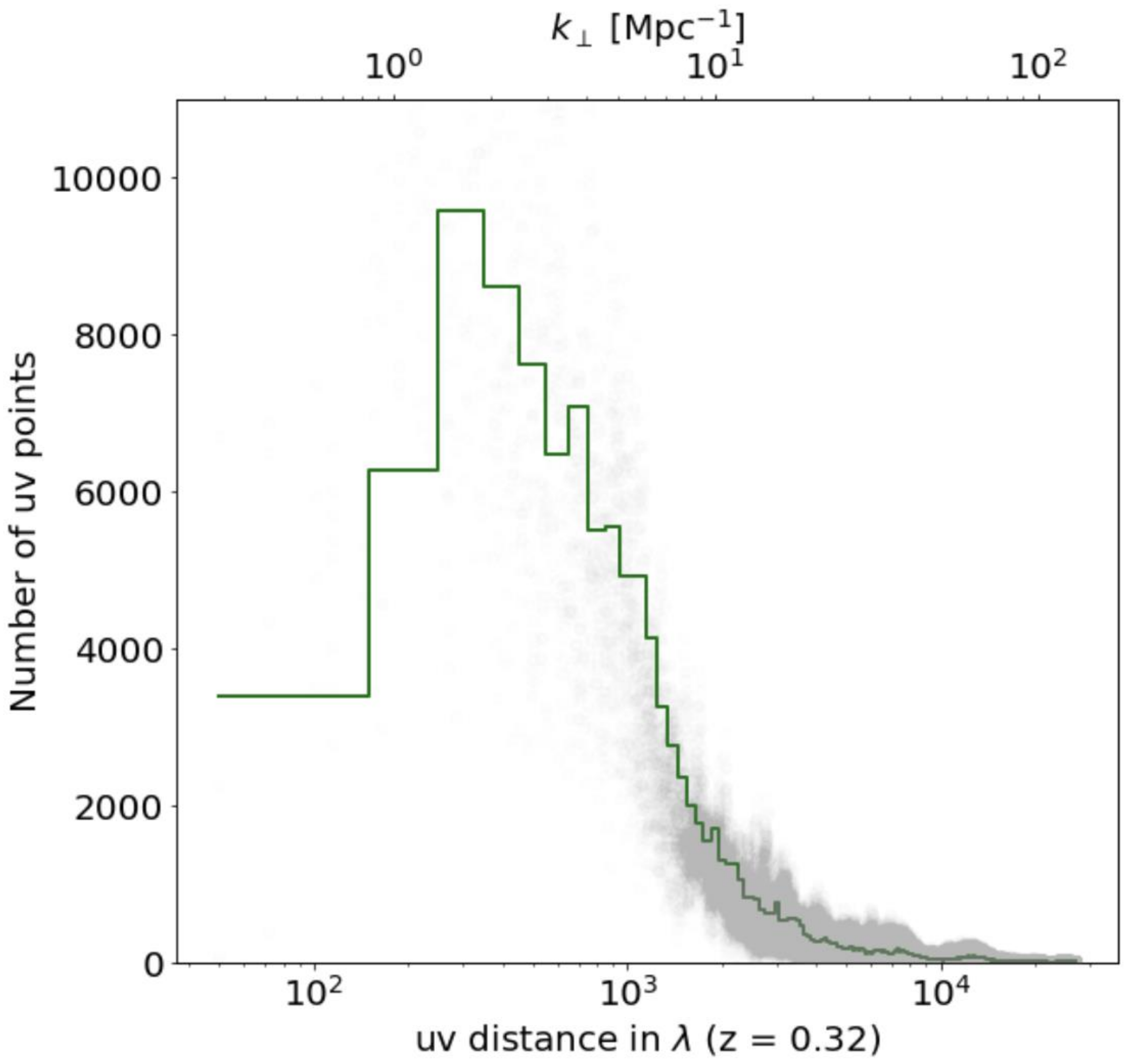
The halo model and HI mass function at z~0.32





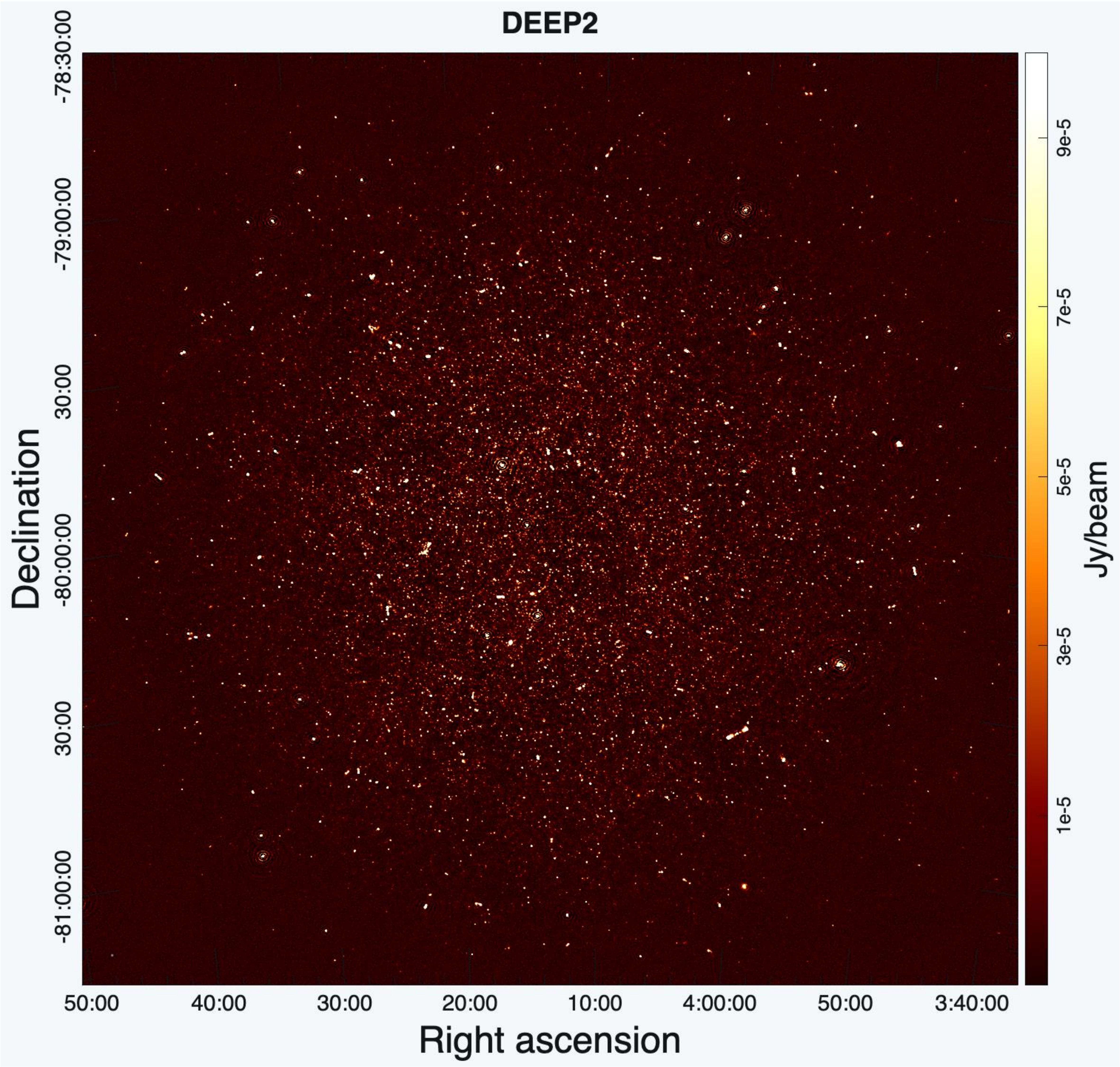
MeerKAT UV coverage





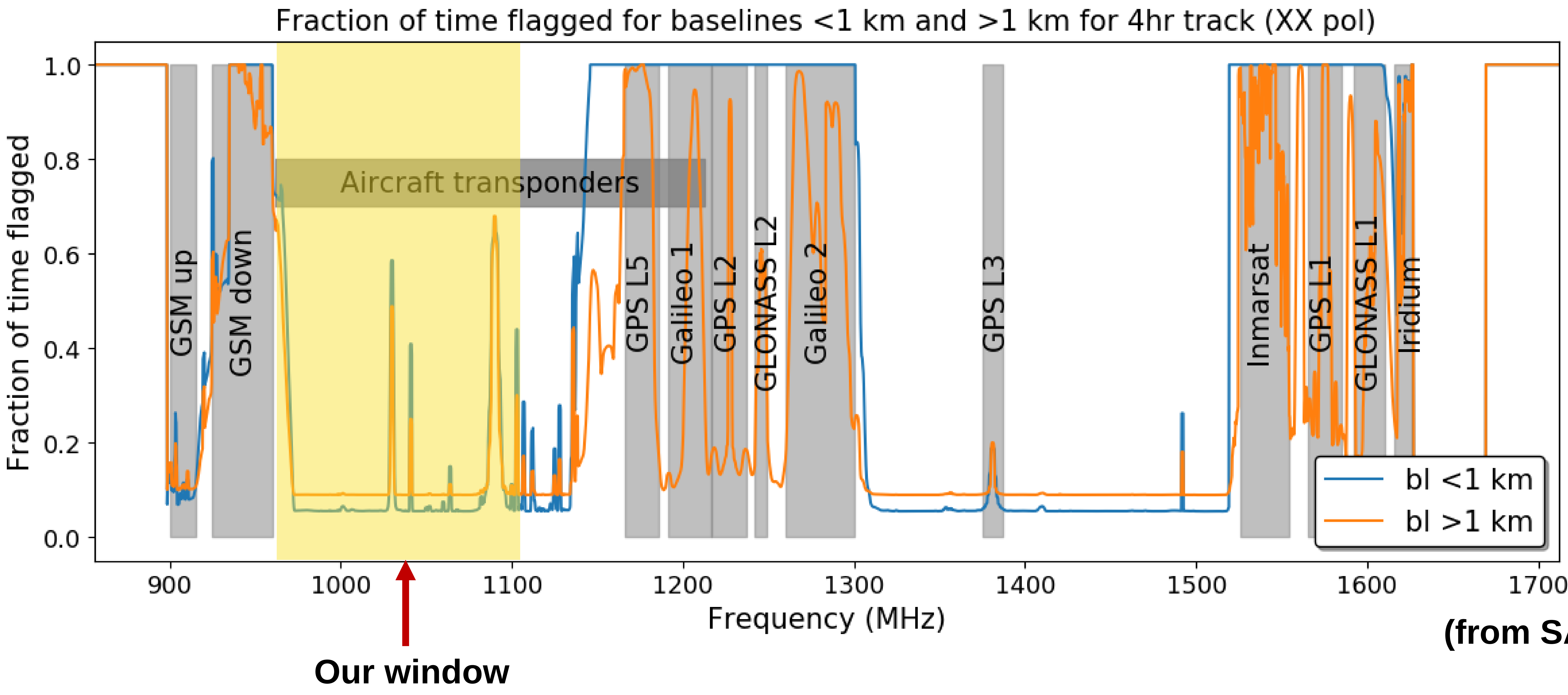
“Deep 2” field

- Deepest (~ 100 hours) MeerKAT pointing until recently (public commissioning data from 2018)
- α = 04h13m26.4s, δ = −80°0′0″
- Very clean field (no strong point sources)
- 0.21 MHz and 8s resolution
- Used to calculate continuum source counts using P(D) analysis down to 0.25 uJy (arXiv:1912.06212, arXiv:2101.07827)



RMS: 3 µJy/beam

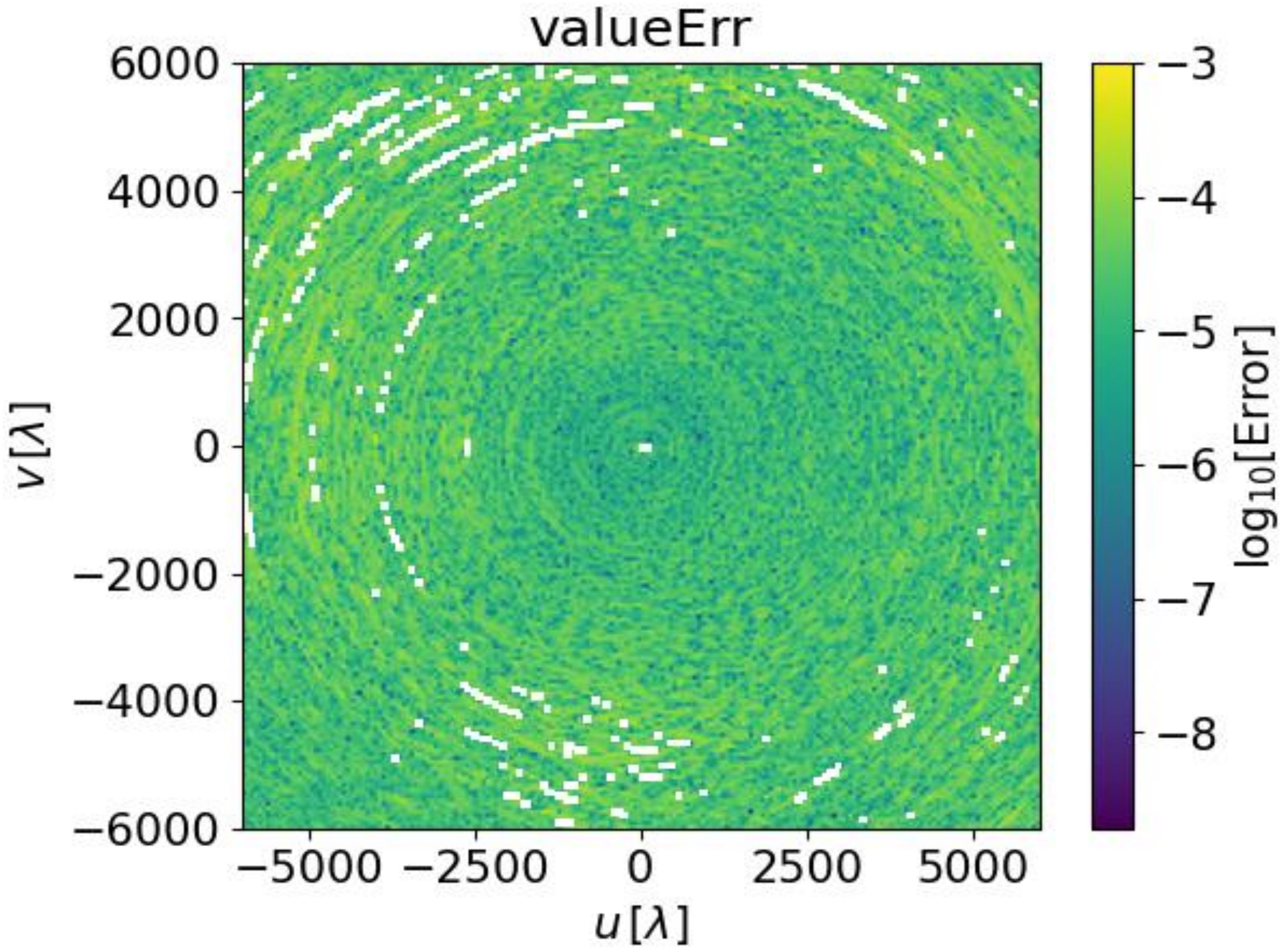
RFI for MeerKAT/SKA-MID…



(from SARAO)

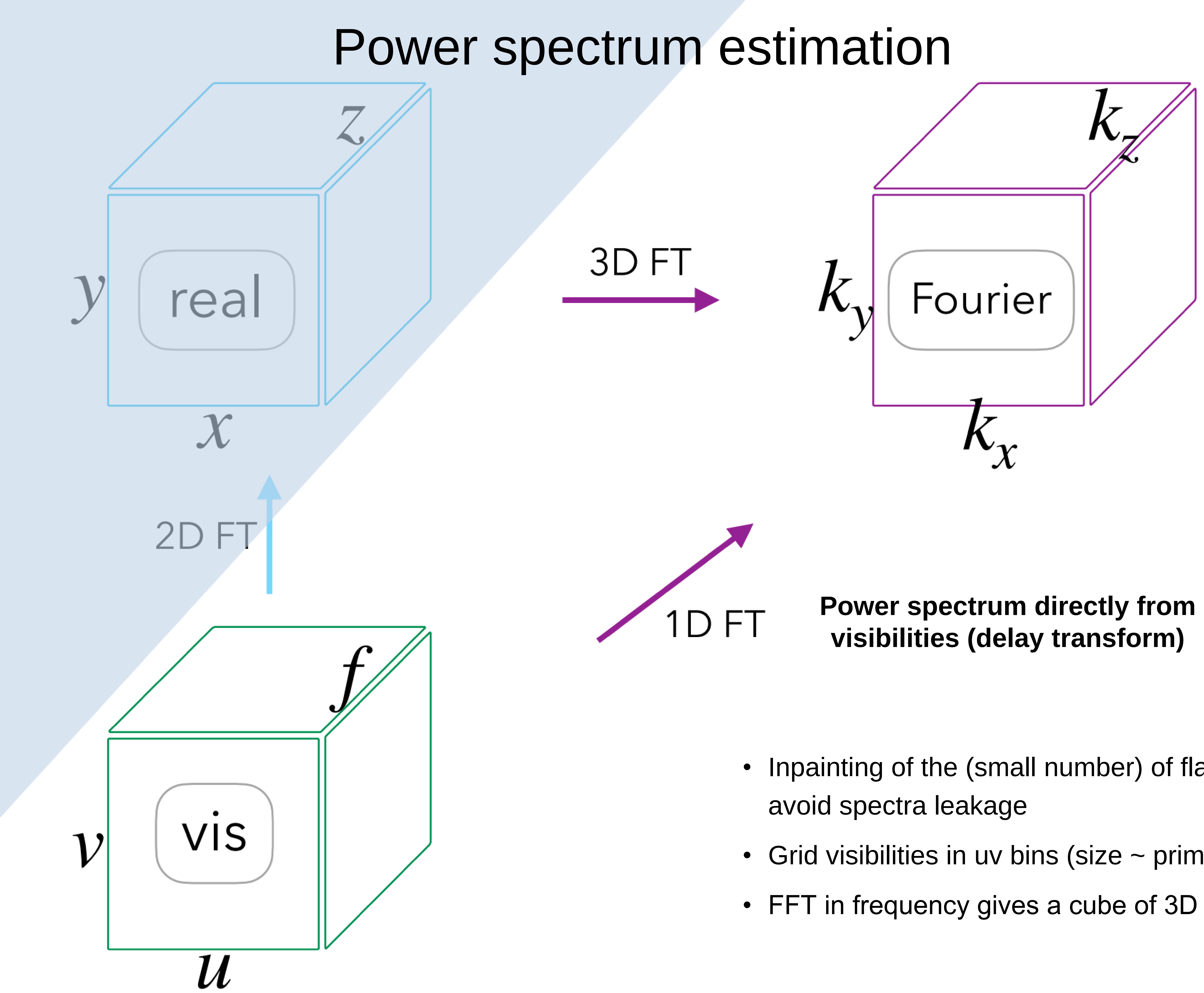
Calibration

- 96 hrs of data spanned through 9 separate observing sessions (data blocks)
- RFI flagging and calibration done using the **processMeerKAT** (Collier et al. 2021) software
- Primary calibrator PKS B1934-638 is observed for 10 minutes every 3 hours for flux and bandpass calibration (main source of error)
- Secondary calibrator is PKS J0252-7104 is observed for 2 minutes every 15 minutes for phase calibration.
- For each data block, an iterative self-calibration is performed for further phase corrections.
- Model visibilities produced from ***clean*** process (our point source/foreground model)
- Full stokes data produced



Calibration error ~ 10^{-5}

Power spectrum estimation



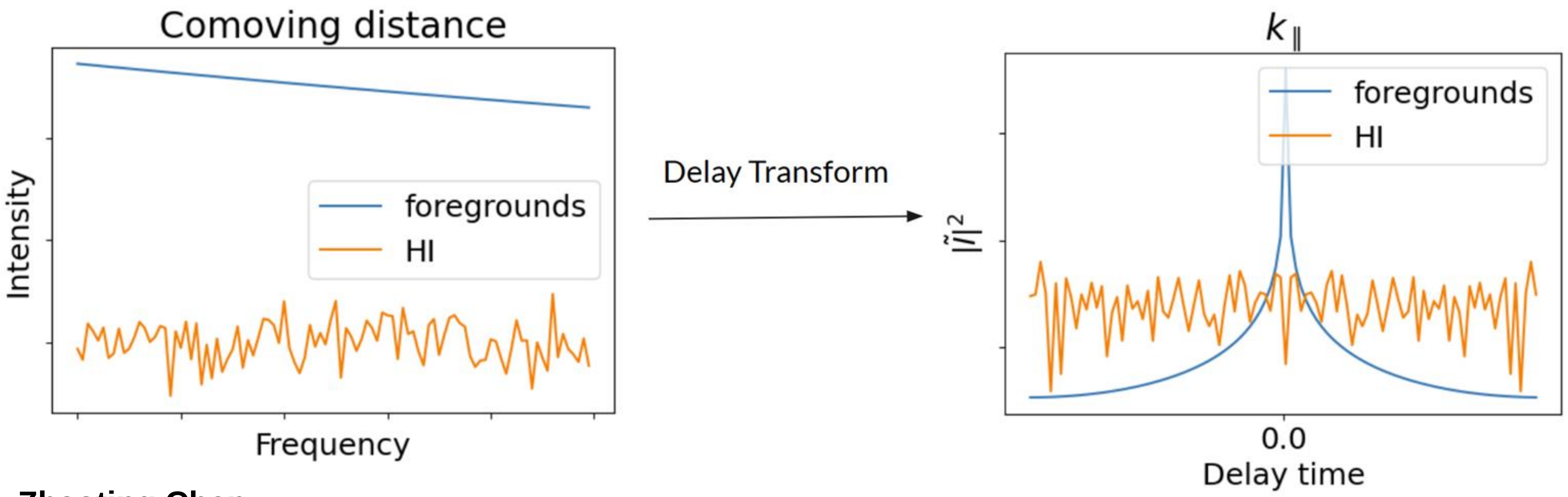
Power spectrum directly from visibilities (delay transform)

- Inpainting of the (small number) of flagged channels to avoid spectra leakage
- Grid visibilities in uv bins (size ~ primary beam)
- FFT in frequency gives a cube of 3D power spectra…

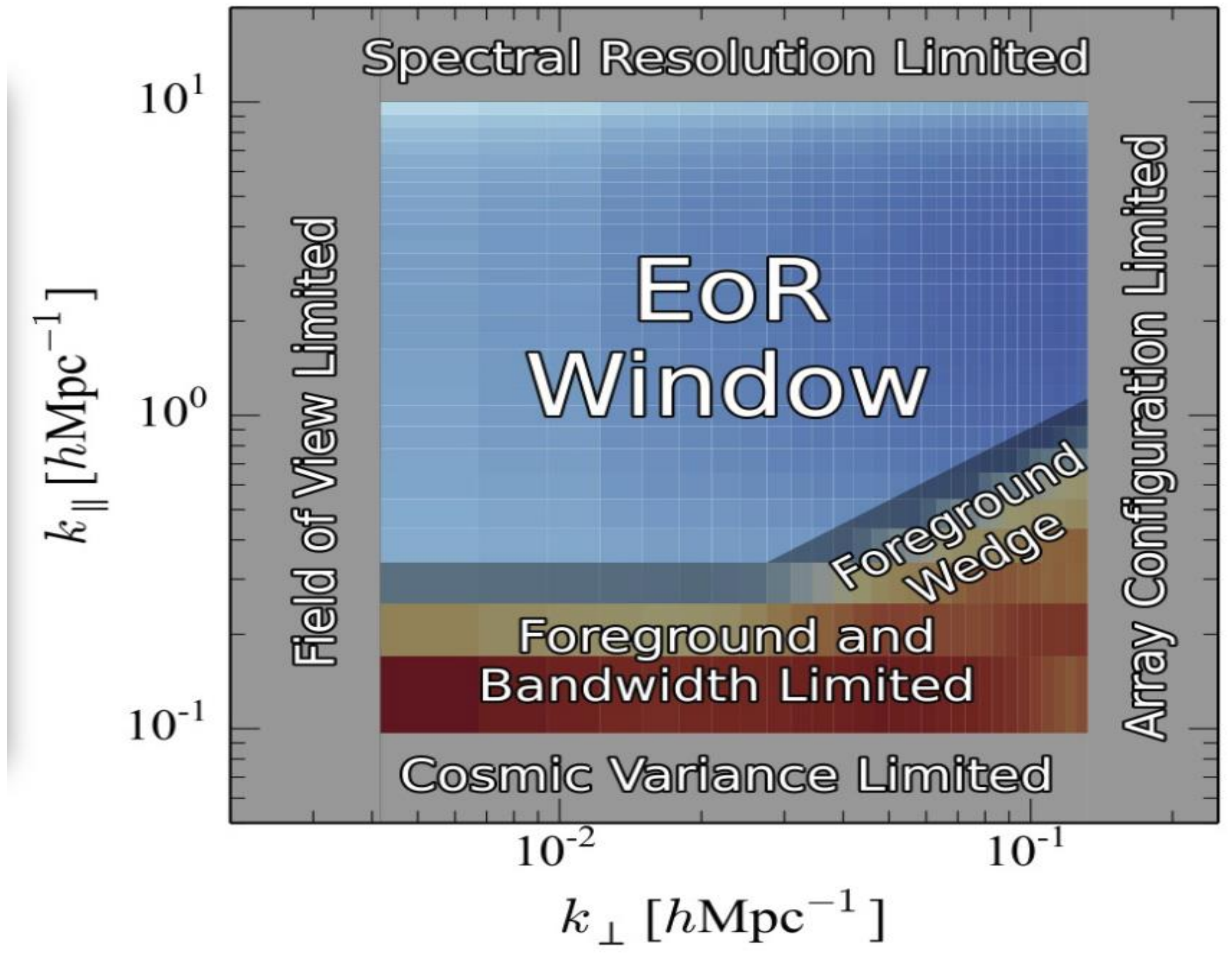
(Sourabh Paul)

Foreground avoidance

- Radio foregrounds are orders of magnitude brighter than HI signal
- Pick a region that is “foreground free”



Zhaoting Chen

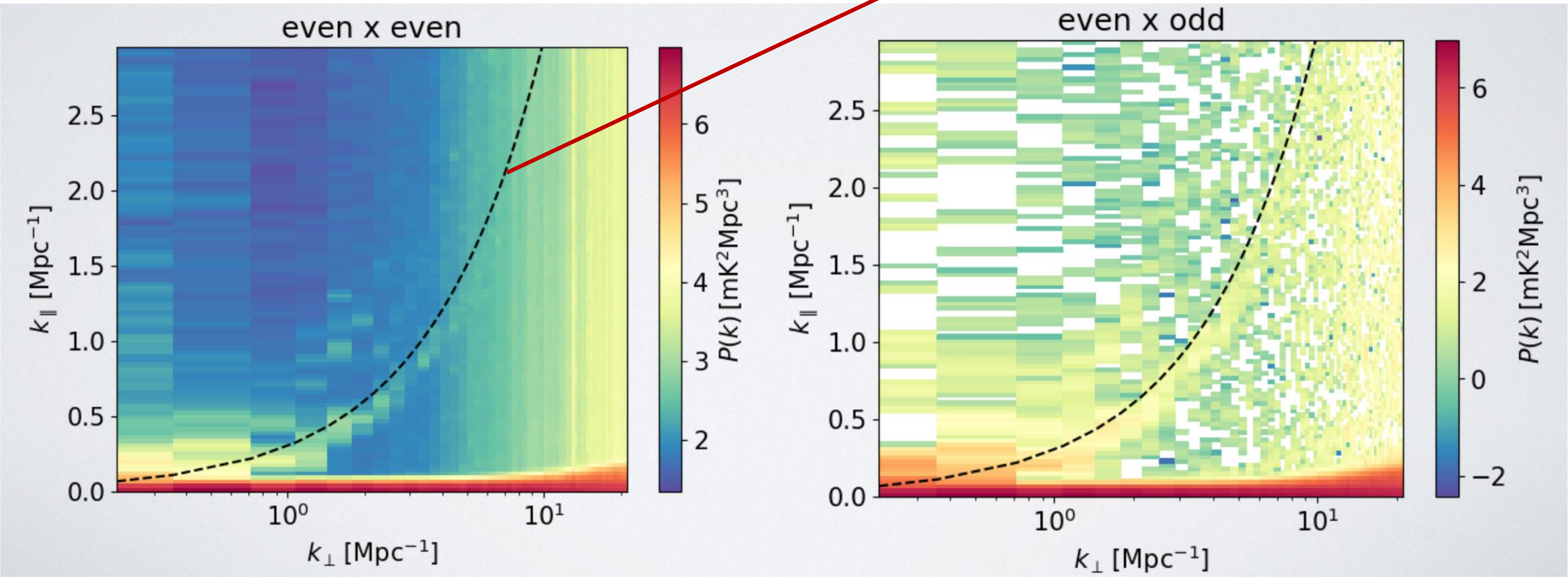


Liu, Parsons & Trott, 1404.2596

2d Power Spectra

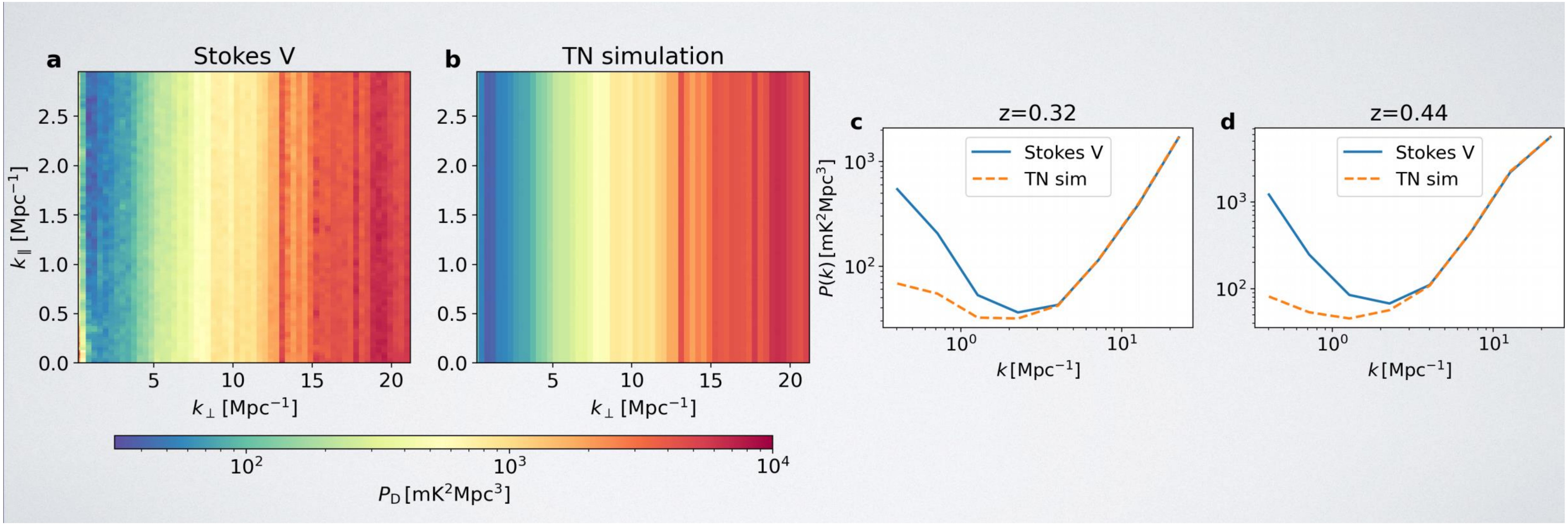
- A clear structure of contamination in the cylindrical power spectrum
- How to remove it?

Foreground wedge (horizon)

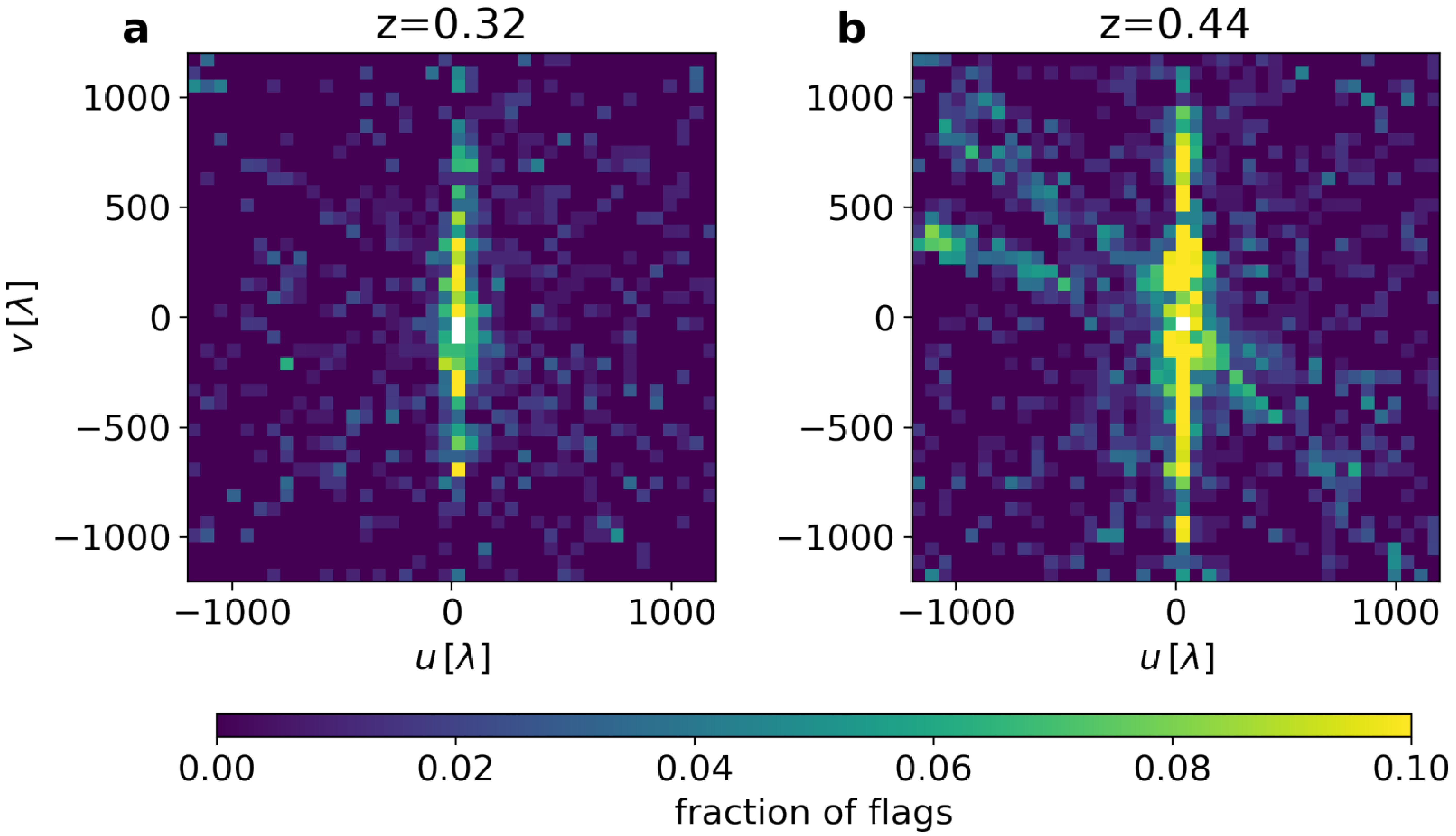


Thermal noise

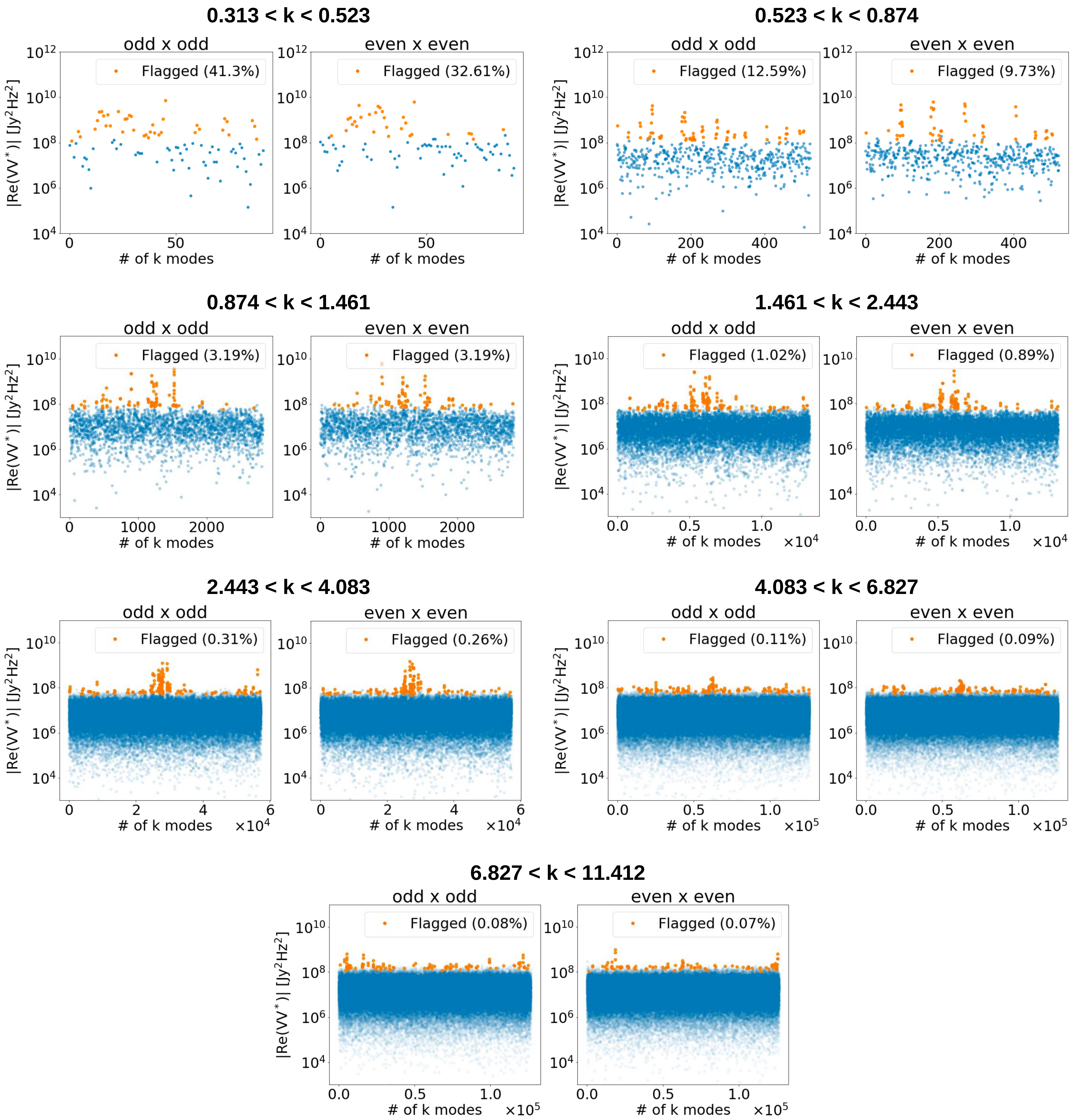
- Use Stokes V data to estimate and simulate thermal noise (end to end simulations developed in Paul et al. https://arxiv.org/abs/2009.13550)
- Flag strong deviations from thermal noise (5 sigma cut) – tested with simulations



Fraction of k_{par} pixels flagged



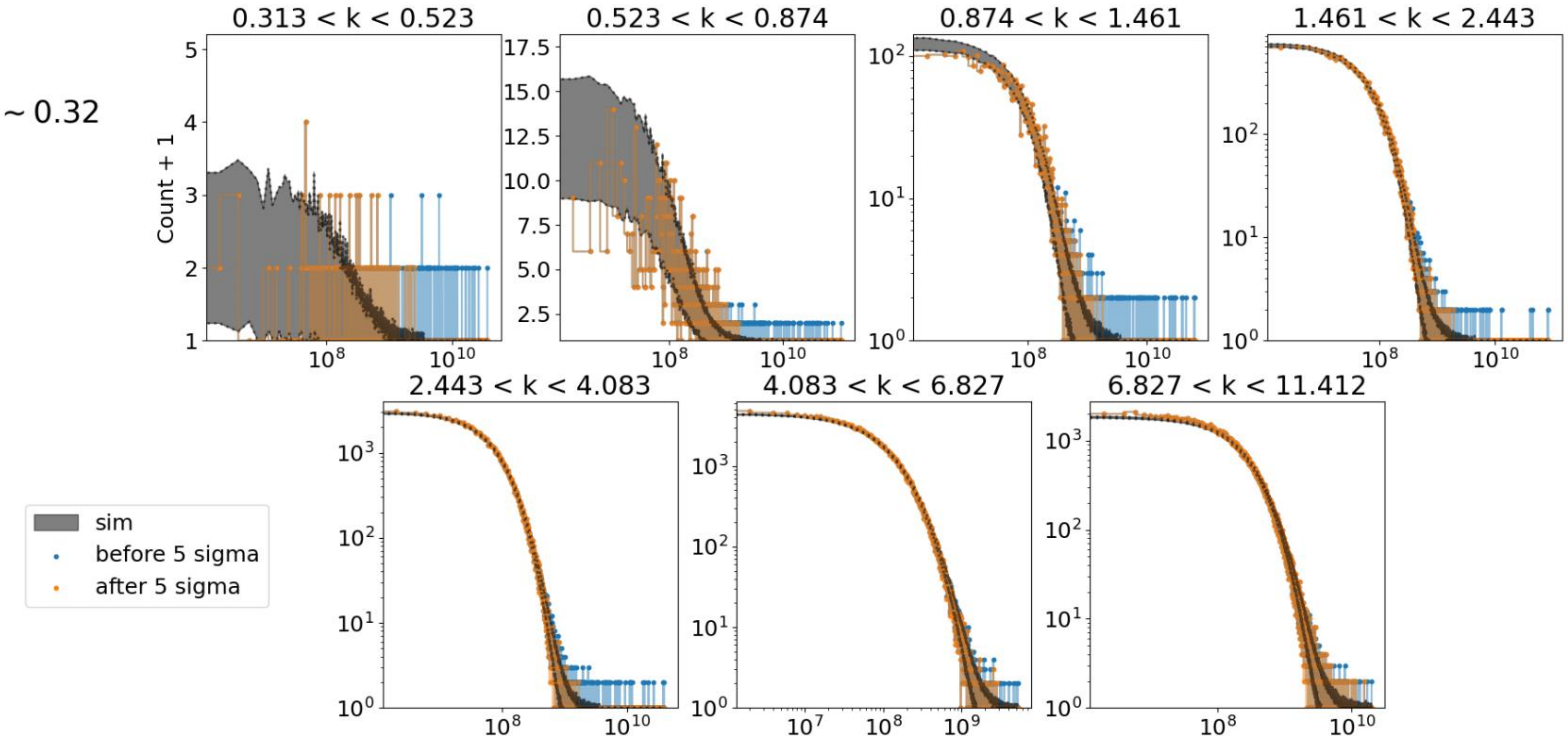
Using thermal noise simulations, we can identify the k-pixels that are outliers and flag them. Tracing these k-pixels, we find that the excess power comes from a weak wide-band RFI which shows up at u=0.

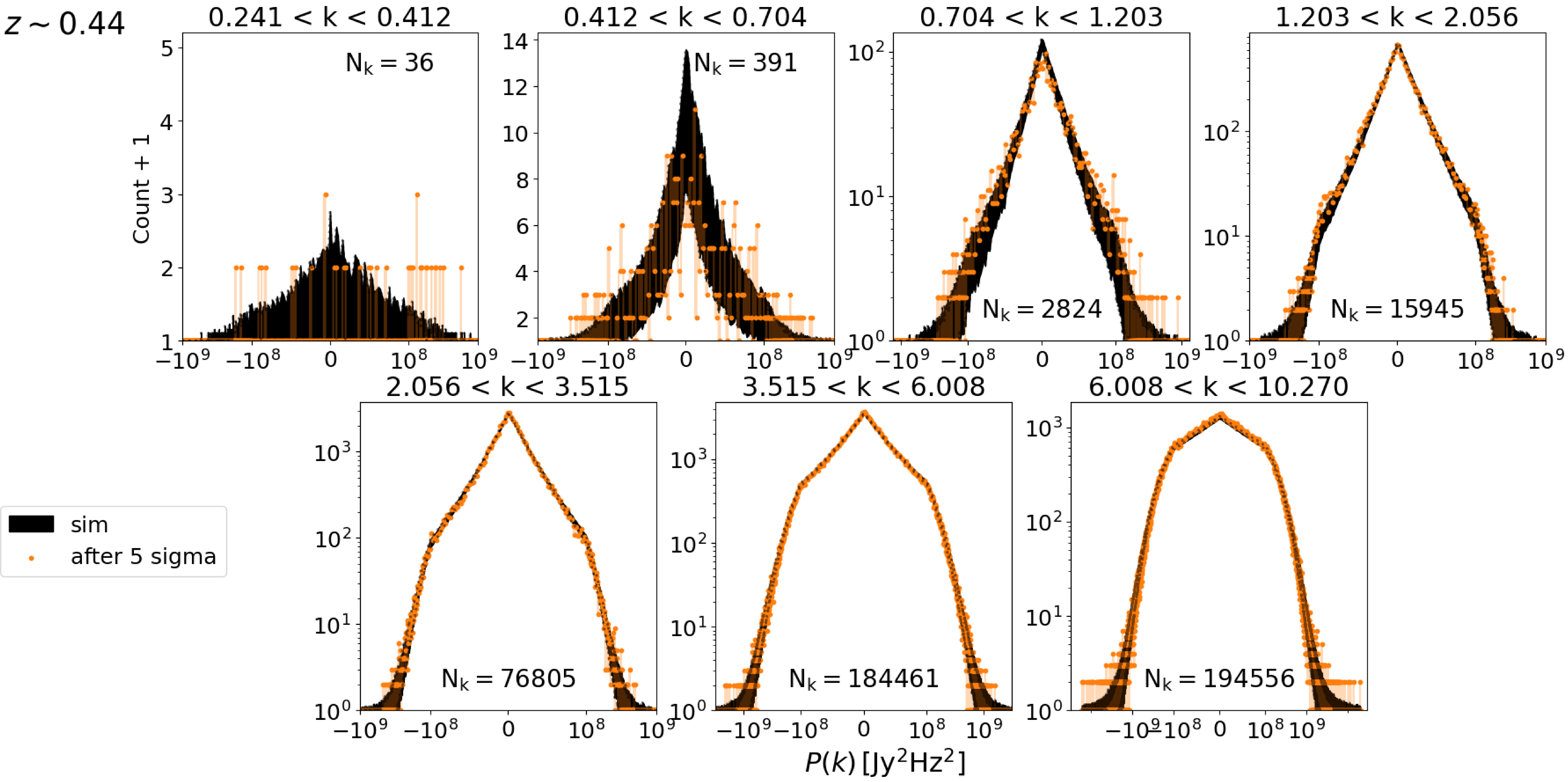


0.313 < k < 0.523
odd x odd
Flagged (41.3%)
even x even
Flagged (32.61%)
0.523 < k < 0.874
odd x odd
Flagged (12.59%)
even x even
Flagged (9.73%)
0.874 < k < 1.461
odd x odd
Flagged (3.19%)
even x even
Flagged (3.19%)
1.461 < k < 2.443
odd x odd
Flagged (1.02%)
even x even
Flagged (0.89%)
2.443 < k < 4.083
odd x odd
Flagged (0.31%)
even x even
Flagged (0.26%)
4.083 < k < 6.827
odd x odd
Flagged (0.11%)
even x even
Flagged (0.09%)
6.827 < k < 11.412
odd x odd
Flagged (0.08%)
even x even
Flagged (0.07%)
|Re(VV*)| [Jy²Hz²]
# of k modes

Flagging in the auto Pk

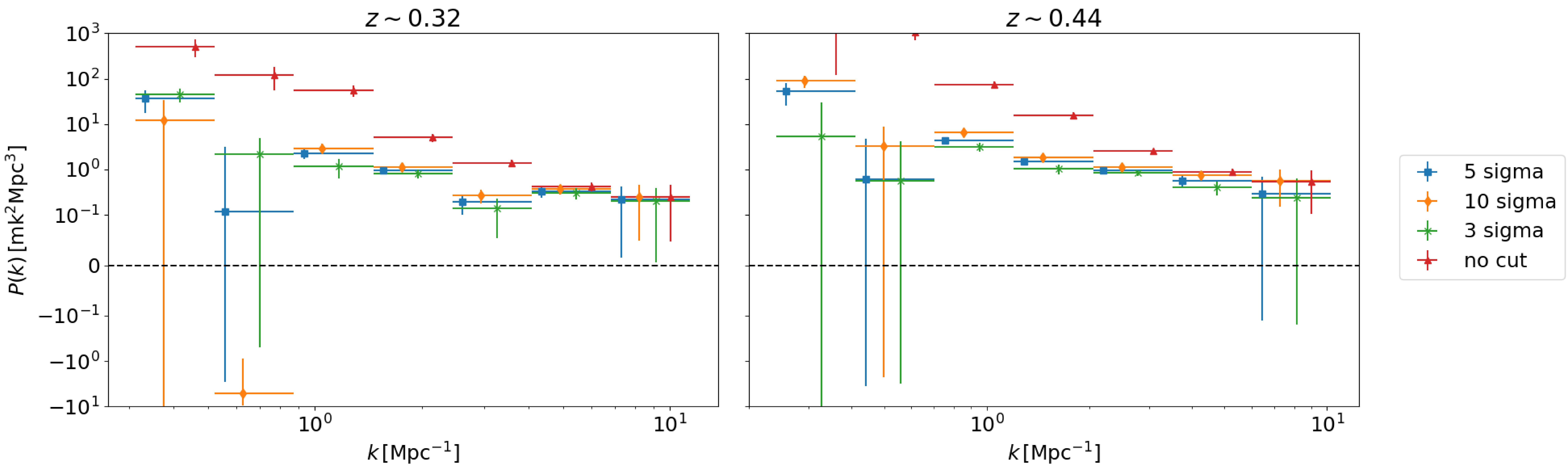
$z \sim 0.32$





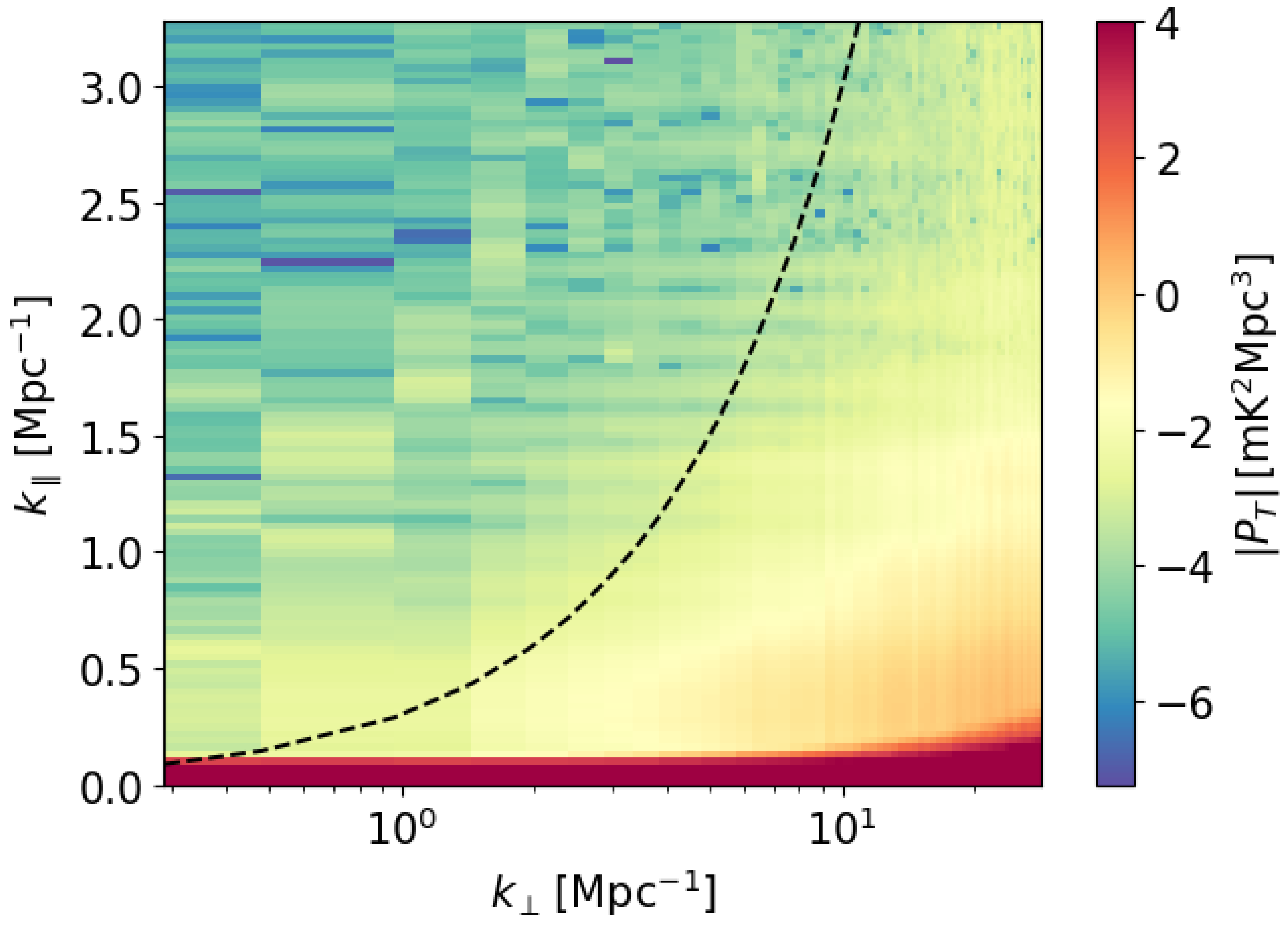
z ~ 0.44
0.241 < k < 0.412
N_k = 36
0.412 < k < 0.704
N_k = 391
0.704 < k < 1.203
N_k = 2824
1.203 < k < 2.056
N_k = 15945
2.056 < k < 3.515
N_k = 76805
3.515 < k < 6.008
N_k = 184461
6.008 < k < 10.270
N_k = 194556
Count + 1
P(k) [Jy²Hz²]
sim
after 5 sigma

Effect of 5 sigma cut

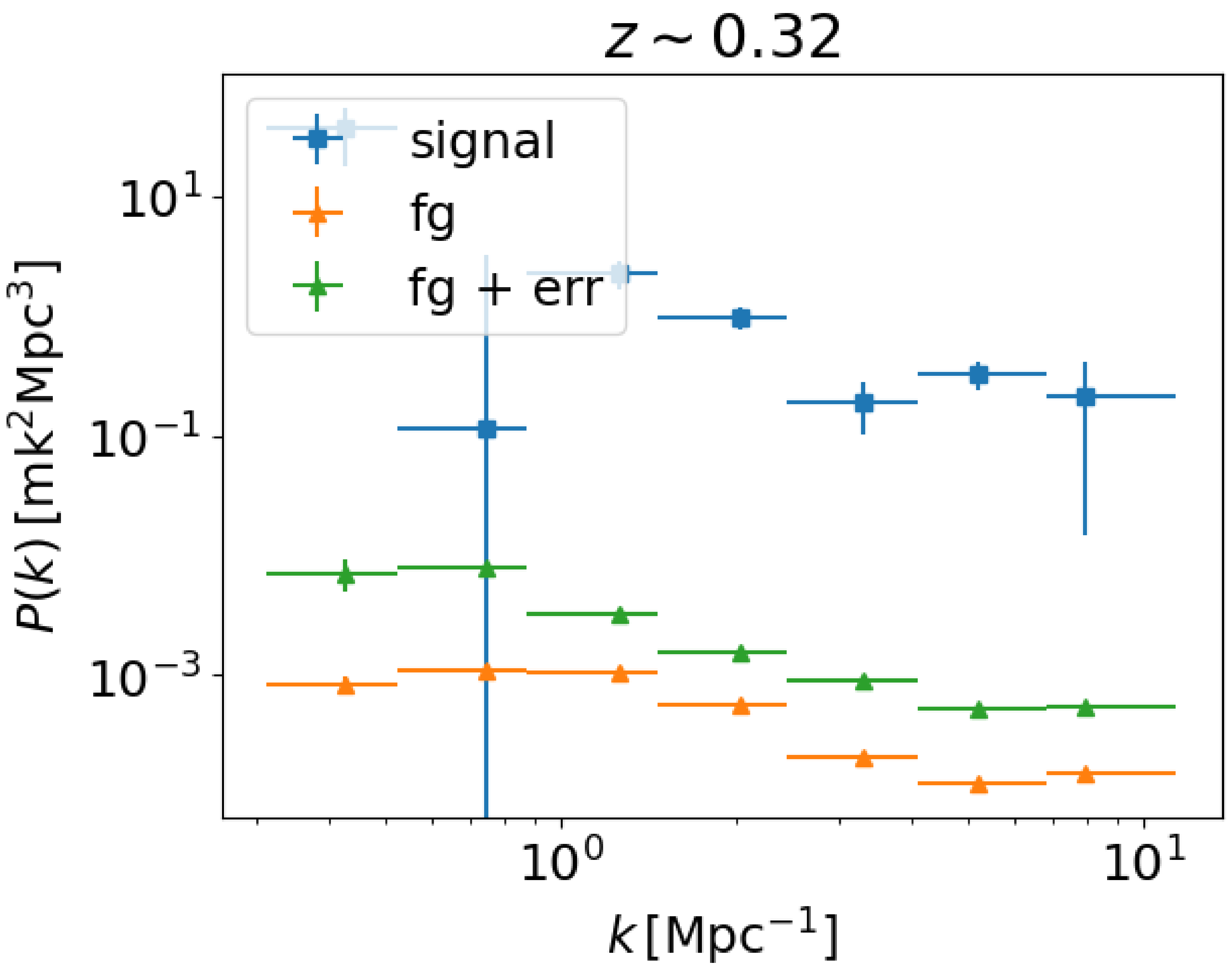


Foreground leakage?

The foreground power, with leakage from inpainting, RFI flagging, calibration error, is still insignificant compared to the measured signal -> no need for point source subtraction

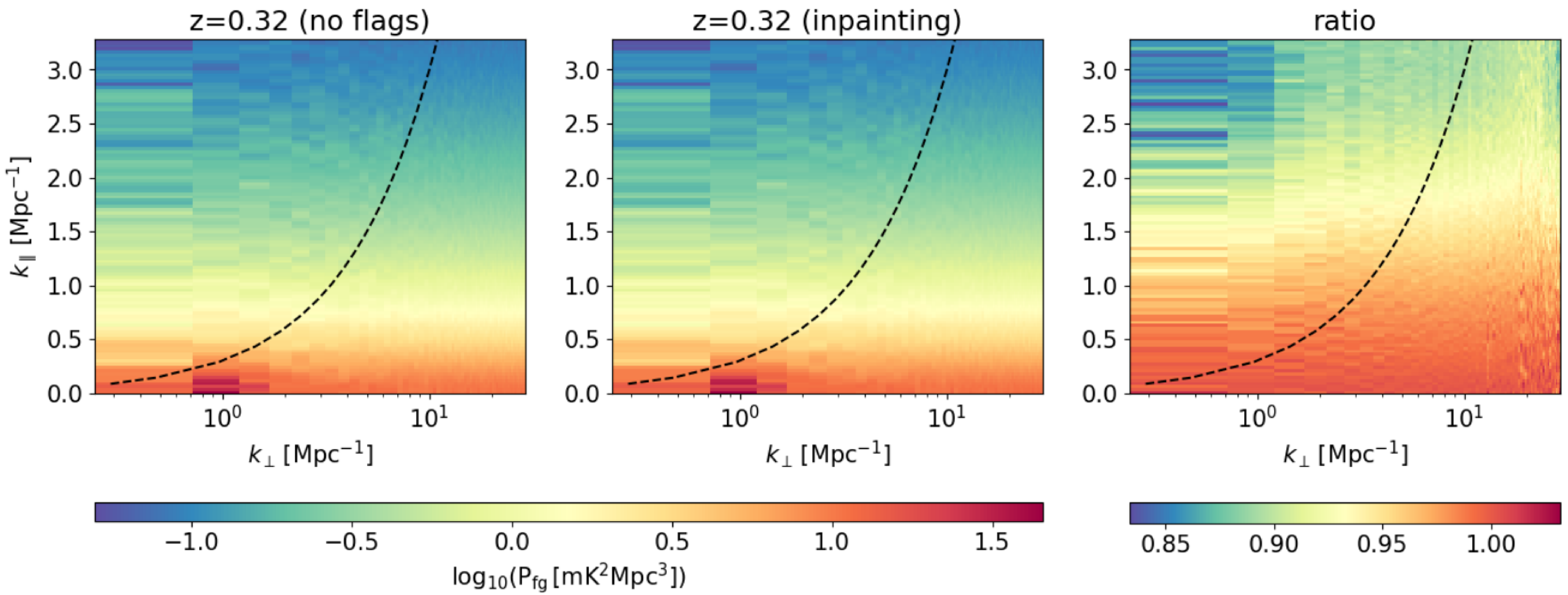


Foregrounds using point source model from Clean (color scale saturated)

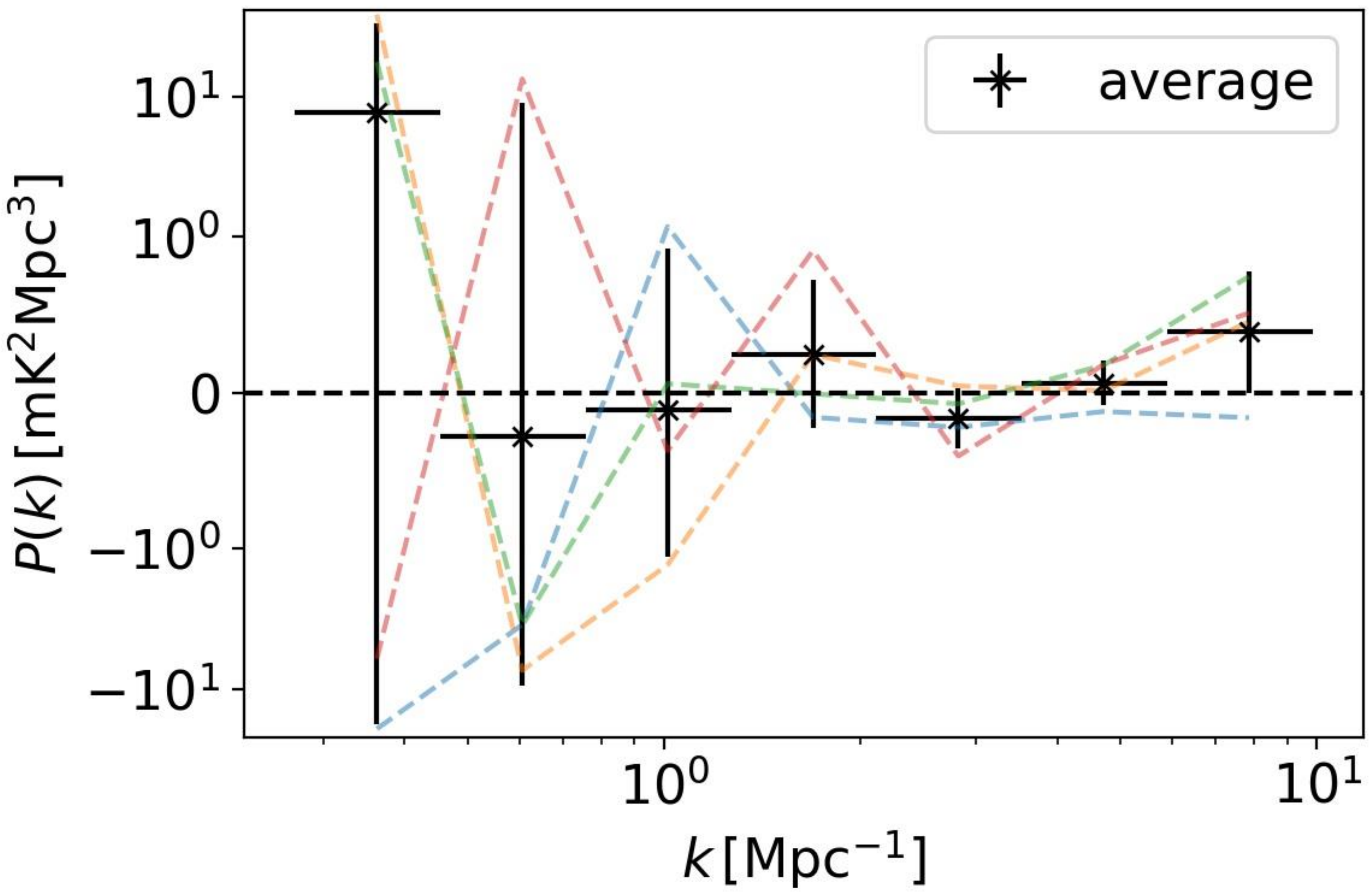


Foreground power spectrum

Effect of flagging and inpainting on the HI signal



Null test: cross-correlate visibility data between the two redshifts



No evidence of excess power

Jackknife of different data blocks