Enabling hydrogen intensity mapping

Isabella Paola Carucci

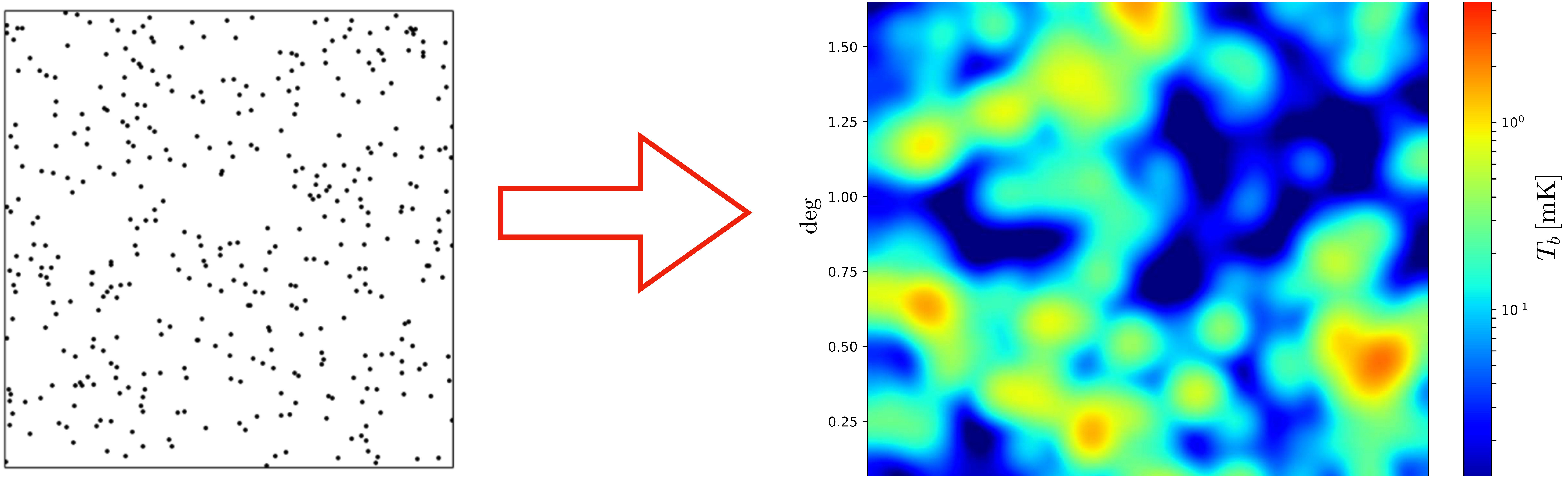
Università di Torino

INFN Torino – Istituto Nazionale di Fisica Nucleare, Sezione di Torino

8th November 2022

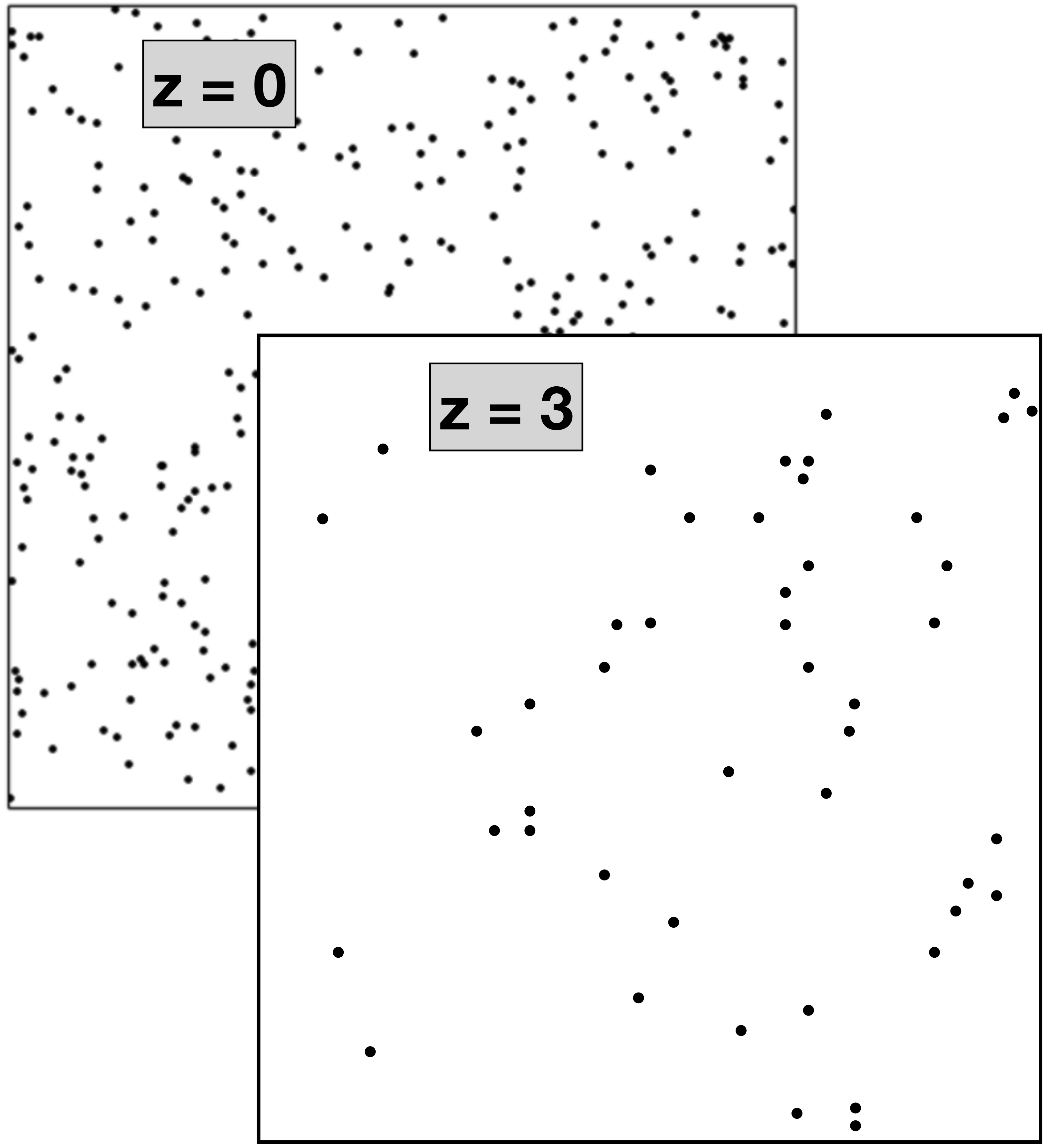
Paris-Saclay Astroparticle Symposium 2022

Hydrogen intensity mapping

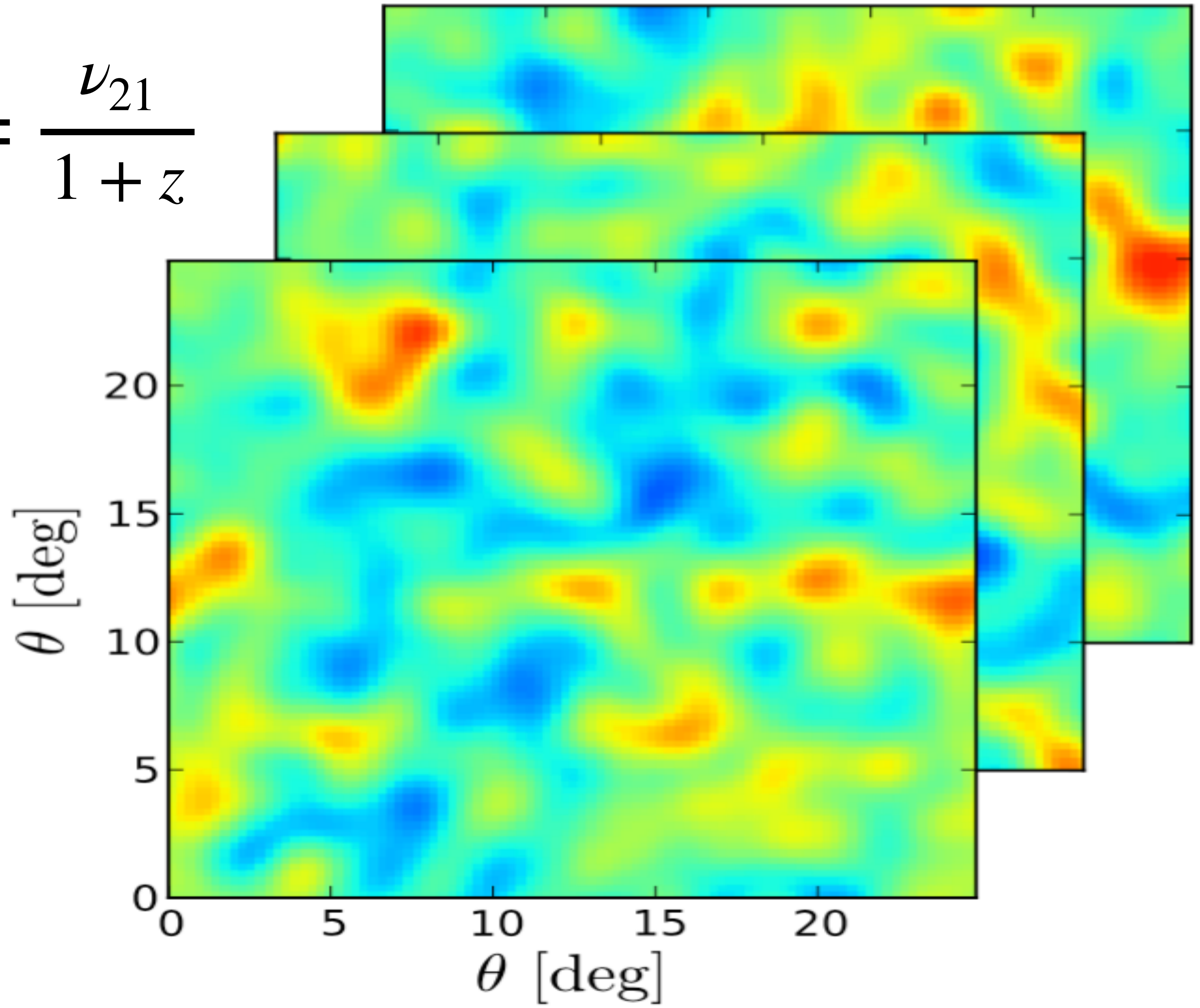


Put signal-to-noise where you really need it: **linear large scale modes**

Hydrogen intensity mapping

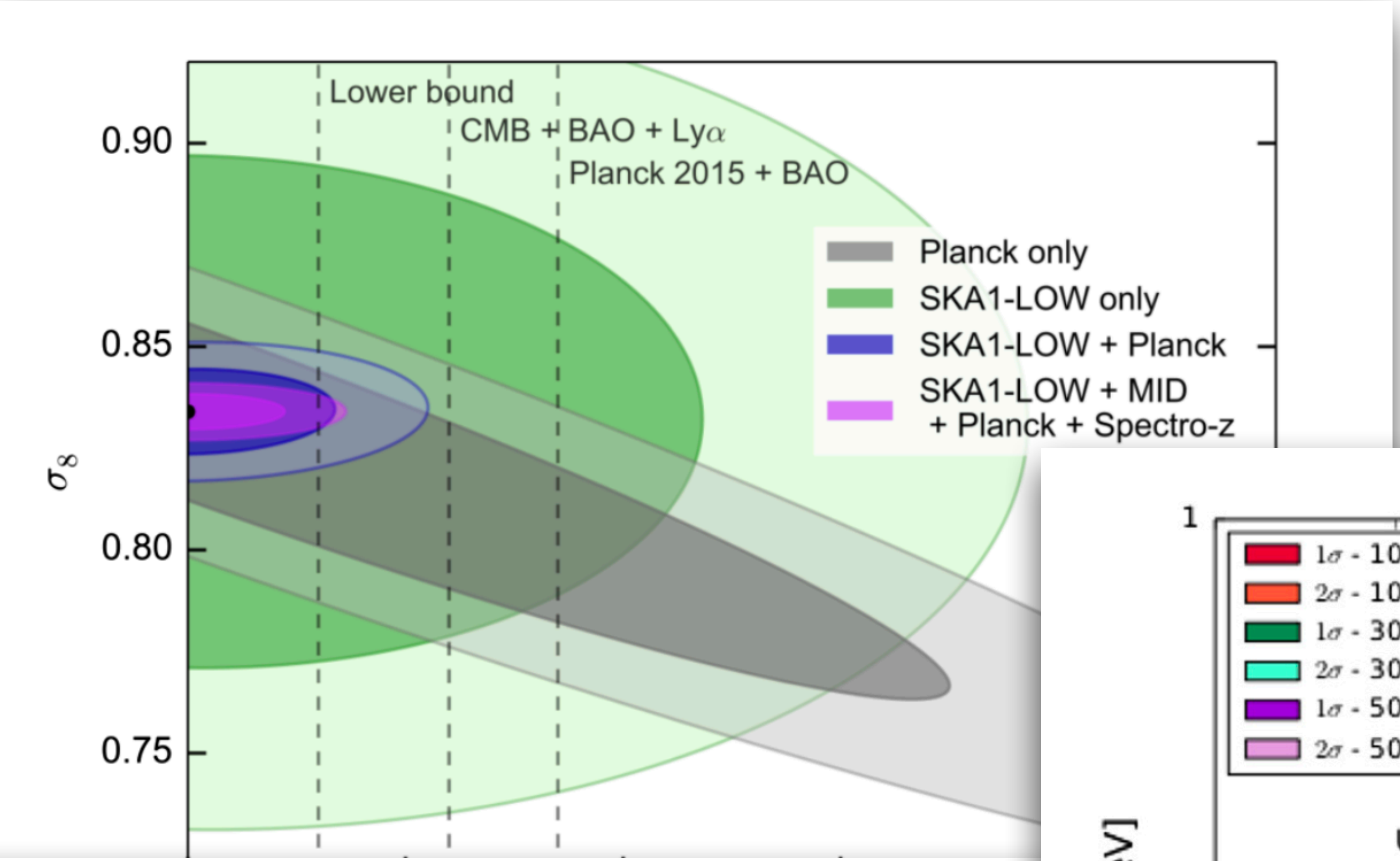


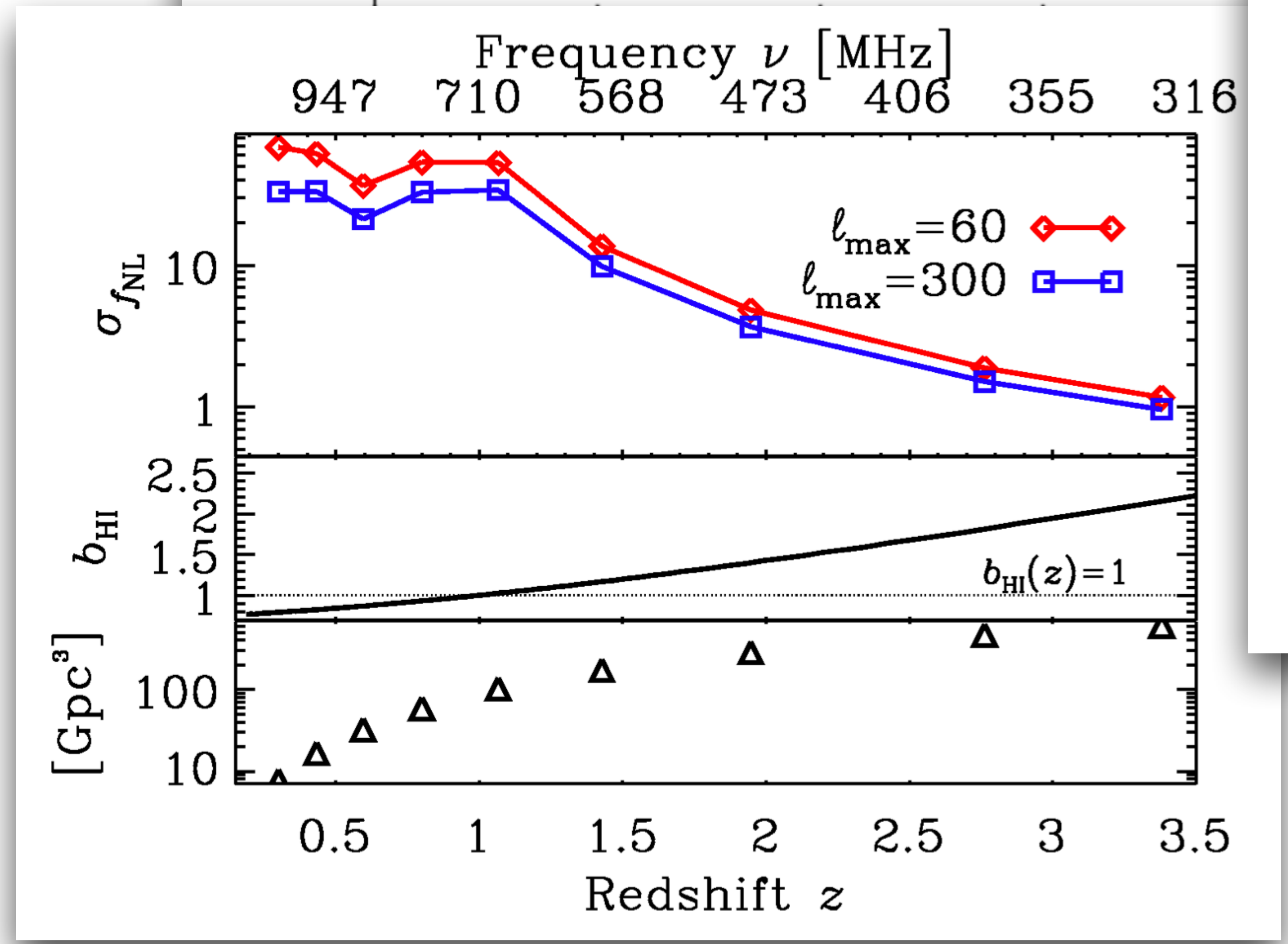
$$\nu = \frac{\nu_{21}}{1+z}$$

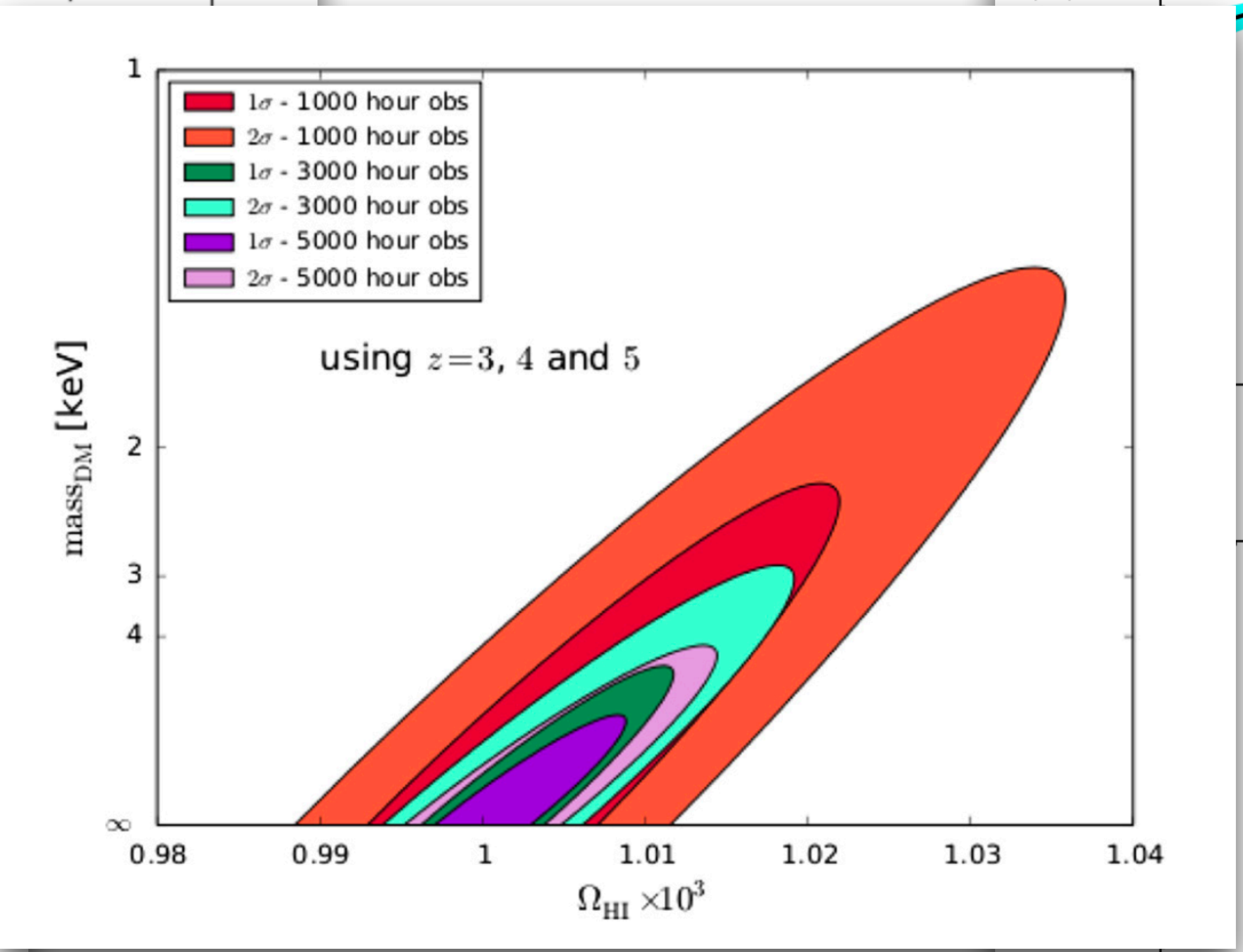


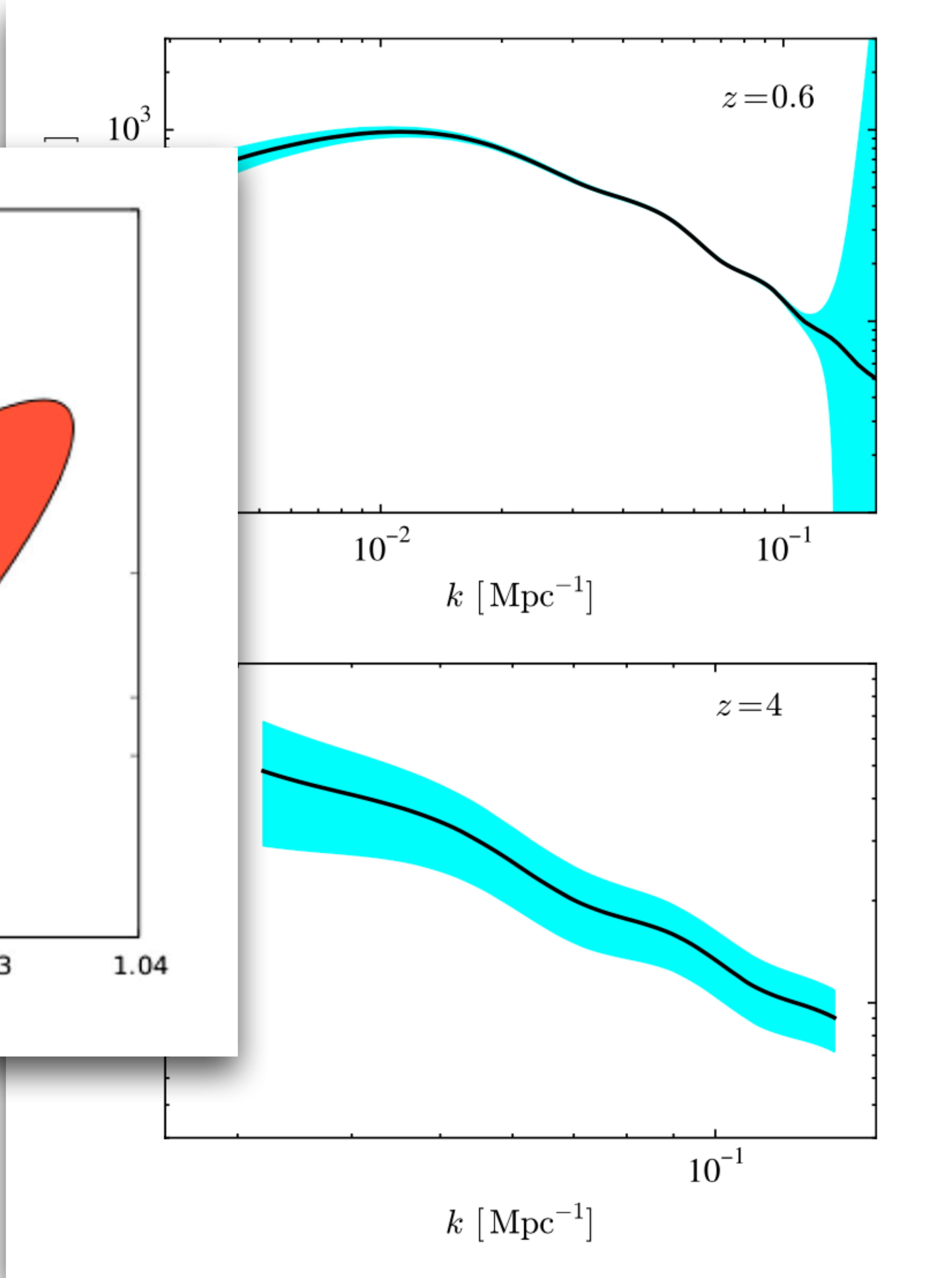
Big volumes (for cheap) and high redshift resolution

All plots from “Cosmology with Phase 1 of the Square Kilometre Array **Red Book”,** 2018



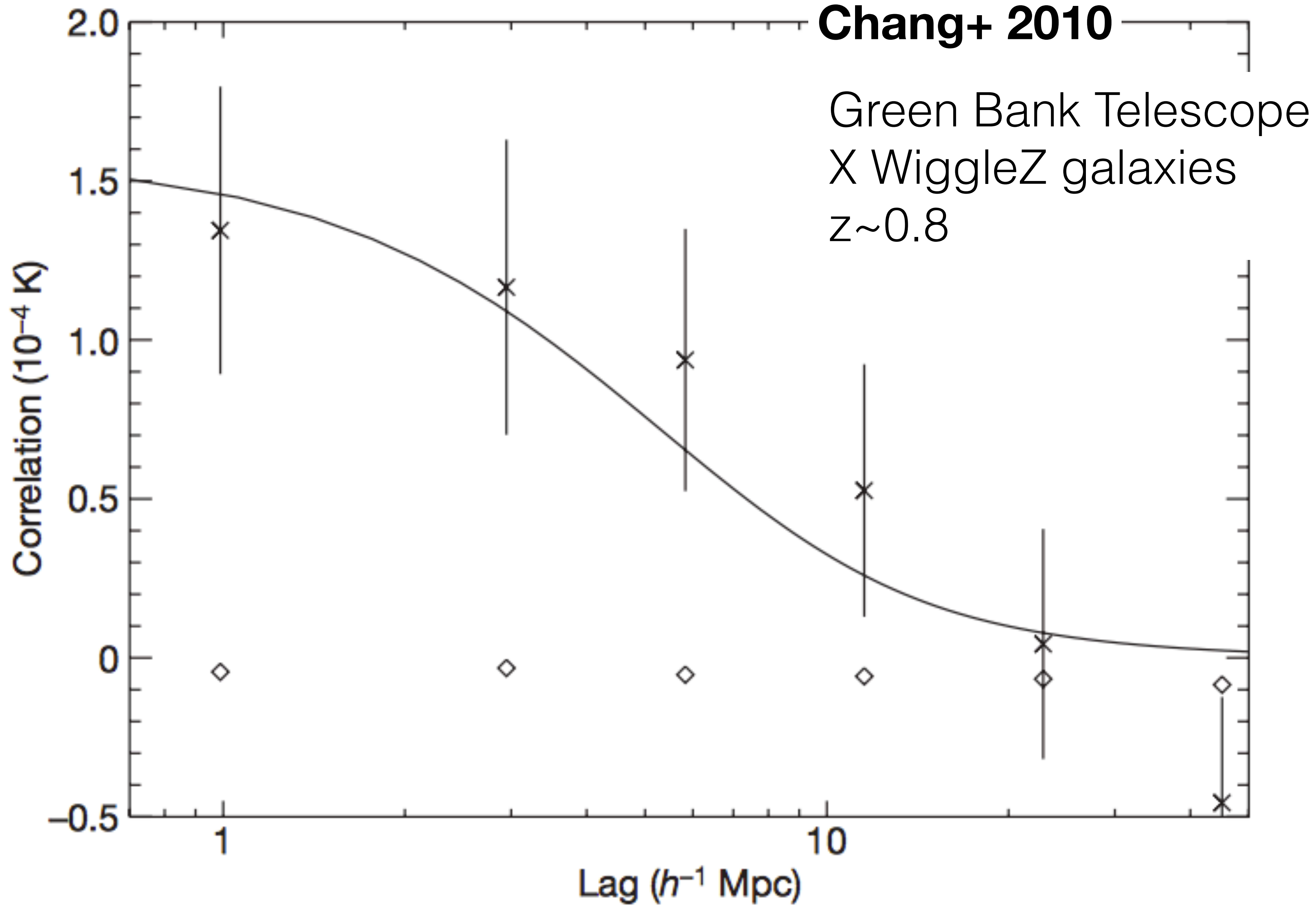




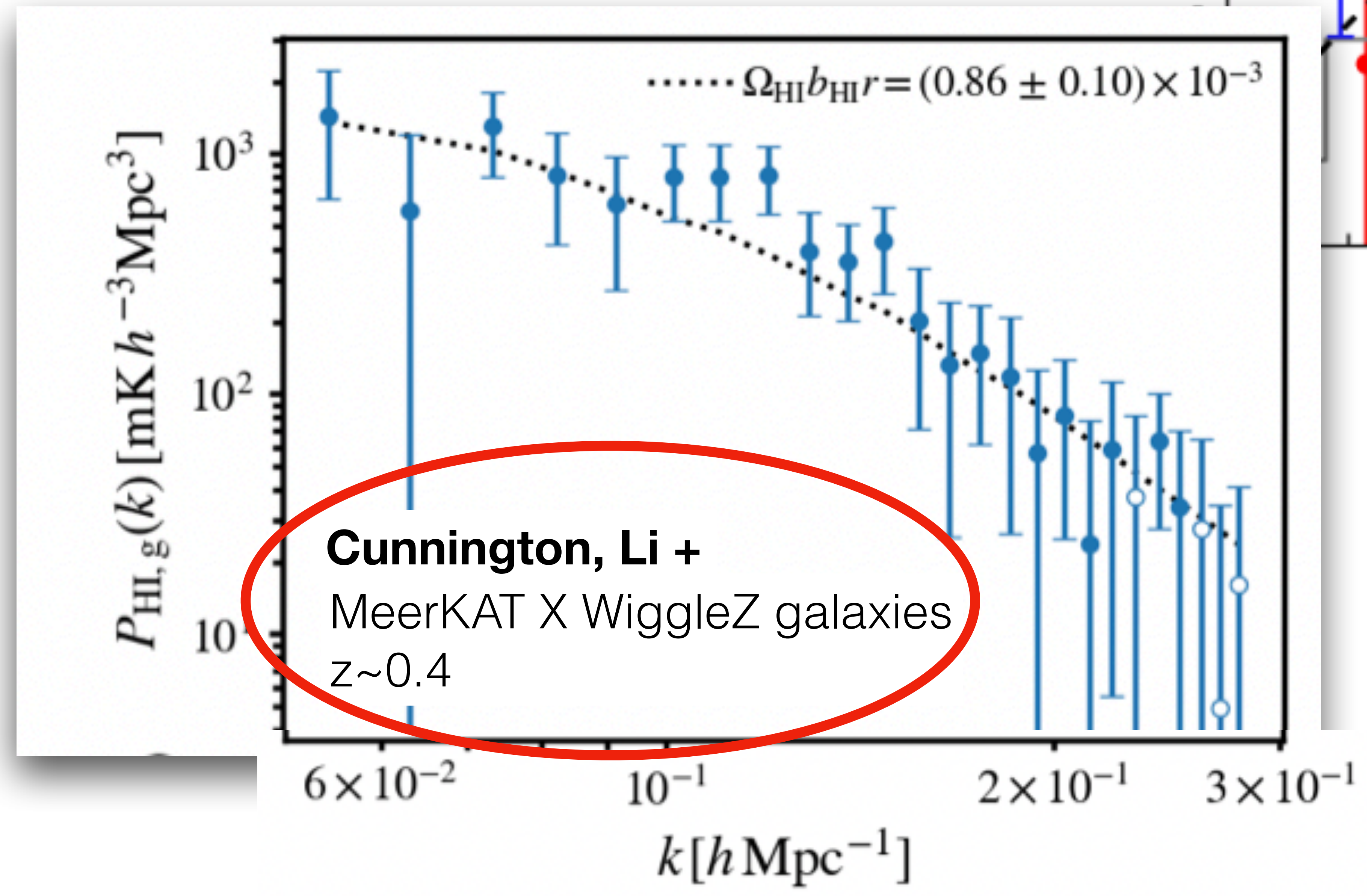


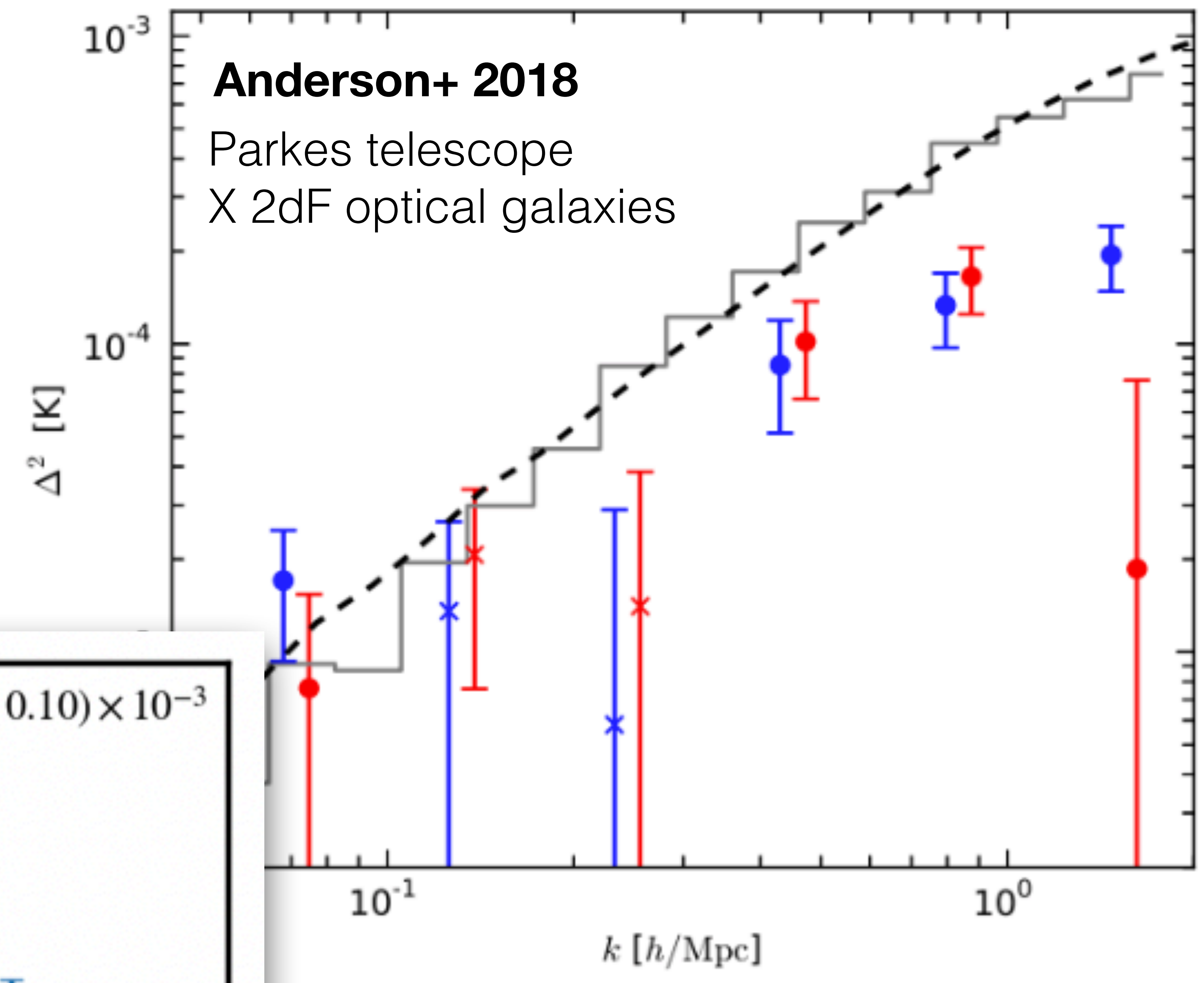
HI intensity mapping

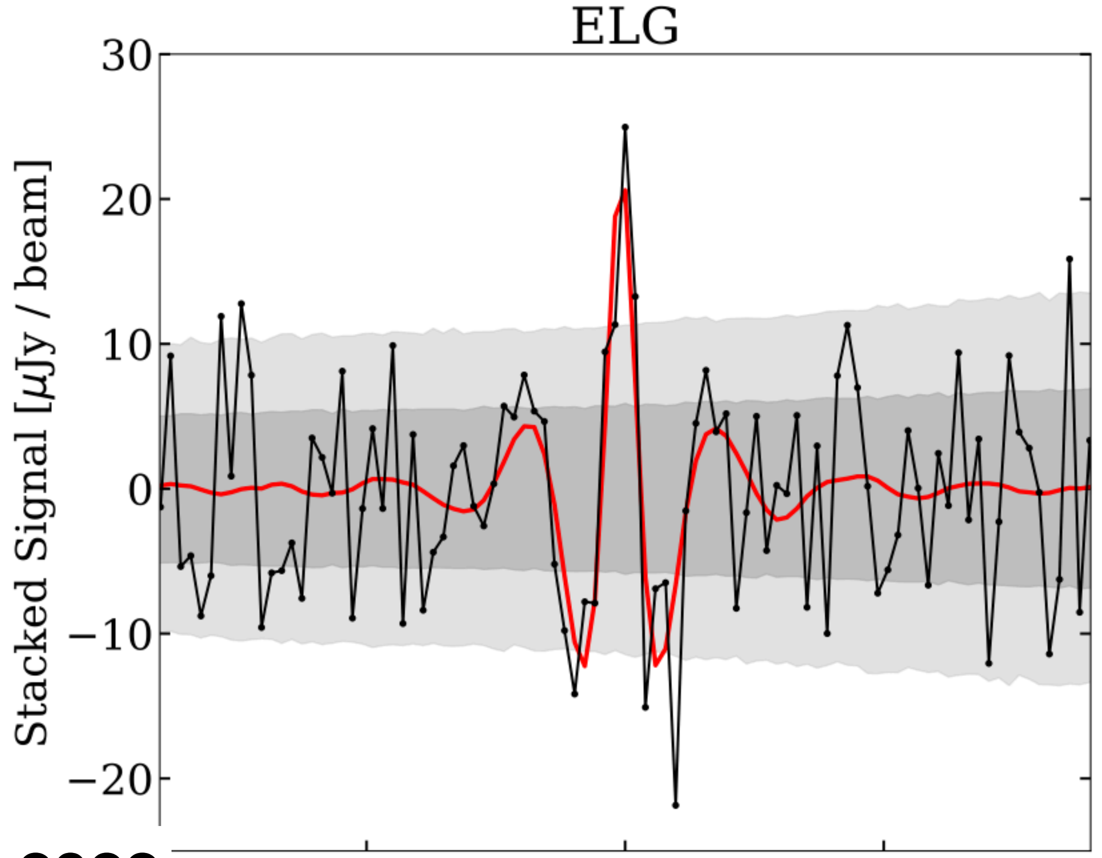
State-of-the-art



also Masui+ 2013, Switzer+ 2013,
Wolz+ 2017,2022





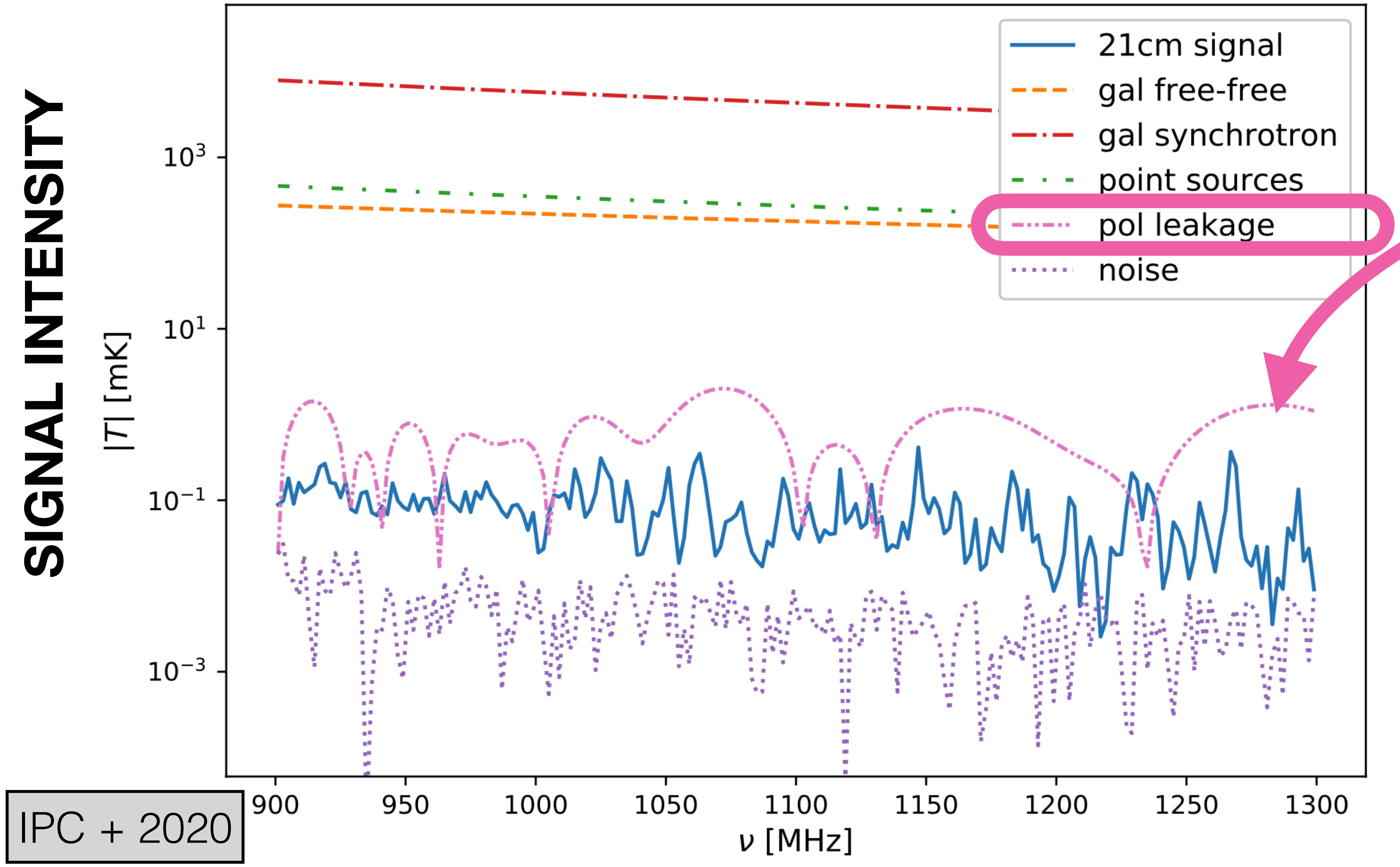


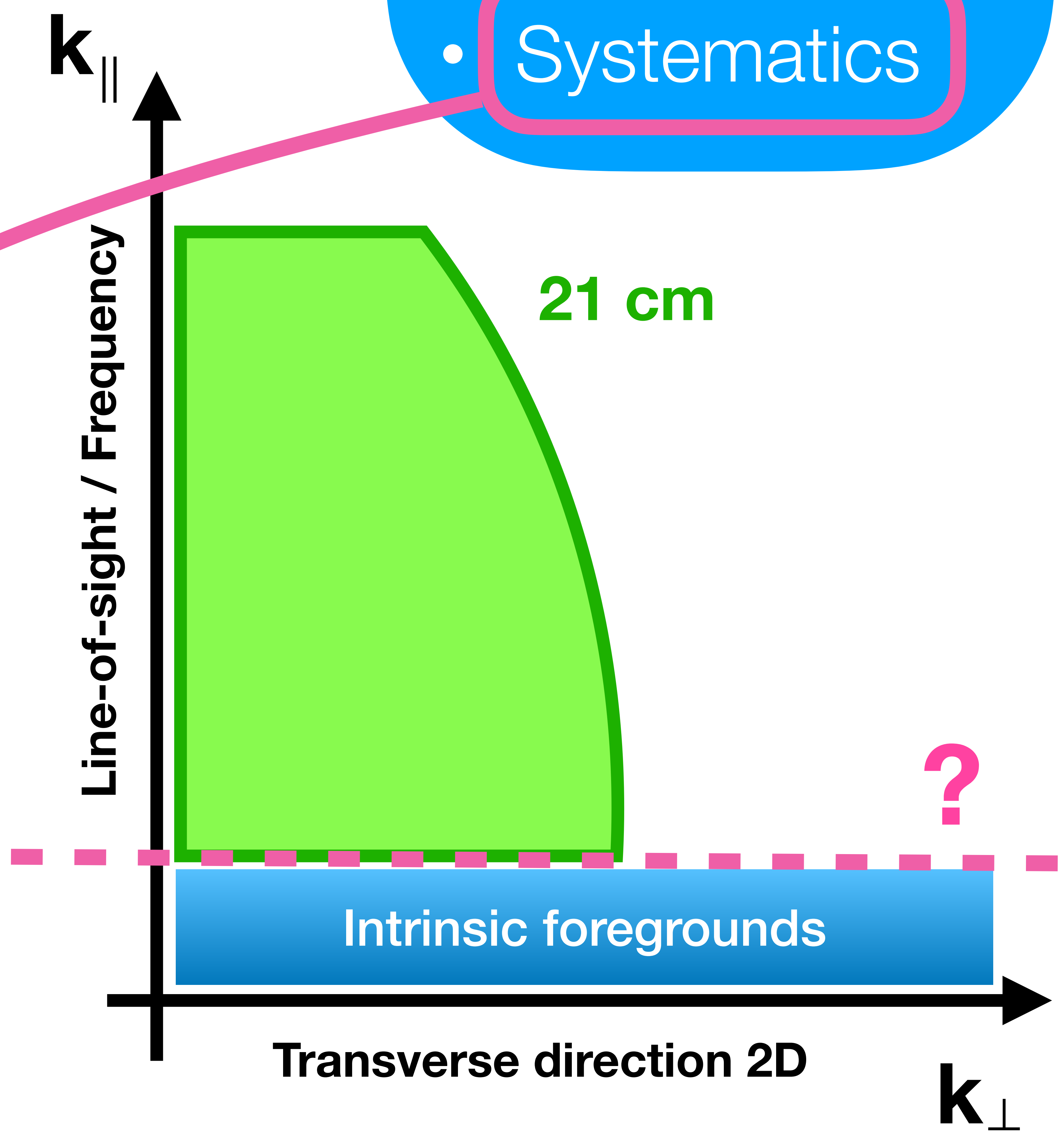
CHIME collaboration, 2022
stacking LRGs, ELG and QSOs from eBOSS
0.8<z<1.5

HI intensity mapping:
buried under the foregrounds

CHALLENGES:

- Foregrounds
- Systematics





HI intensity mapping:
buried under the foregrounds

1. **Ongoing** efforts:
   - statistical learning techniques (borrowed from the signal processing community) and
   - optimised statistical estimators, based on 2- and 3-point correlations.
2. Testing on simulations & **data** — MeerKLASS

Blind Source Separation algorithms

The separation of a set of source signals (contaminants) from a set of mixed signals (the maps), with little or no info about the source signal or the mixing process.

mixing matrix (f,n)

HI signal!

$$\mathbf{X} = \mathbf{A}\,\mathbf{S} + \mathbf{N}$$

signal (f,p)

sources (n,p)

- **Decorrelation —>** diagonalise the covariance matrix
- **Independence —>** as more independent sources are mixed the signal becomes more Gaussian (central limit theorem). So, let's maximise the non-gaussianity of the sources to *unmix* them.
- **Sparsity —>** mixtures are less sparse than sources!

Principal Component Analysis (**PCA**)

Independent Component Analysis (**ICA**)

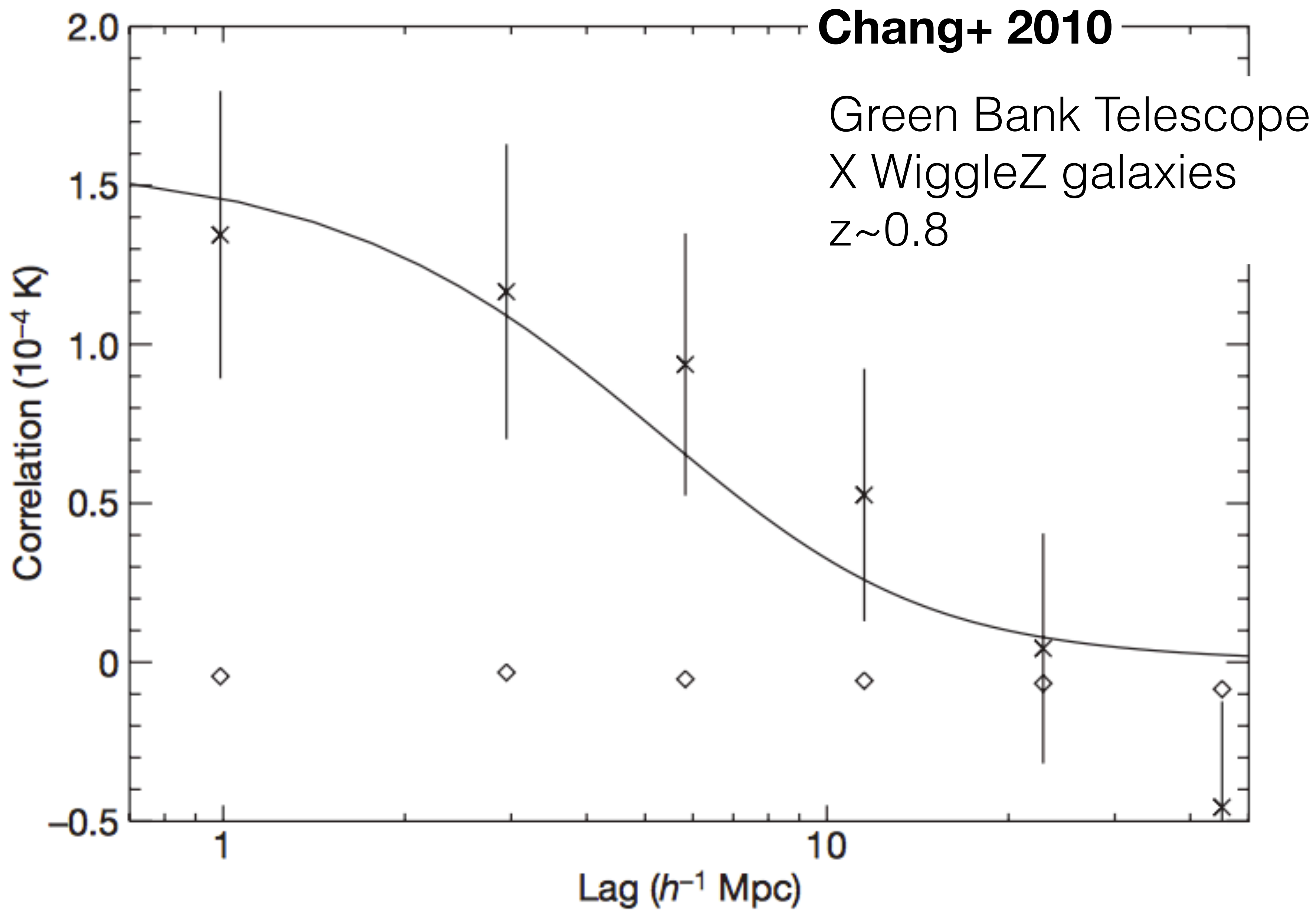
Hyvarinen + 1999

Generalised Morphological Component Analysis (**GMCA**)

Bobin + 2007

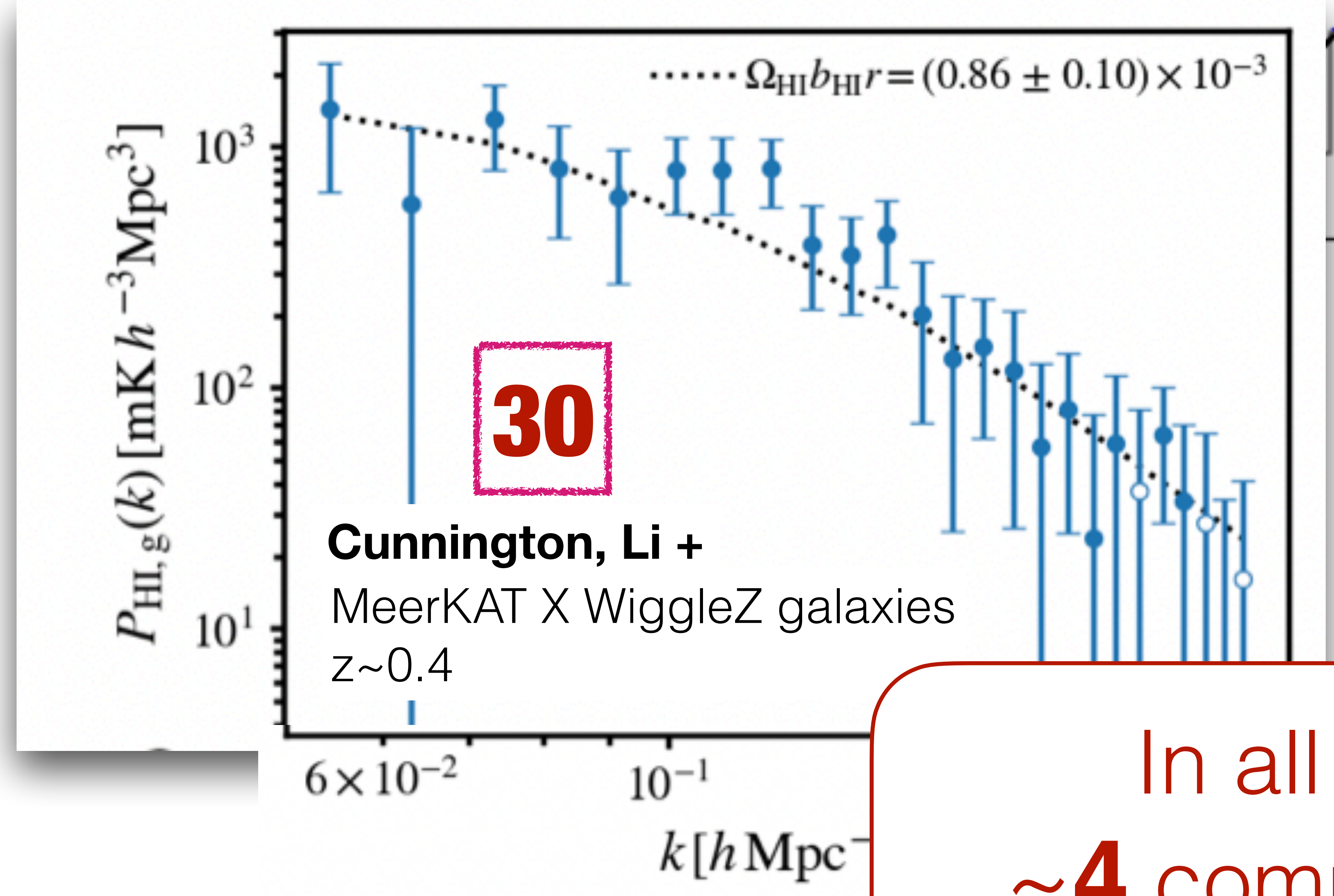
HI intensity mapping

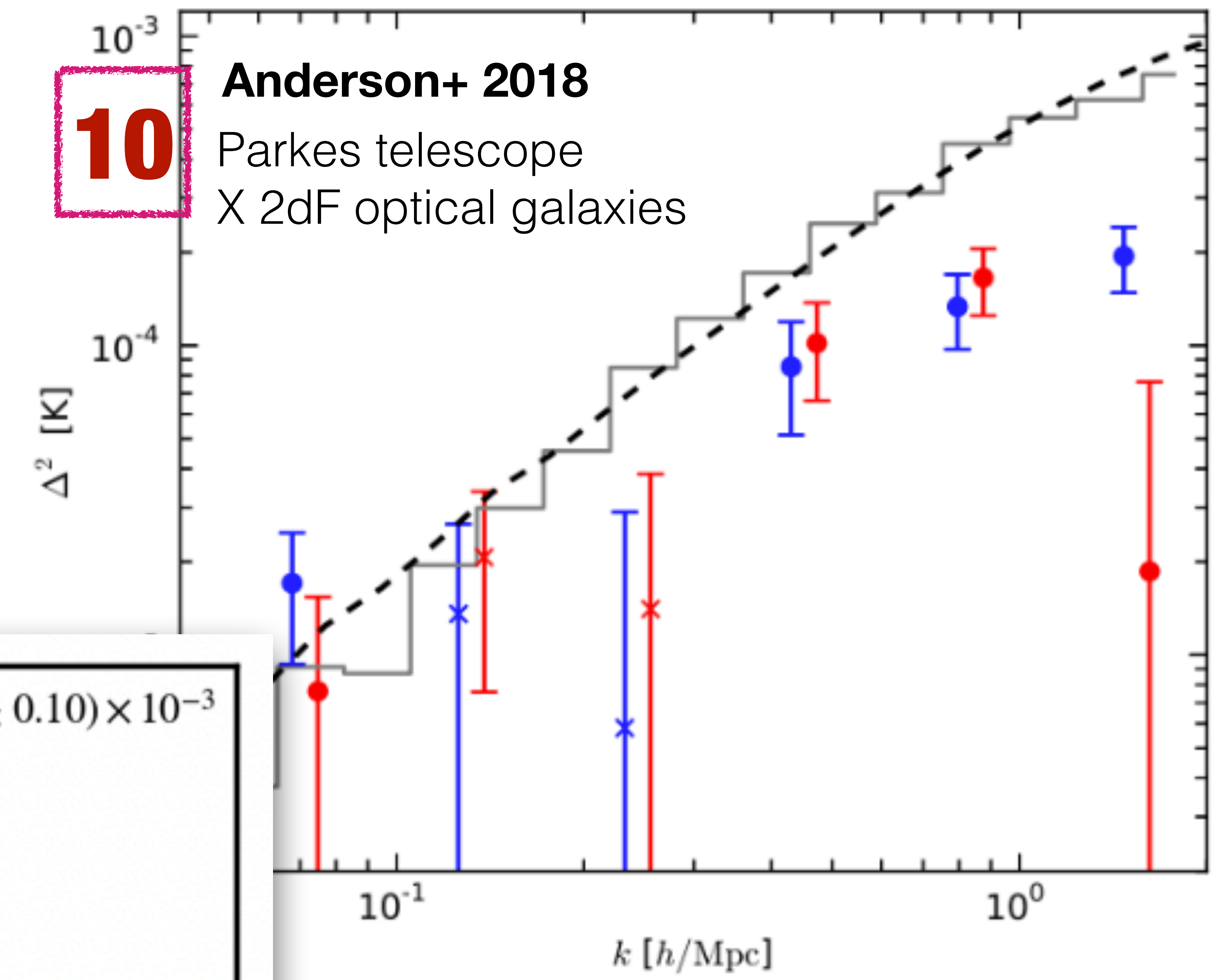
State-of-the-art

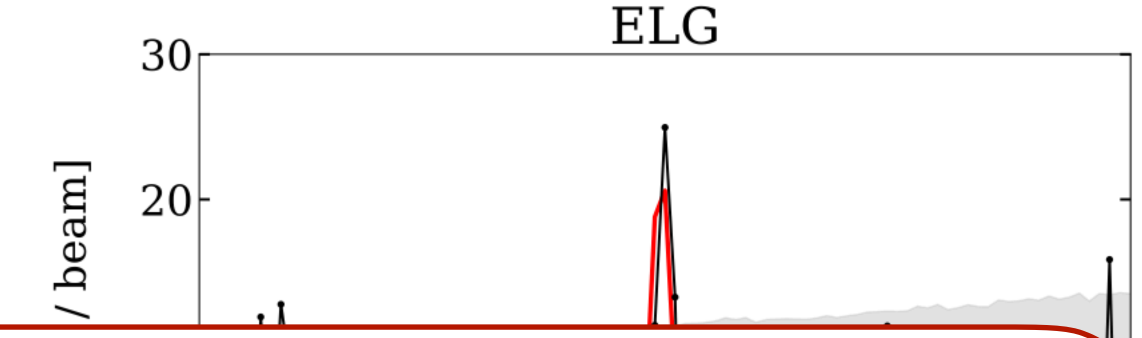


also Masui+ 2013, Switzer+ 2013, Wolz+ 2017,2022

10 - 20 **20 - 36**



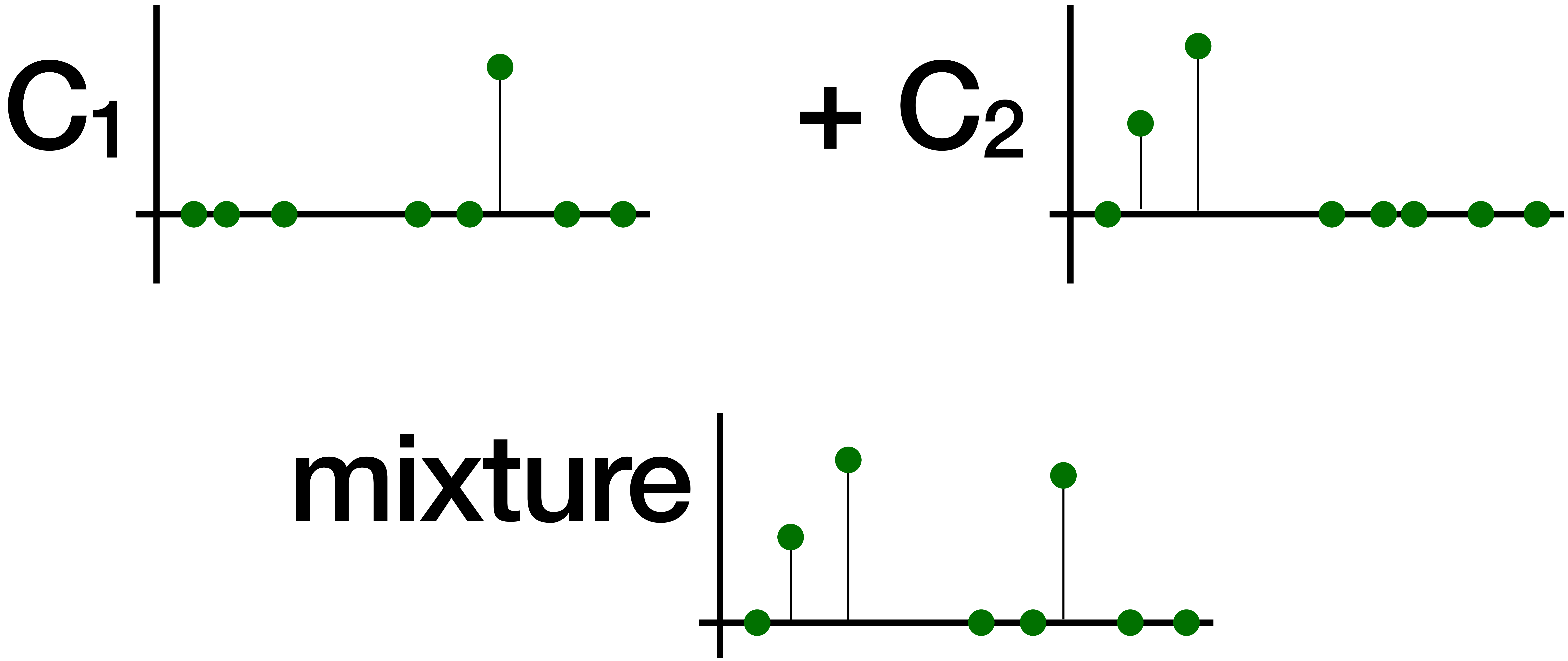




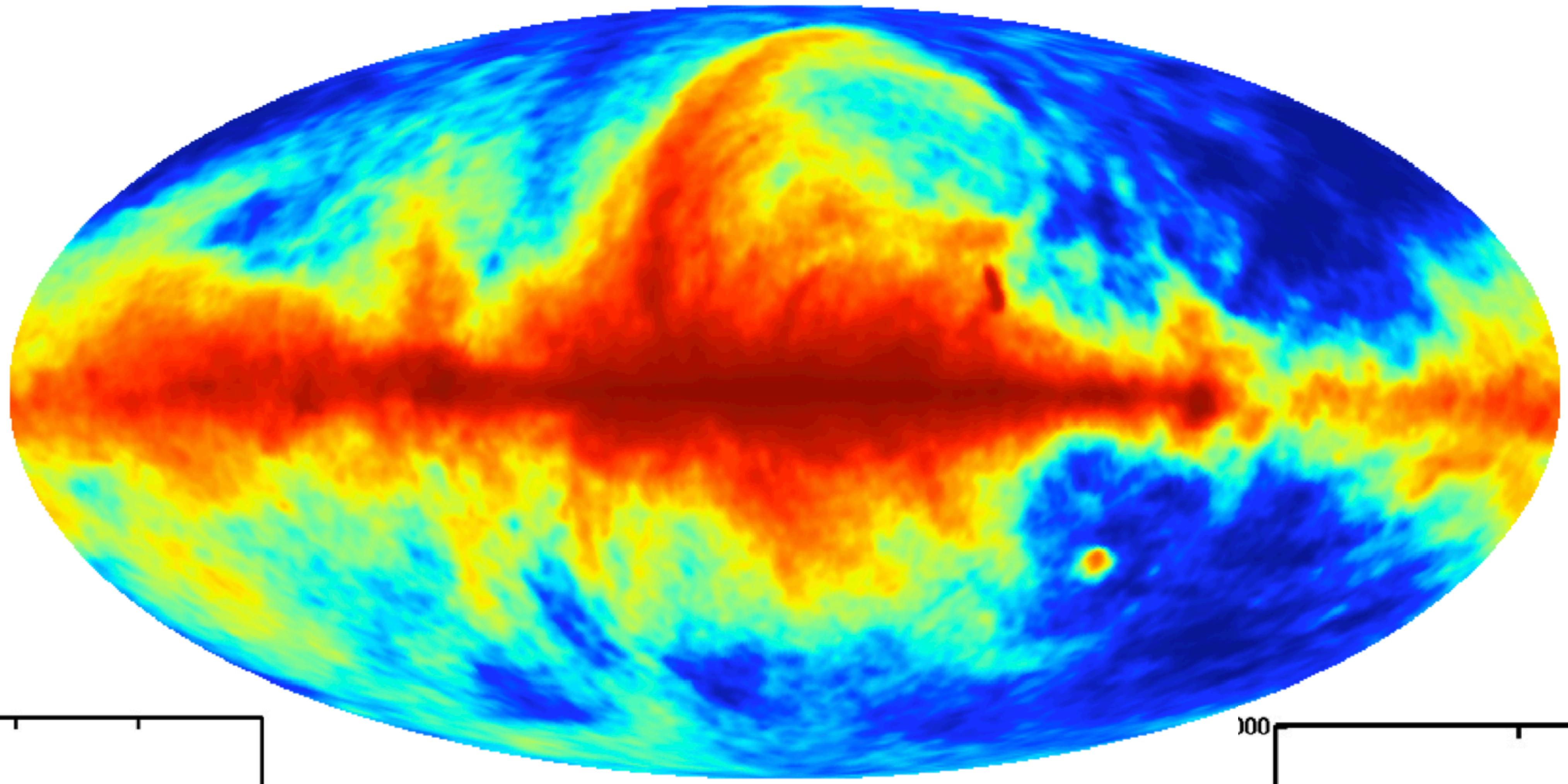
In all theoretical works:
~**4** components removed are enough

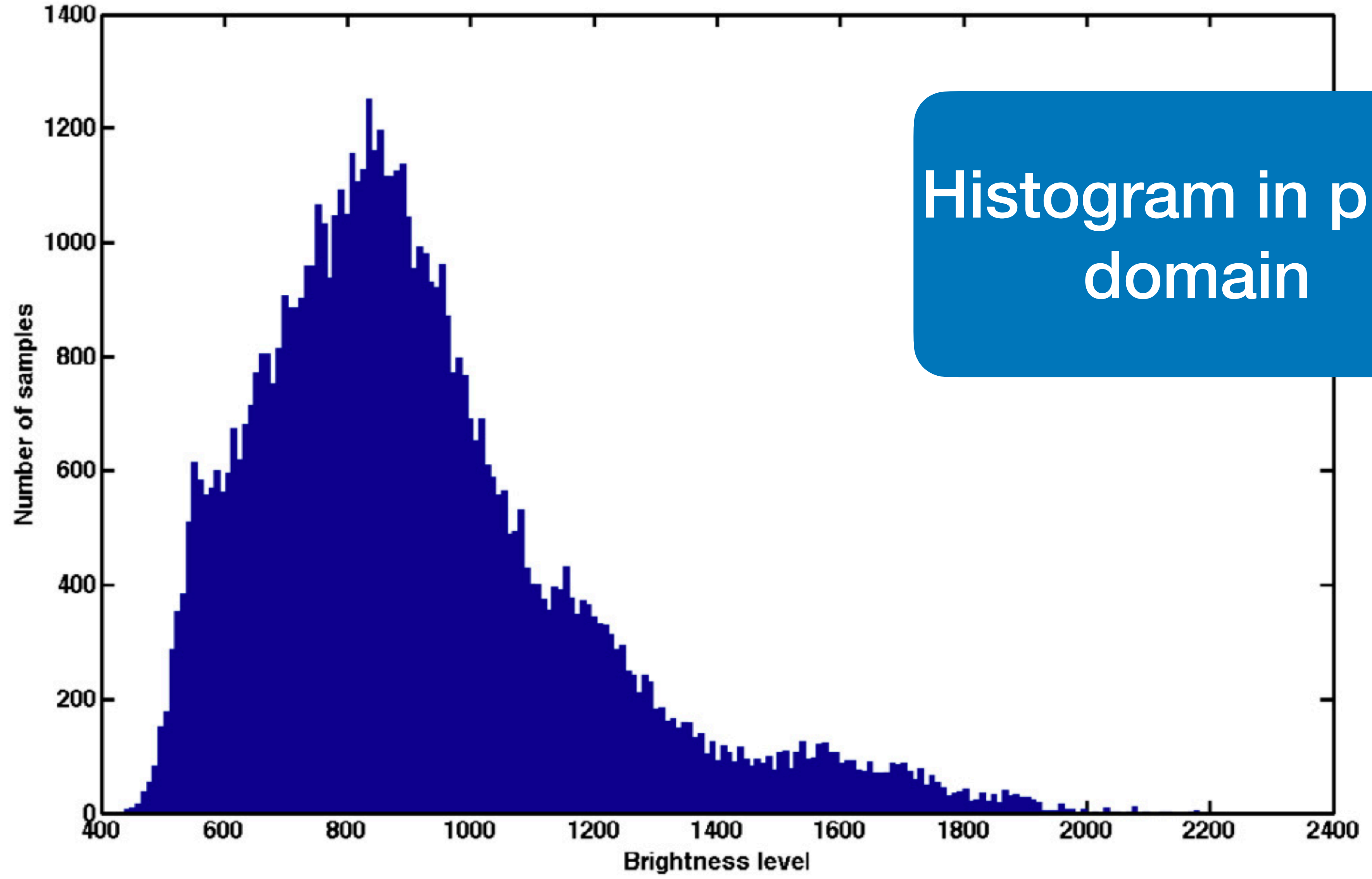
(e.g., Wolz+ 2014, Alonso+ 2015, Cunnington+ 2019) why sparsity?

mixtures are less sparse than components

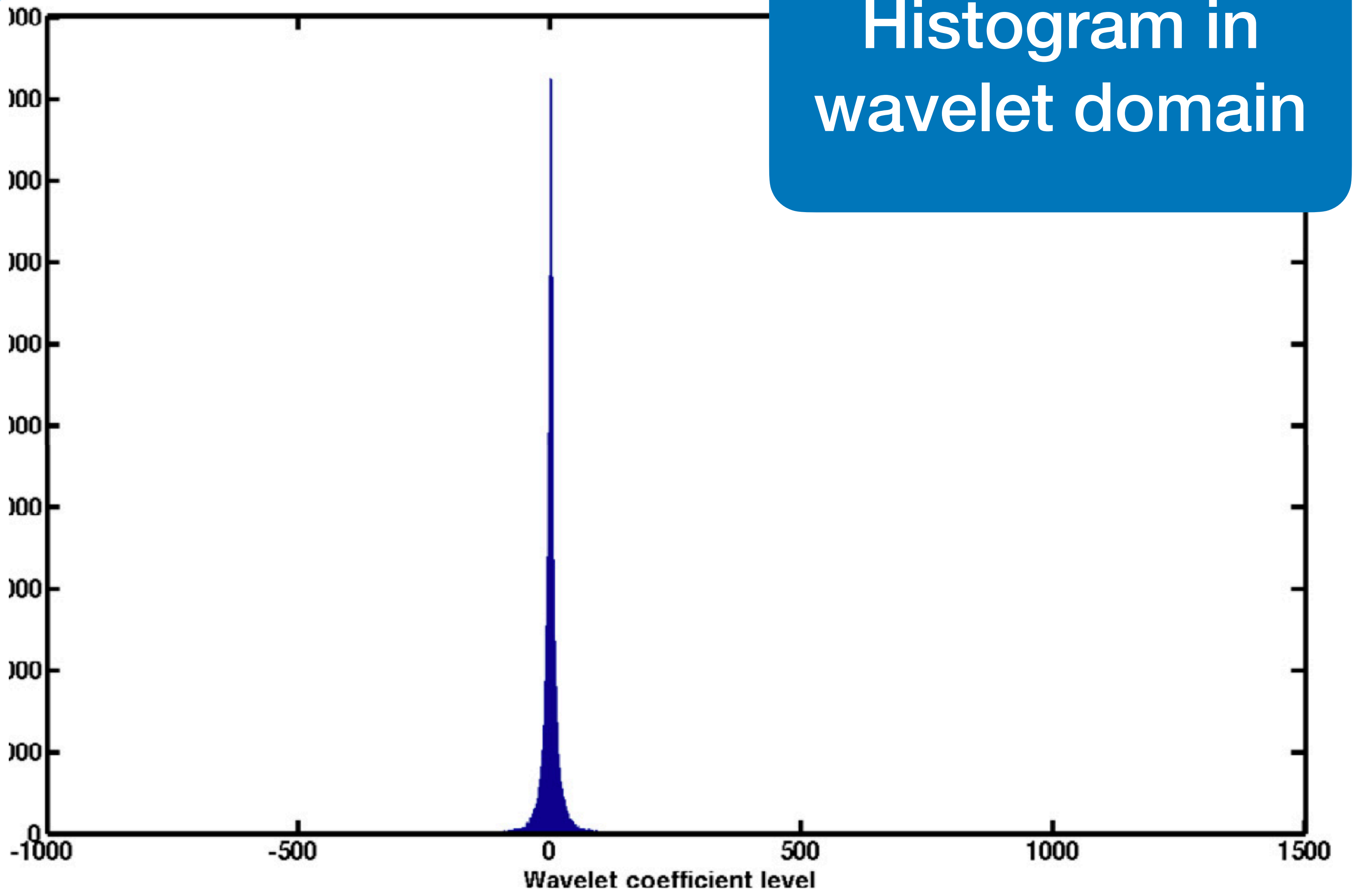


Enforcing sparsity: in which domain?



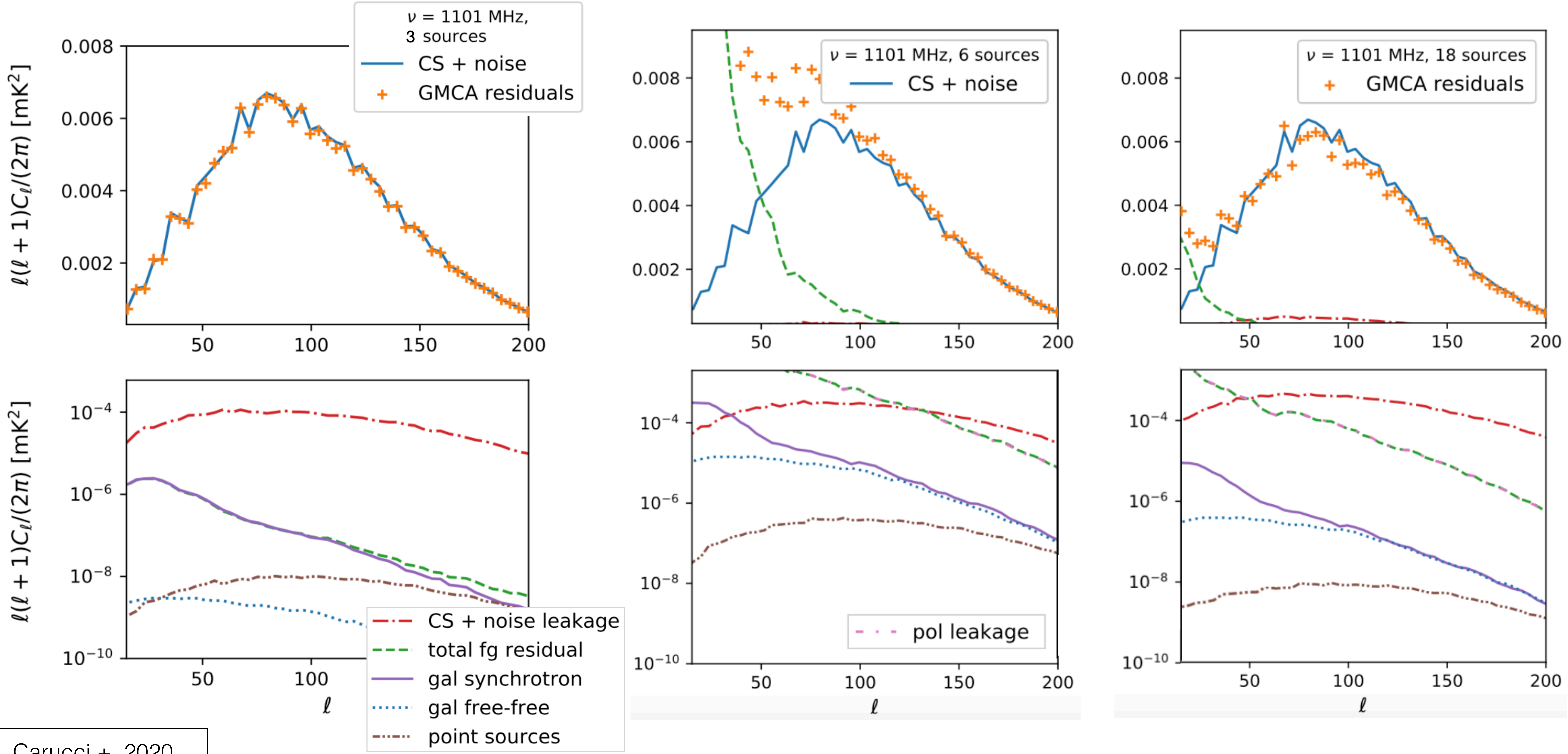


Histogram in pixel domain



Histogram in wavelet domain

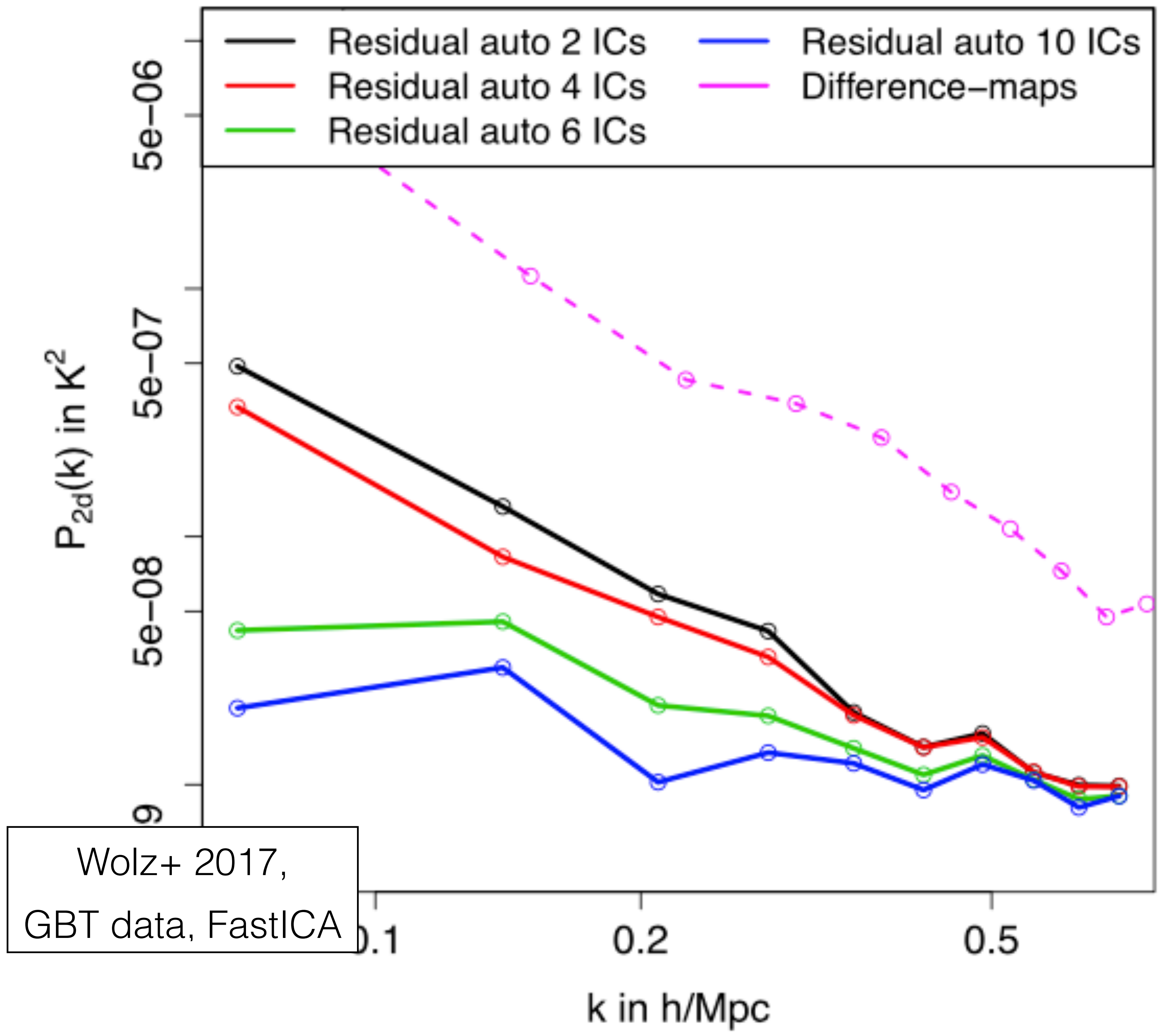
Different scales need different care



Carucci +, 2020

Different scales need different care

The wavelet domain is a multi-scale framework!



Wolz+ 2017,
GBT data, FastICA

- GMCA performs very well on small scales, can fail at the large scale
- PCA / ICA —> overfit the large scales

PCA on the large scale

+

GMCA on the small scales

mixGMCA

See also Hothi+2020 with LOFAR data

MeerKLASS: MeerKAT Large Area Synoptic Survey

ArXiv: 1709:06099

Alkistis Pourtsidou, Amadeus Wild, Brandon Engelbrecht, Isabella Carucci, Jingying Wang, Keith Grainge, Laura Wolz, Marta Spinelli, Mel Irfan, Mario Santos, Phil Bull, Stefano Camera, Steve Cunnington, Zé Fonseca, ...

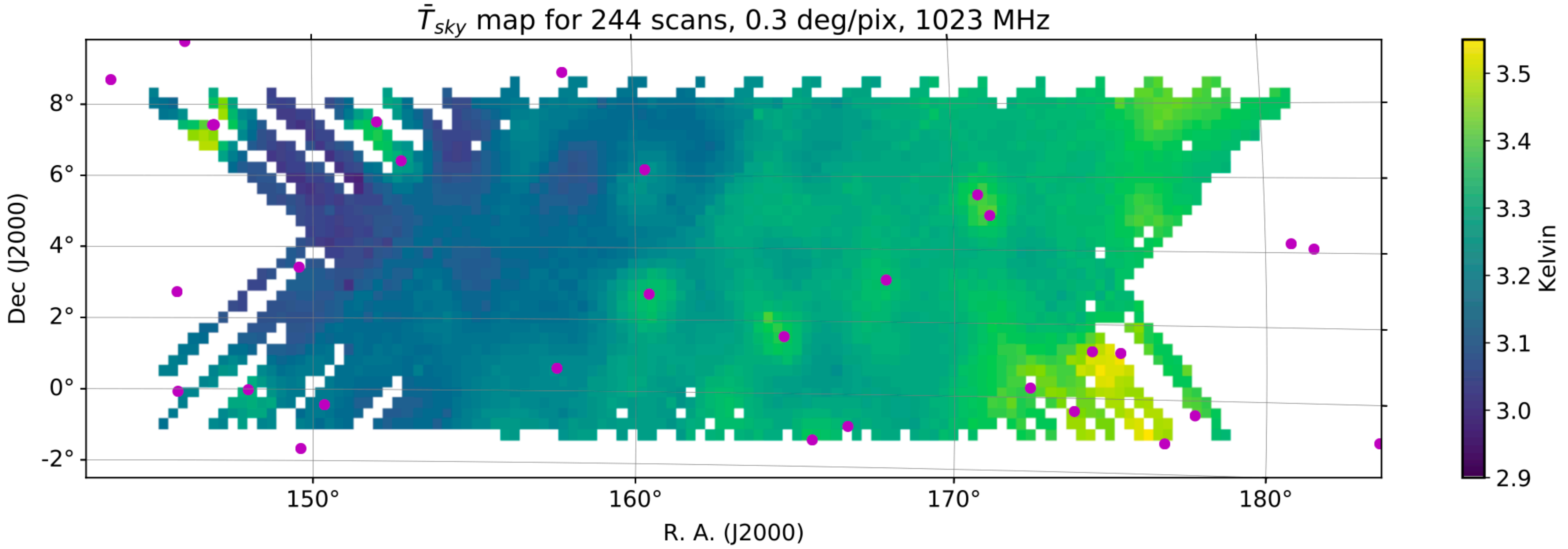
MeerKLASS: MeerKAT Large Area Synoptic Survey

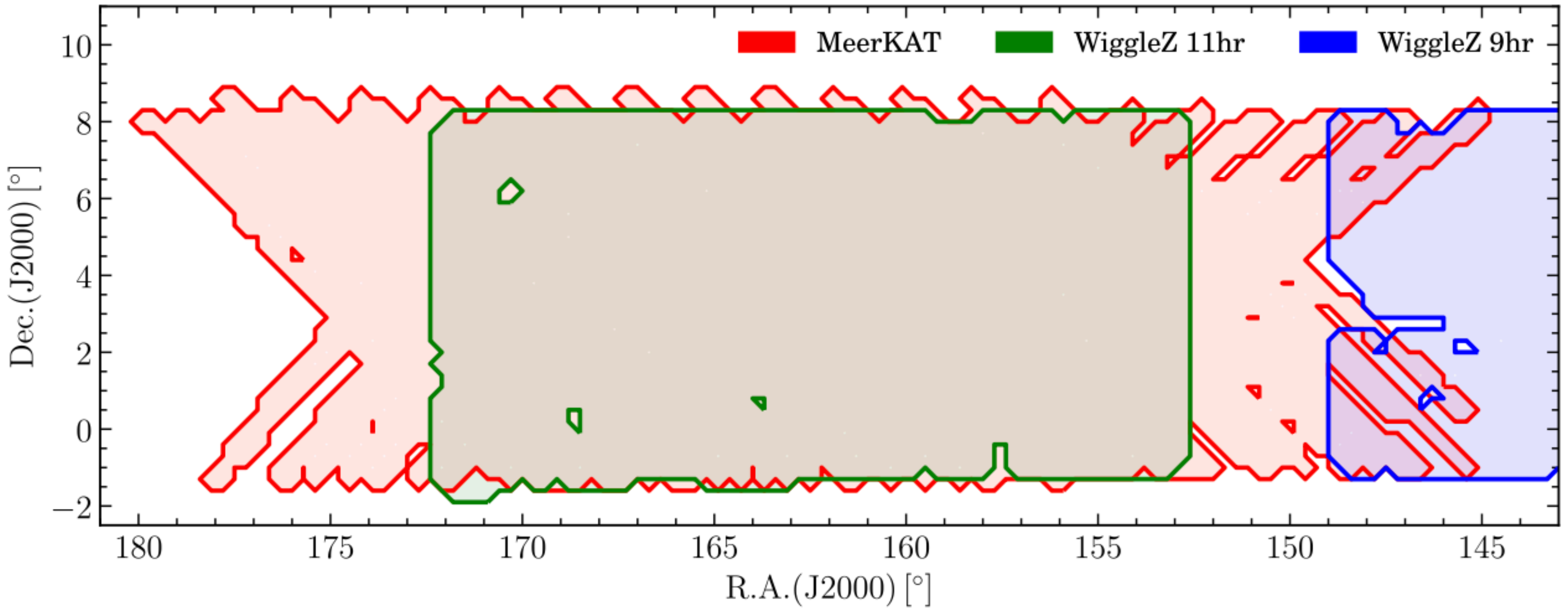
ArXiv: 1709:06099

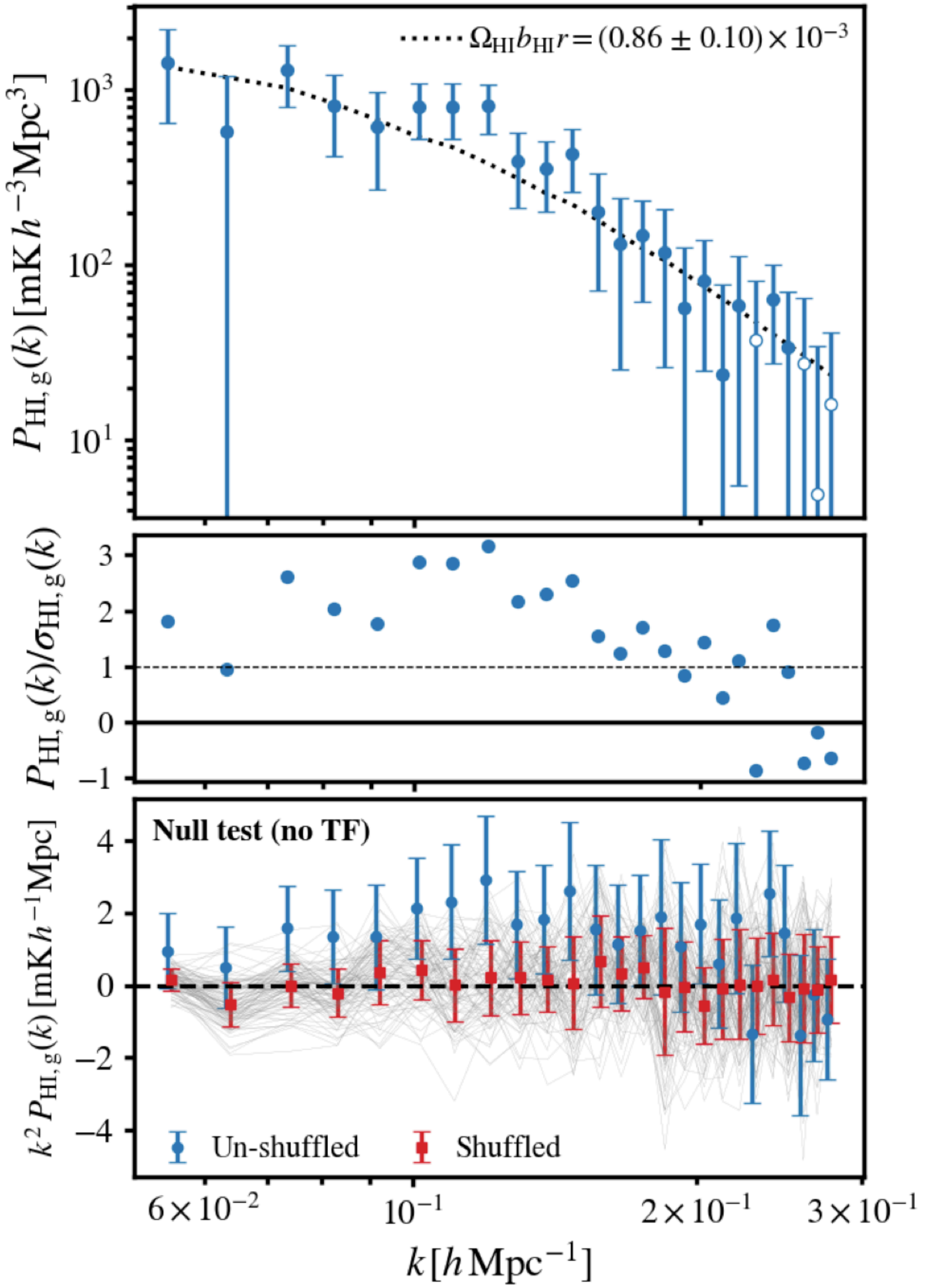
- Wang + 2021 — Calibration paper
- Irfan + 2022 — Synchrotron Spectral Index Measurement
- Cunnington, Li + — Detection in cross-correlation with WiggleZ galaxies

Pilot survey data (2019):

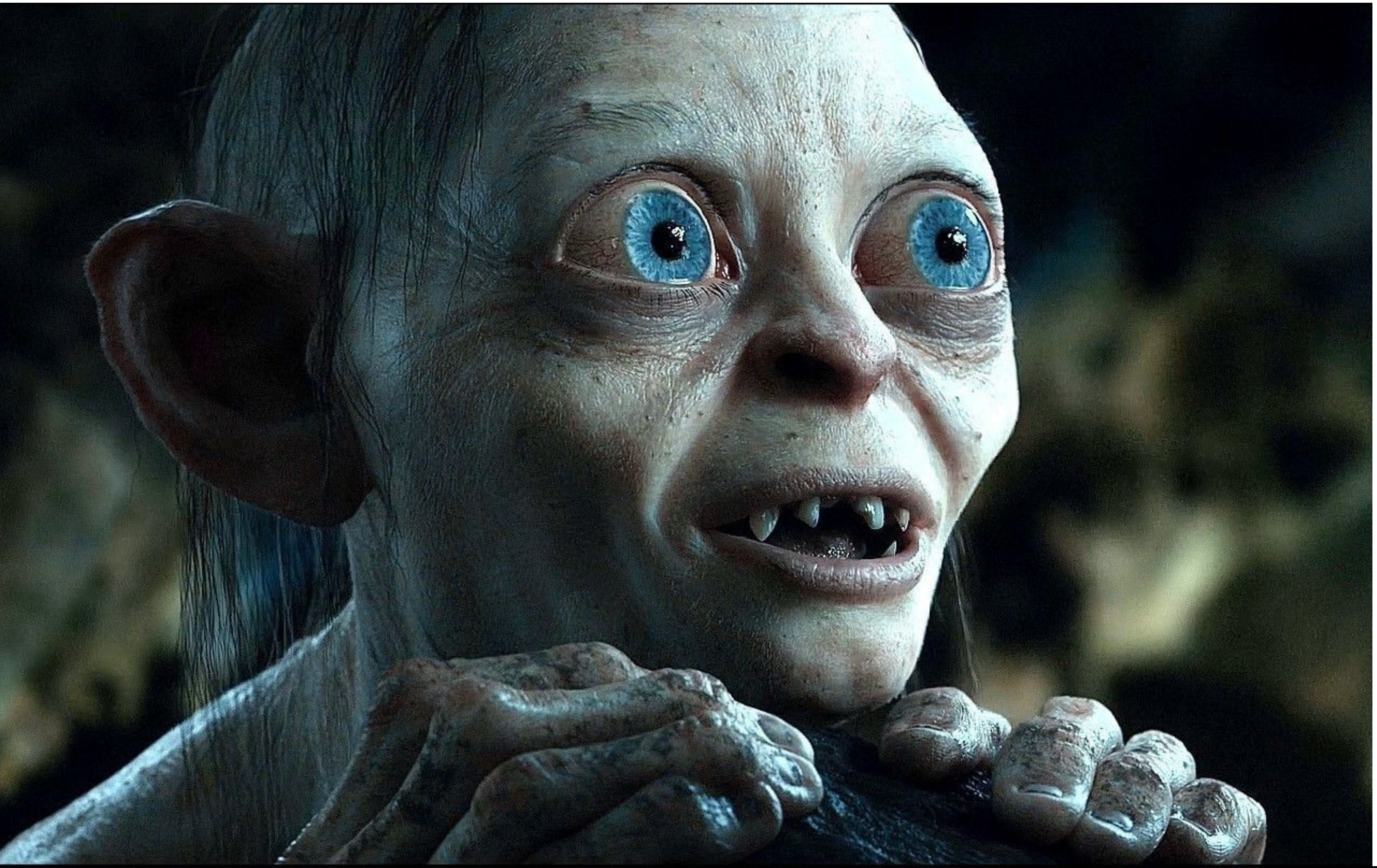
- 10.5 hours of data from six nights of observations
- Overlapping with the WiggleZ11hr field (~200 deg2)
- We use data in range 973-1015 MHz (0.40 < z < 0.46)

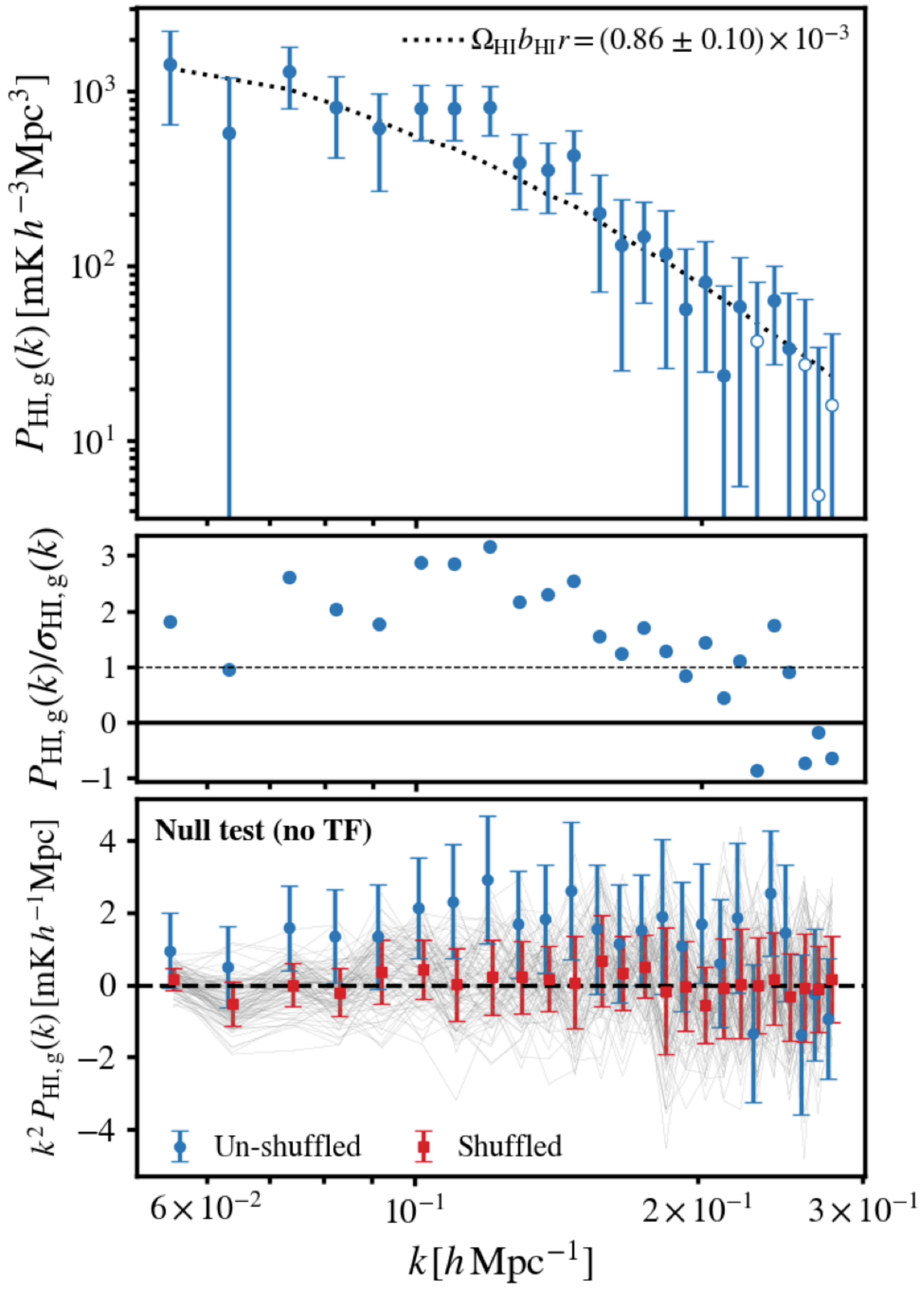






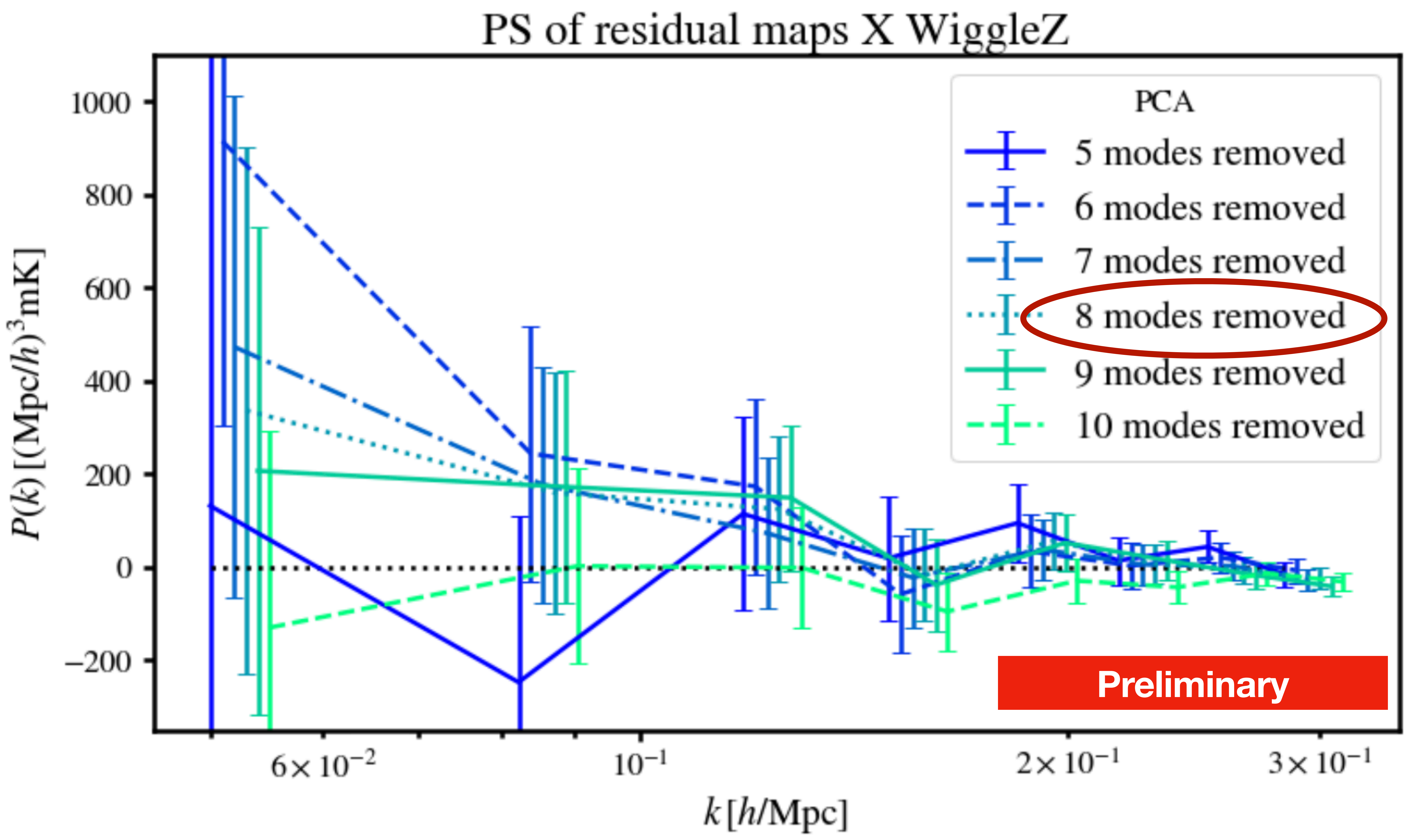
Our precious!!





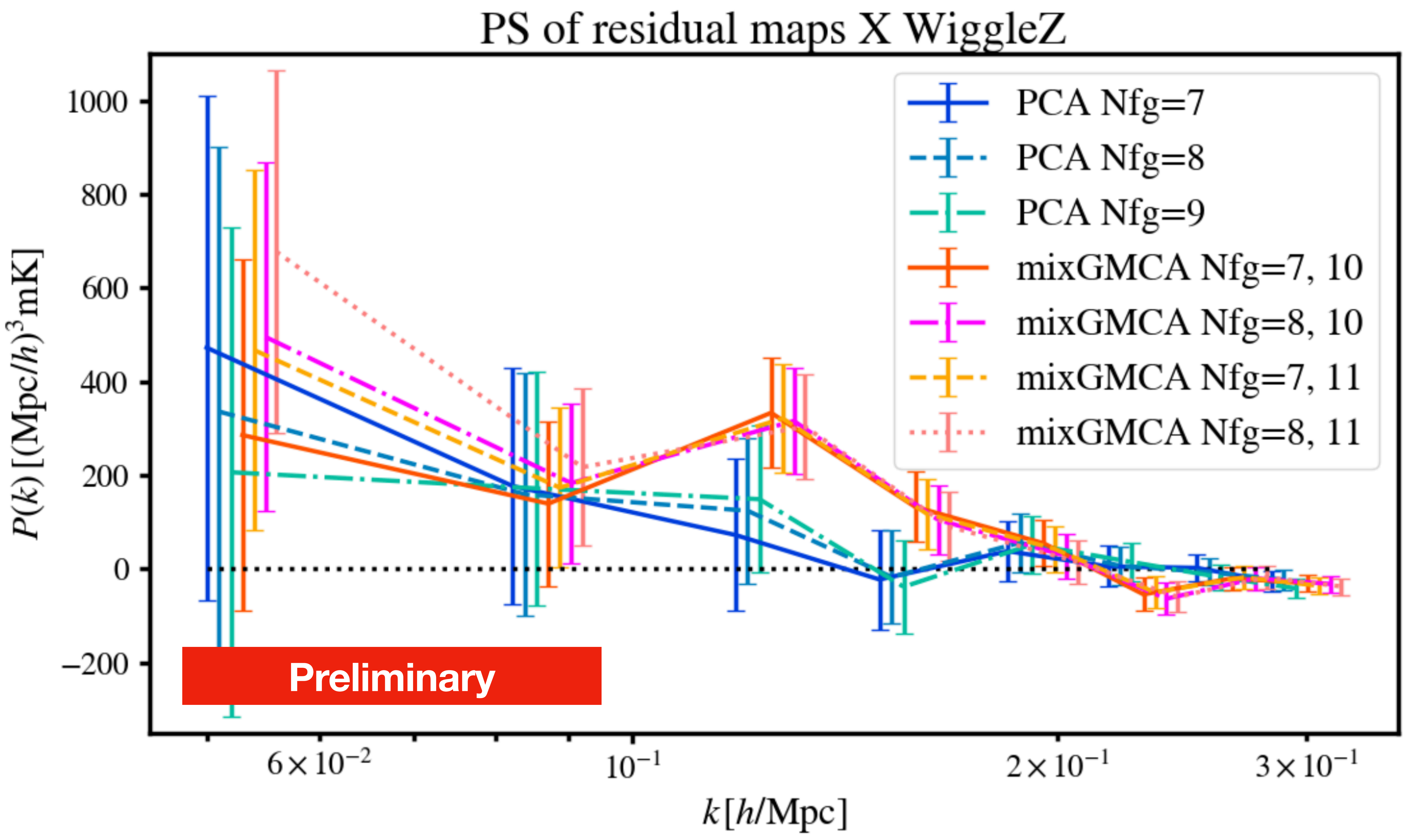
How many modes to subtract with e.g. PCA ?

Seems there's a clear sweet-spot with the cross- measurement

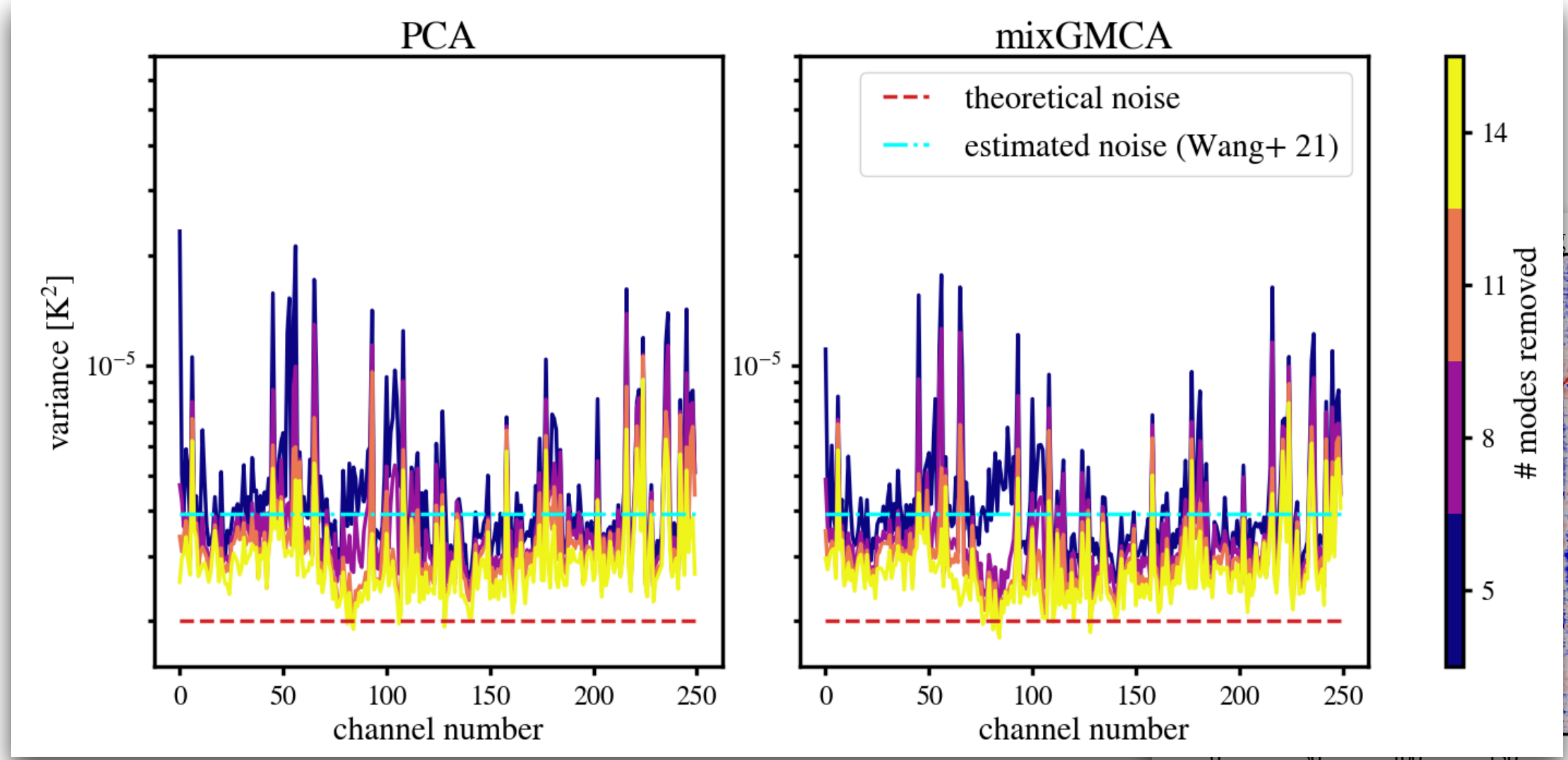


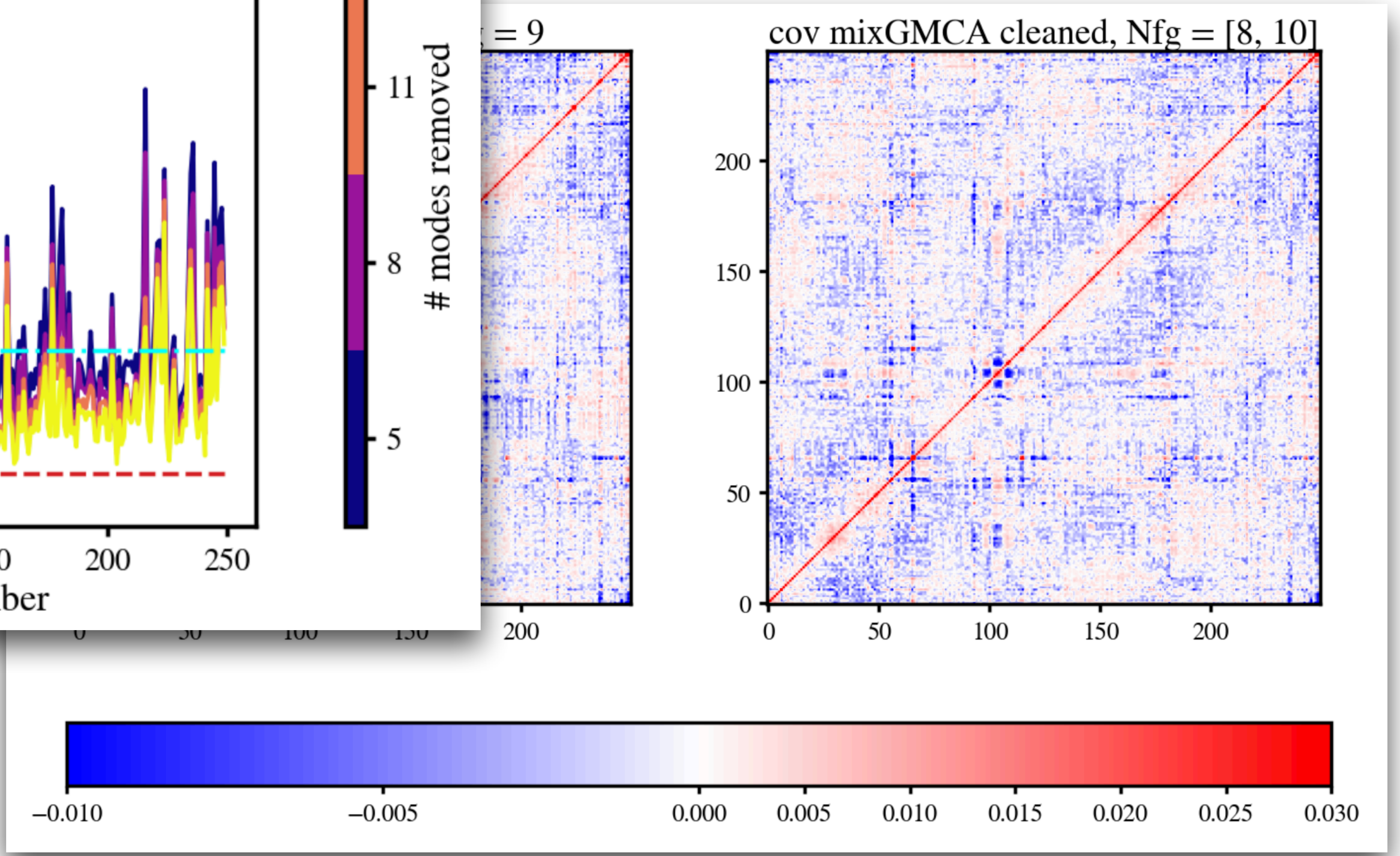
To be compared with 30 !

Comparing the different methods

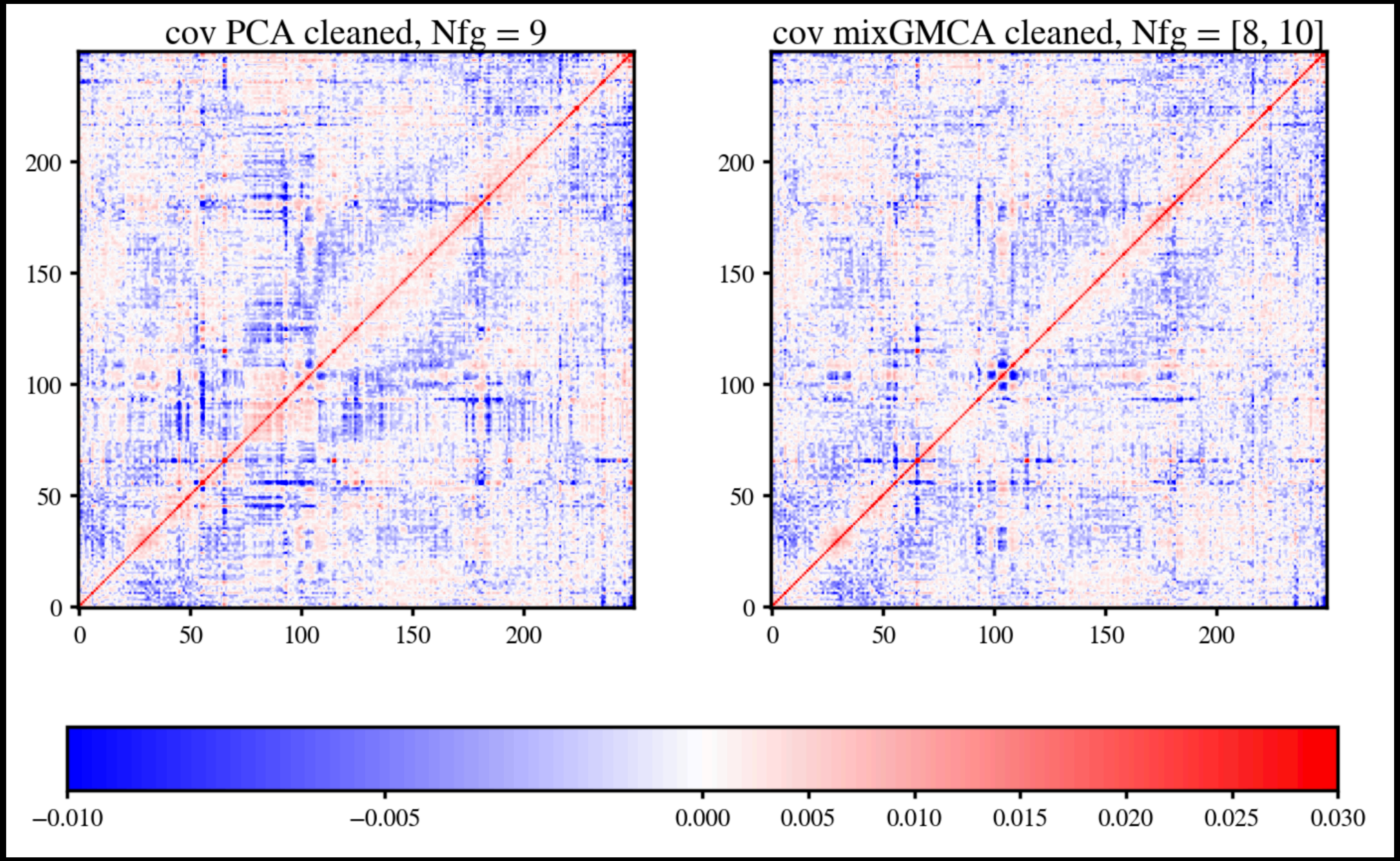


Could we tell about the N_{fg} to remove WITHOUT the cross info?



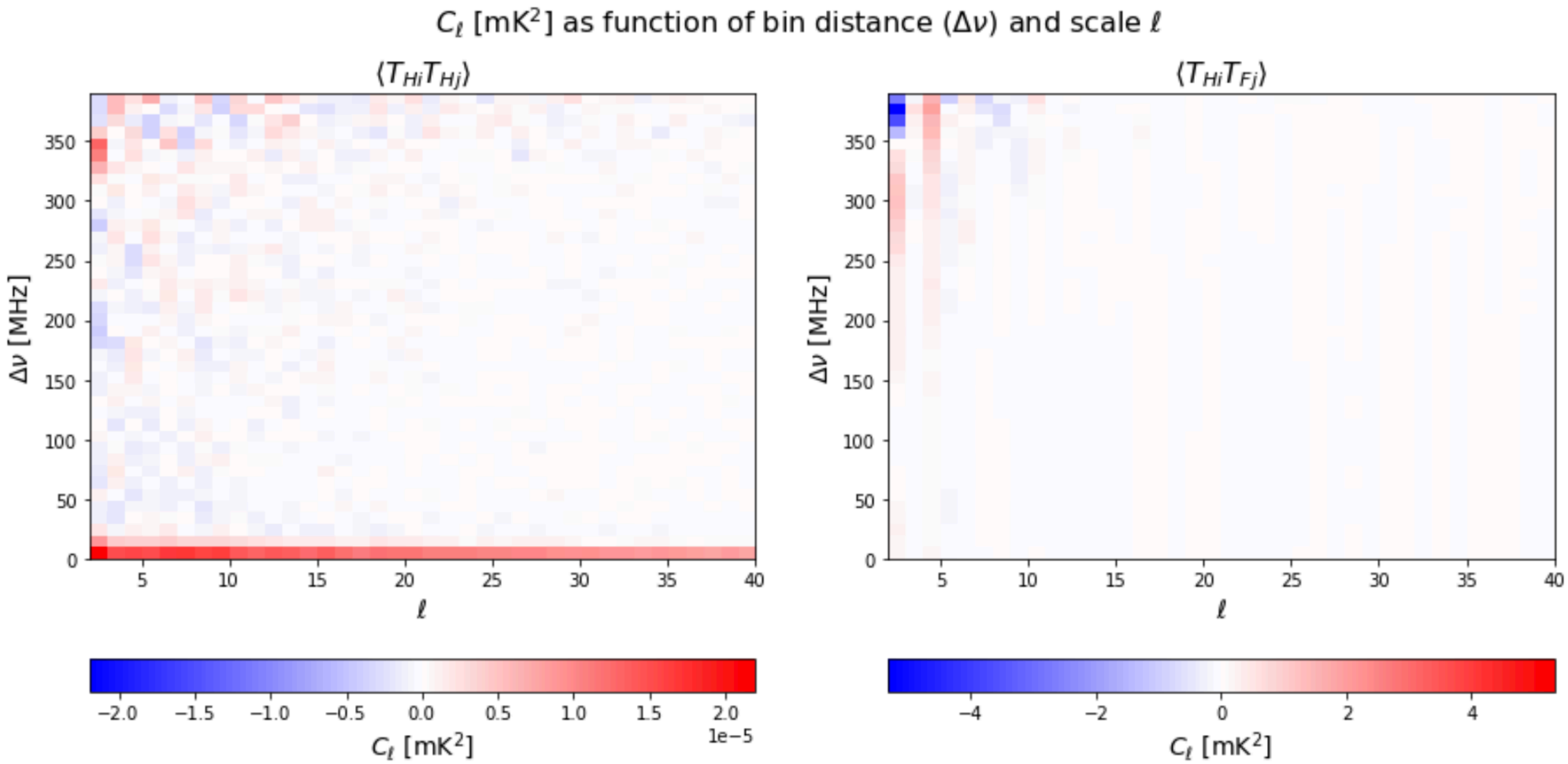


The frequency-frequency covariance of the HI intensity mapping signal is diagonal



$$T_o(\theta, z) = T_H(\theta, z) + T_F(\theta, z)\,.$$

$$\langle T_{o1} T_{o2} \rangle = \cancelto{\emptyset}{\langle T_{H1} T_{H2} \rangle} + \langle T_{F1} T_{F2} \rangle + \cancelto{\emptyset}{\langle T_{H1} T_{F2} \rangle} + \cancelto{\emptyset}{\langle T_{F1} T_{H2} \rangle}\,.$$



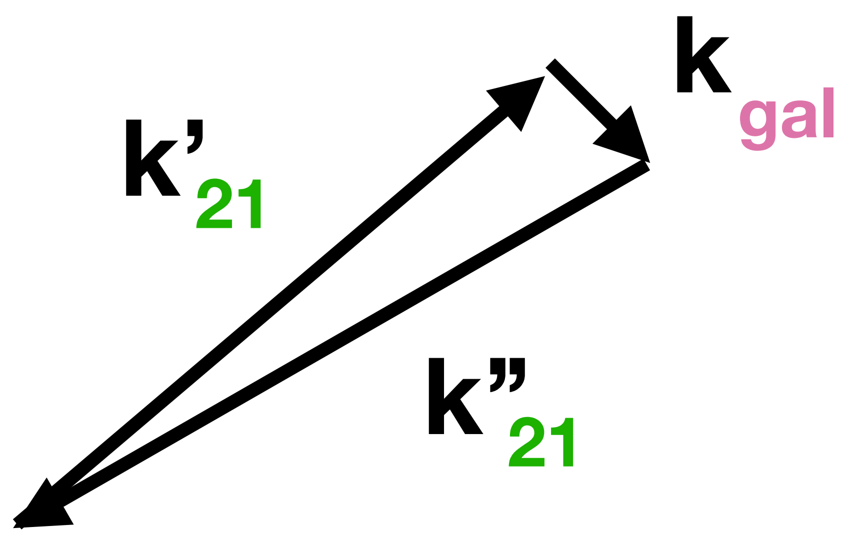
$$\langle T_{o1} T_{o2} \rangle = \langle T_{H1} T_{H2} \rangle + \langle T_{F1} T_{F2} \rangle + \langle T_{H1} T_{F2} \rangle + \langle T_{F1} T_{H2} \rangle .$$

∅ ∅ ∅

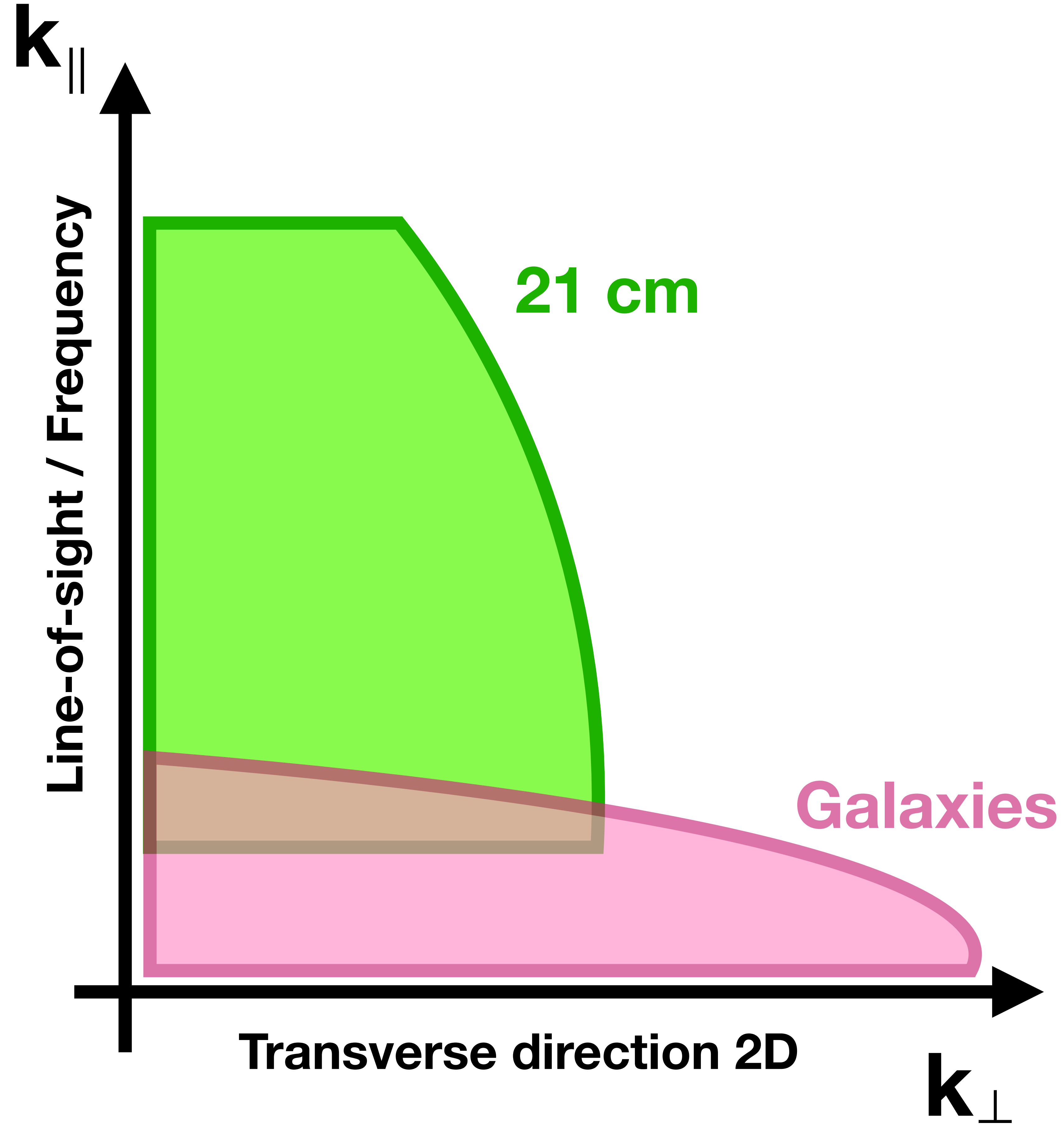
Beat systematics: **cross correlations**

Direct 21cm x galaxies signal vanishes due to foregrounds in long wavelength line-of-sight modes. Need to use higher order correlations.

- e.g., a *squeezed* bispectrum estimator:
  1 low-k mode from galaxy survey X 2 high-k 21 cm modes.

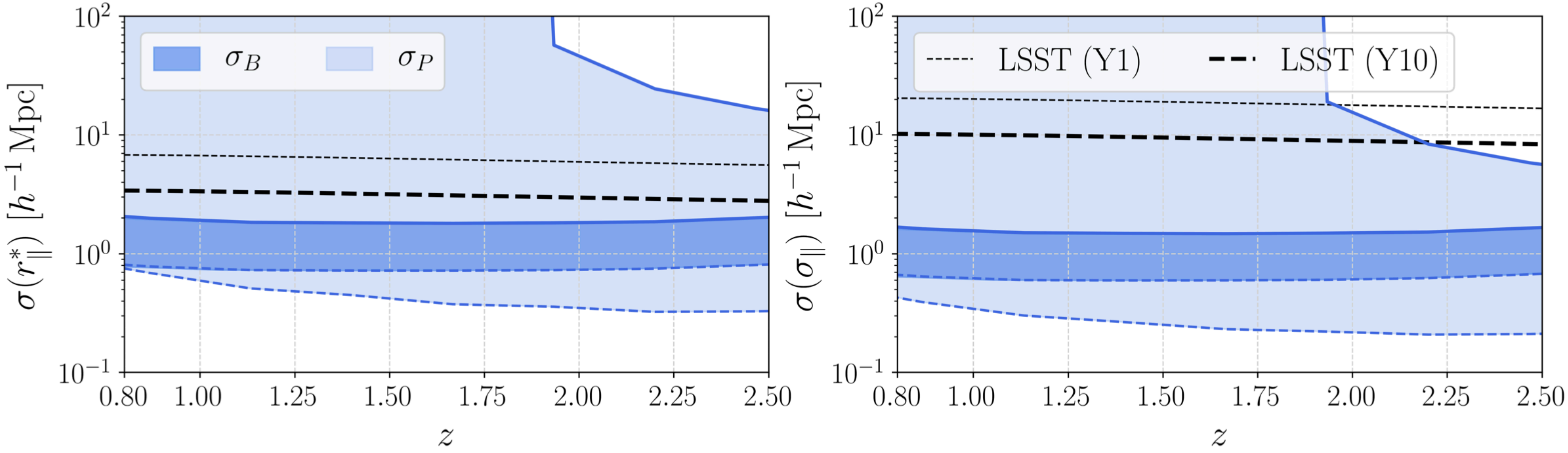


$$<\delta(\mathbf{k})\ \delta(\mathbf{k'})\ \delta(\mathbf{k''})>= \delta_D(\mathbf{k}+\mathbf{k'}+\mathbf{k''})\ B(\mathbf{k}+\mathbf{k'})$$



Beat systematics: **cross correlations**

Direct 21cm x galaxies signal vanishes due to foregrounds in long wavelength line-of-sight modes. Need to use higher order correlations.



Guandalin+ 2022

Summary

- **HI IM** will be game changer in cosmology
- Contaminants-removal is the biggest problem, lots of efforts devoted to this (also collectively within the SKA Cosmology SWG - IM Focus Group)
- Smart estimators + cleaver cross possibilities: we can get plenty of new/complementary info out of the maps
- **MeerKLASS** ongoing!

Enabling hydrogen intensity mapping

Proposed SKA1 Cosmology Surveys

a) Medium–Deep Survey of 5,000 deg² at 0.95-1.4 GHz for
- HI galaxy redshift survey with 3.5 million objects
- Weak Lensing shape measurements with ~50 million objects
- Continuum galaxy survey with ~60 million objects

b) Wide Survey of 20,000 deg² at 0.35-1.05 GHz for
- Continuum galaxy survey with ~100 million objects
- HI intensity maps for $0.35<z<3$

c) Deep Survey 100 deg² at 200-350 MHz for
- HI intensity maps for $3<z<6$

Cosmology with Phase 1 of the Square Kilometre Array **Red Book** 2018: Technical specifications and performance forecasts