

Studying the large-scale structure of the Universe with hydrogen intensity mapping

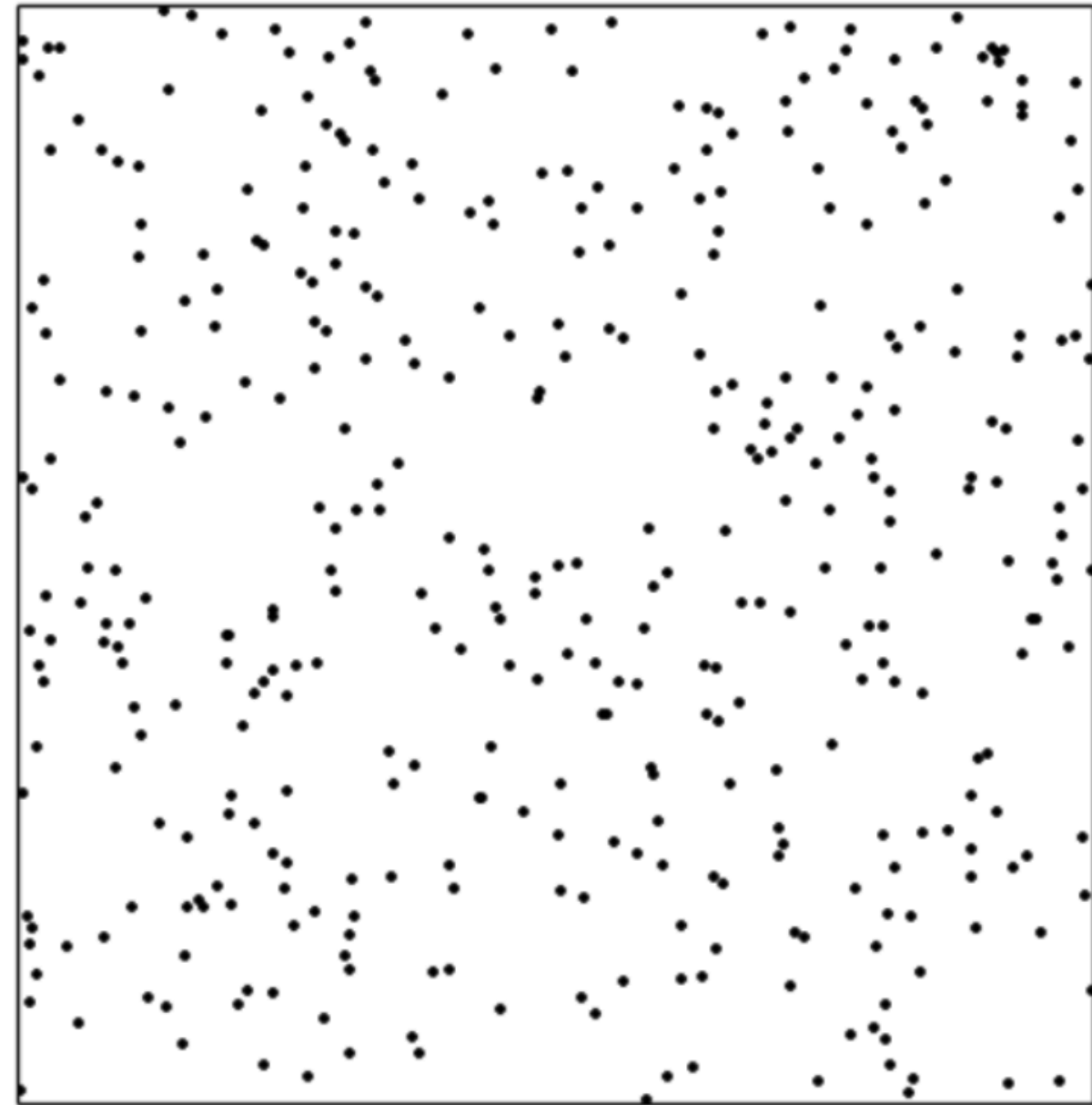
Isabella Paola Carucci (INAF - Trieste)



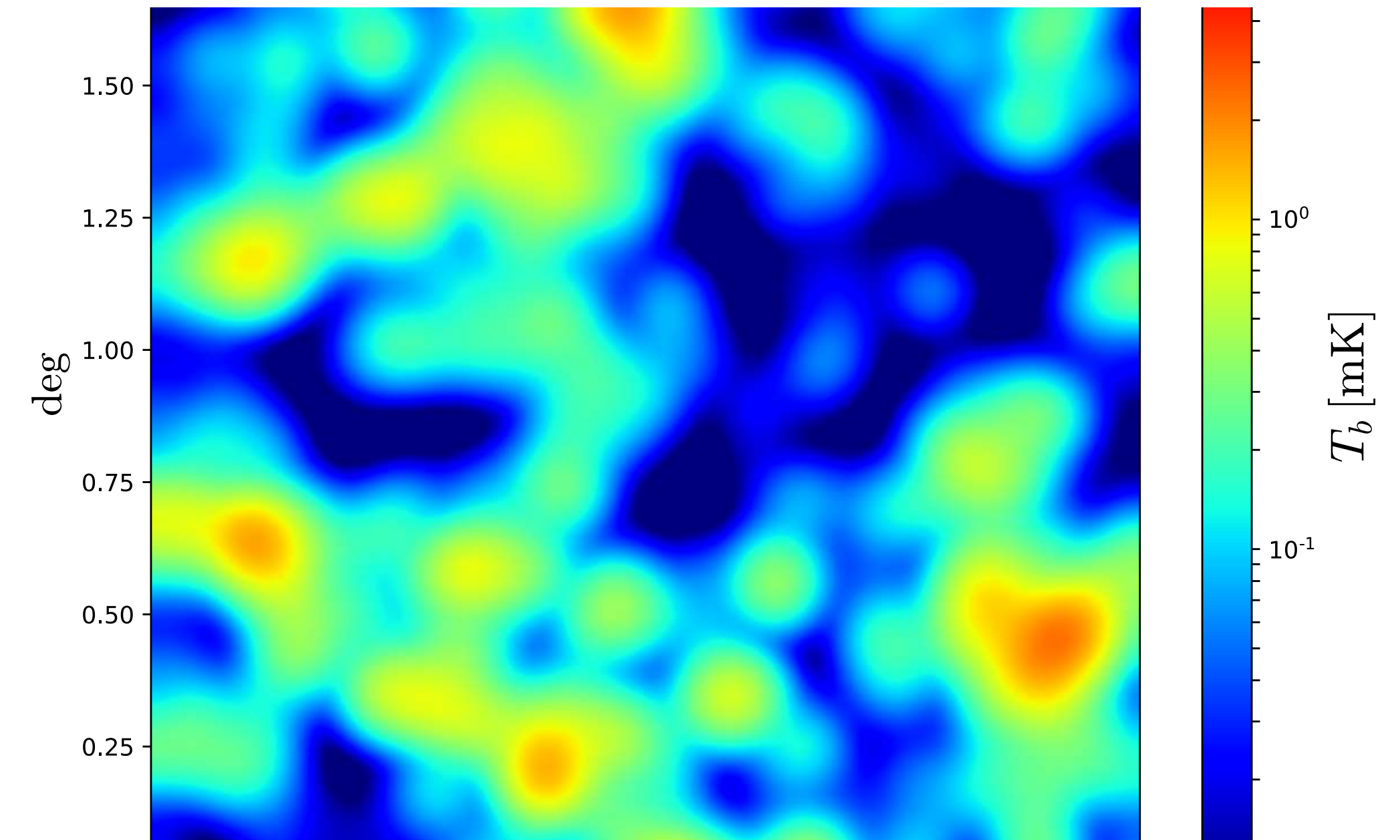
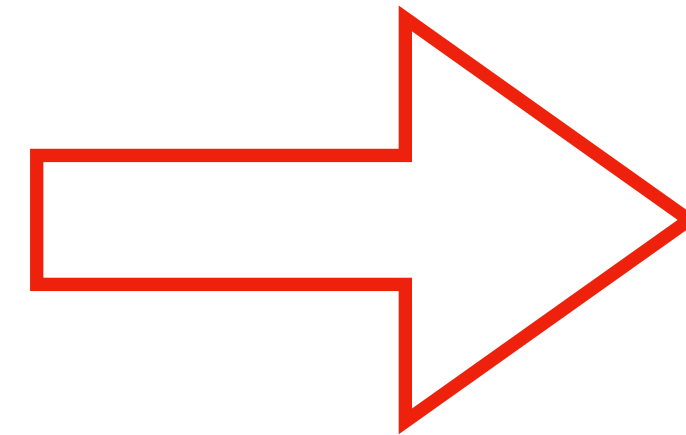
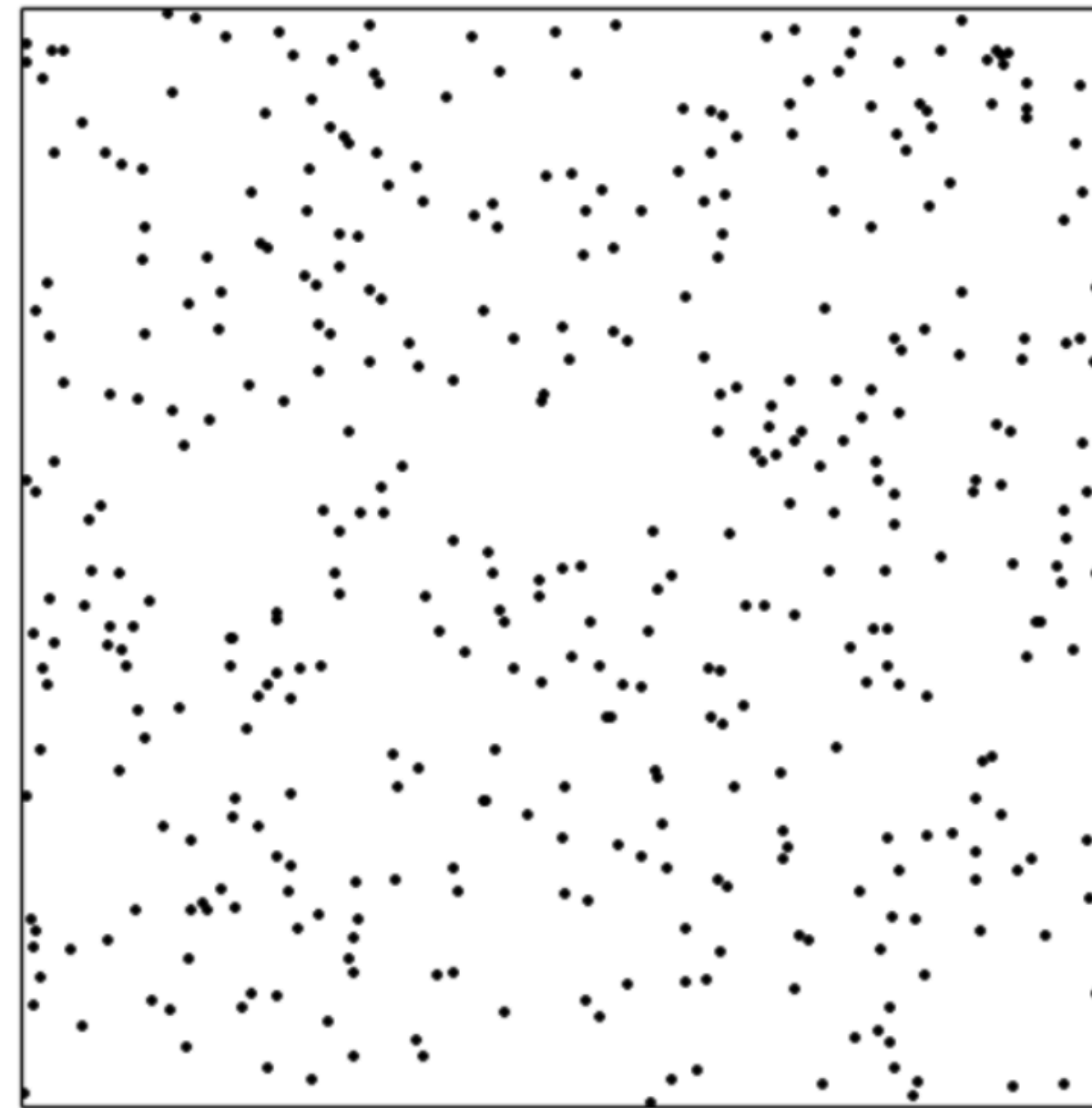
Funded by
the European Union
NextGenerationEU



Understanding the Galaxy/Matter Connection in the Era of Large Surveys



Understanding the Galaxy/Matter Connection in the Era of Large Surveys



1. **Hydrogen Intensity Mapping (IM):** what is it and why to do it

1. **Hydrogen Intensity Mapping (IM):**
what is it and why to do it

2. **IM is hard...**
(Biggest challenge: weakness of the IM
signal compared to contaminants)

1. **Hydrogen Intensity Mapping (IM):**

what is it and why to do it

2. **IM is hard...**

(Biggest challenge: weakness of the IM signal compared to contaminants)

3. **...yet we are getting there!**

(MeerKAT IM observations, analysis ongoing)

1. **Hydrogen Intensity Mapping (IM):**

what is it and why to do it

2. **IM is hard...**

(Biggest challenge: weakness of the IM signal compared to contaminants)

3. **...yet we are getting there!**

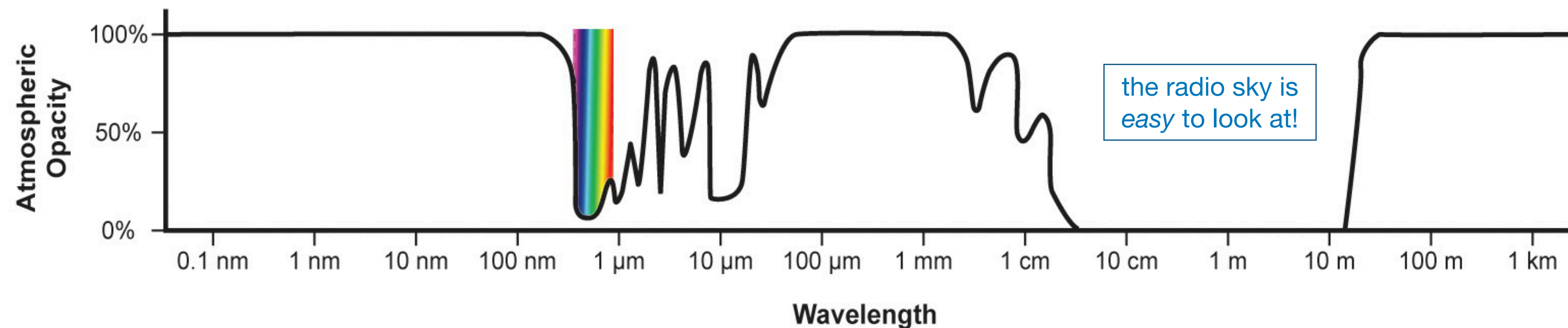
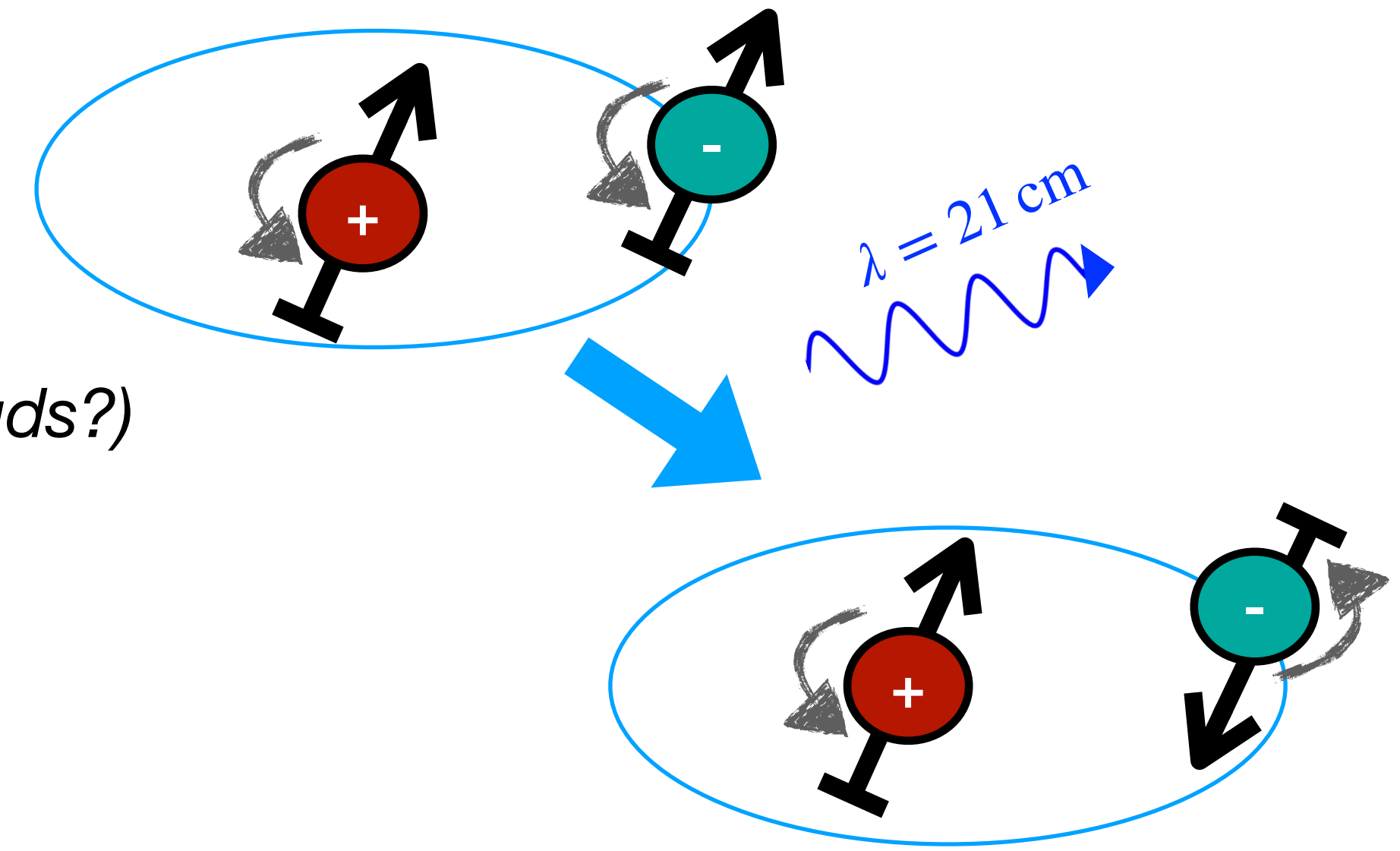
(MeerKAT IM observations, analysis ongoing)

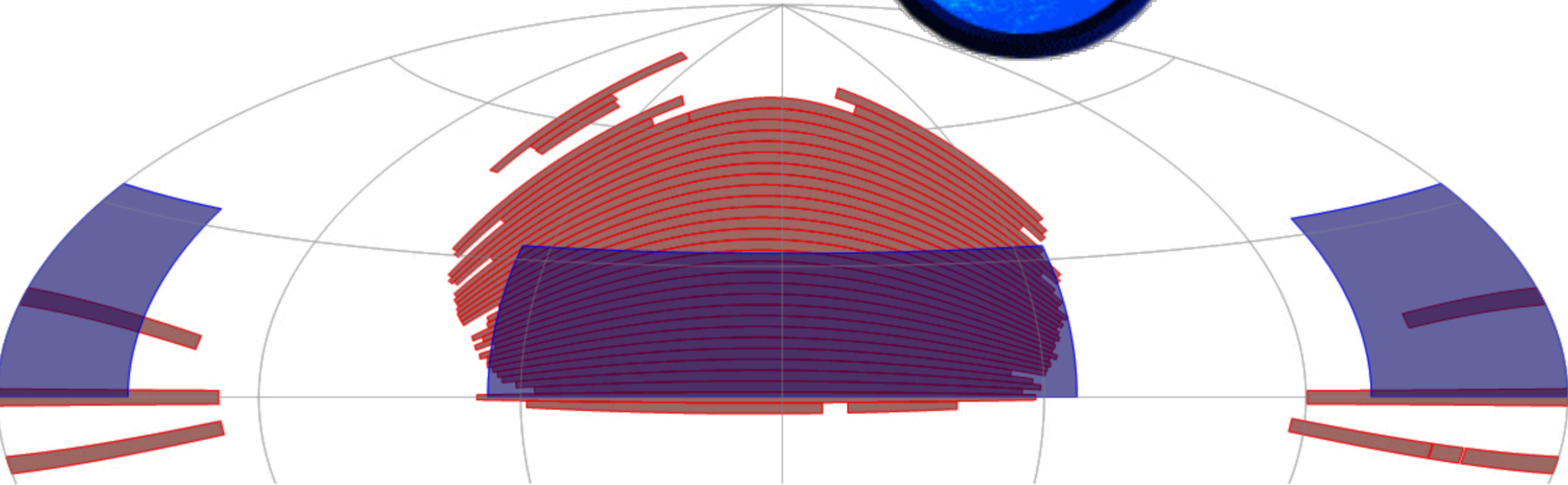
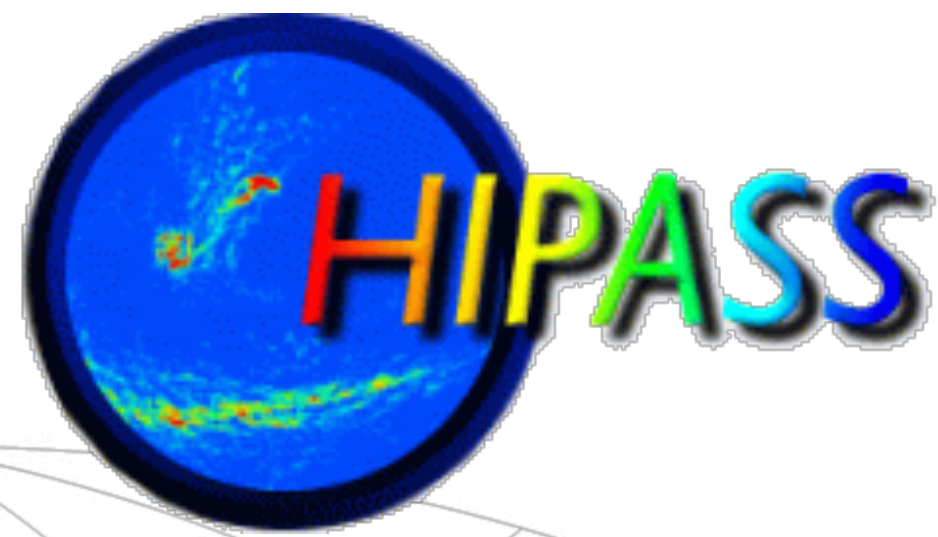


Matilde's talk

neutral hydrogen through its 21cm line

- strongly forbidden: $t_{1/2} \sim 10^7$ years ...
- ... but VERY abundant — all galaxies have some (even *clouds*?)
- Spectrally isolated \longrightarrow no confusion/ interlopers!
- Small obscuration \longrightarrow no atmospheric effects!
- All the way up to EoR (and beyond?)





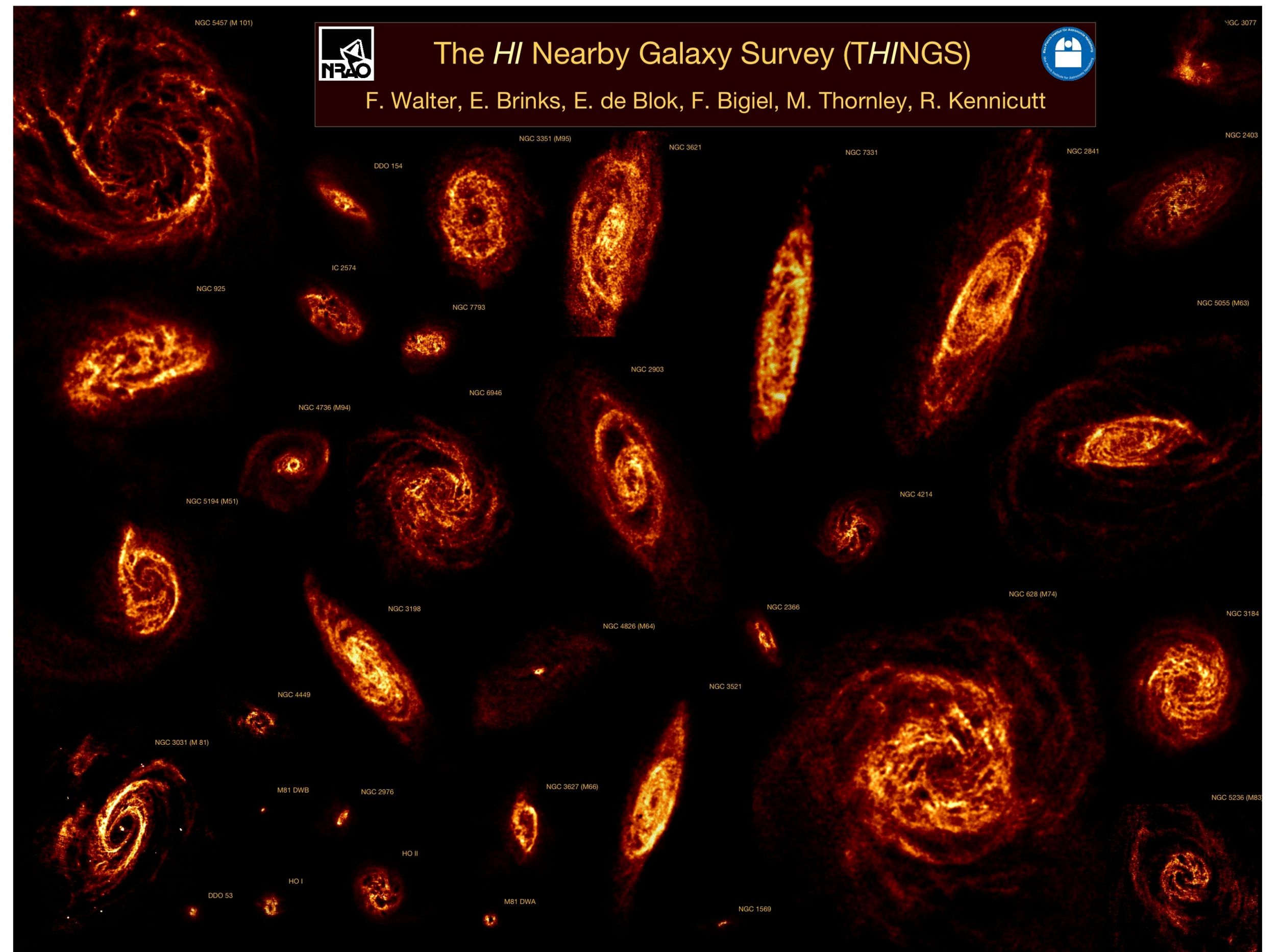
The Arecibo Legacy Fast ALFA Survey

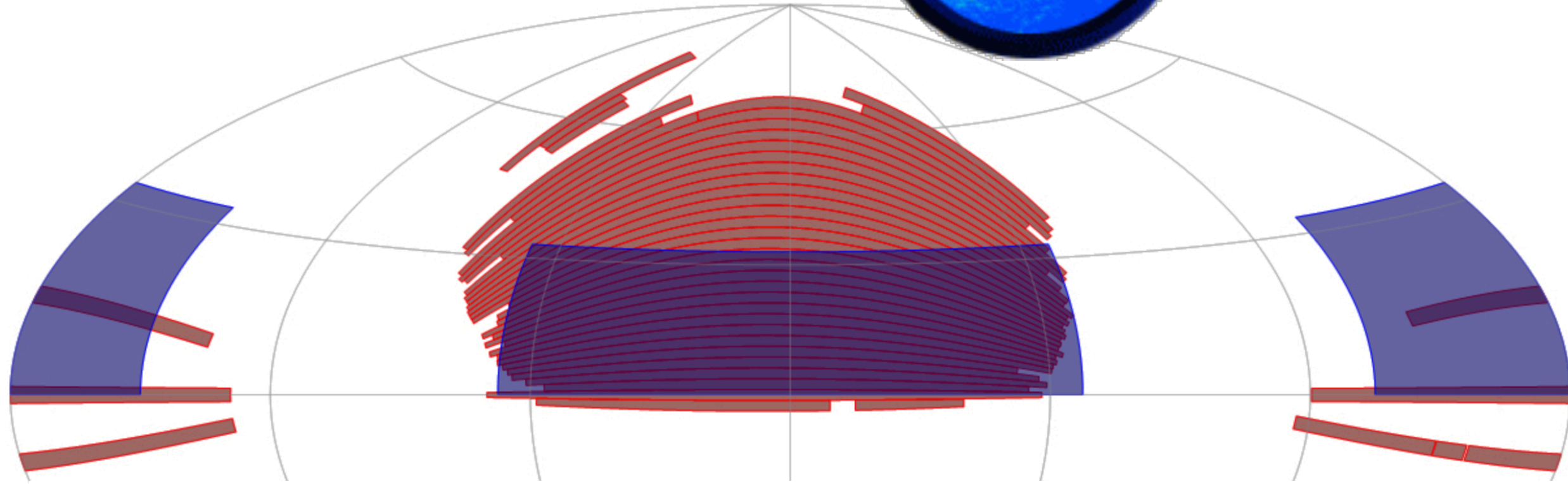
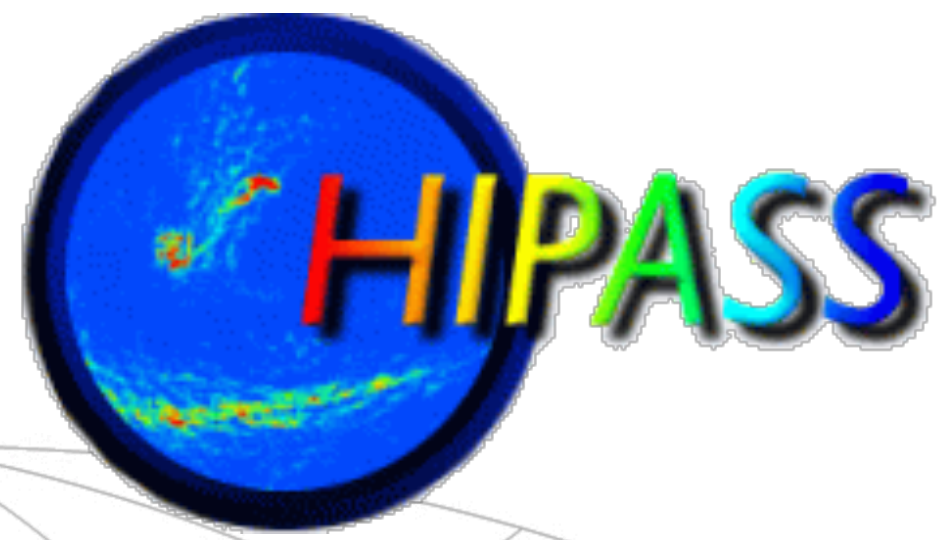


The *HI* Nearby Galaxy Survey (*THINGS*)



F. Walter, E. Brinks, E. de Blok, F. Bigiel, M. Thornley, R. Kennicutt





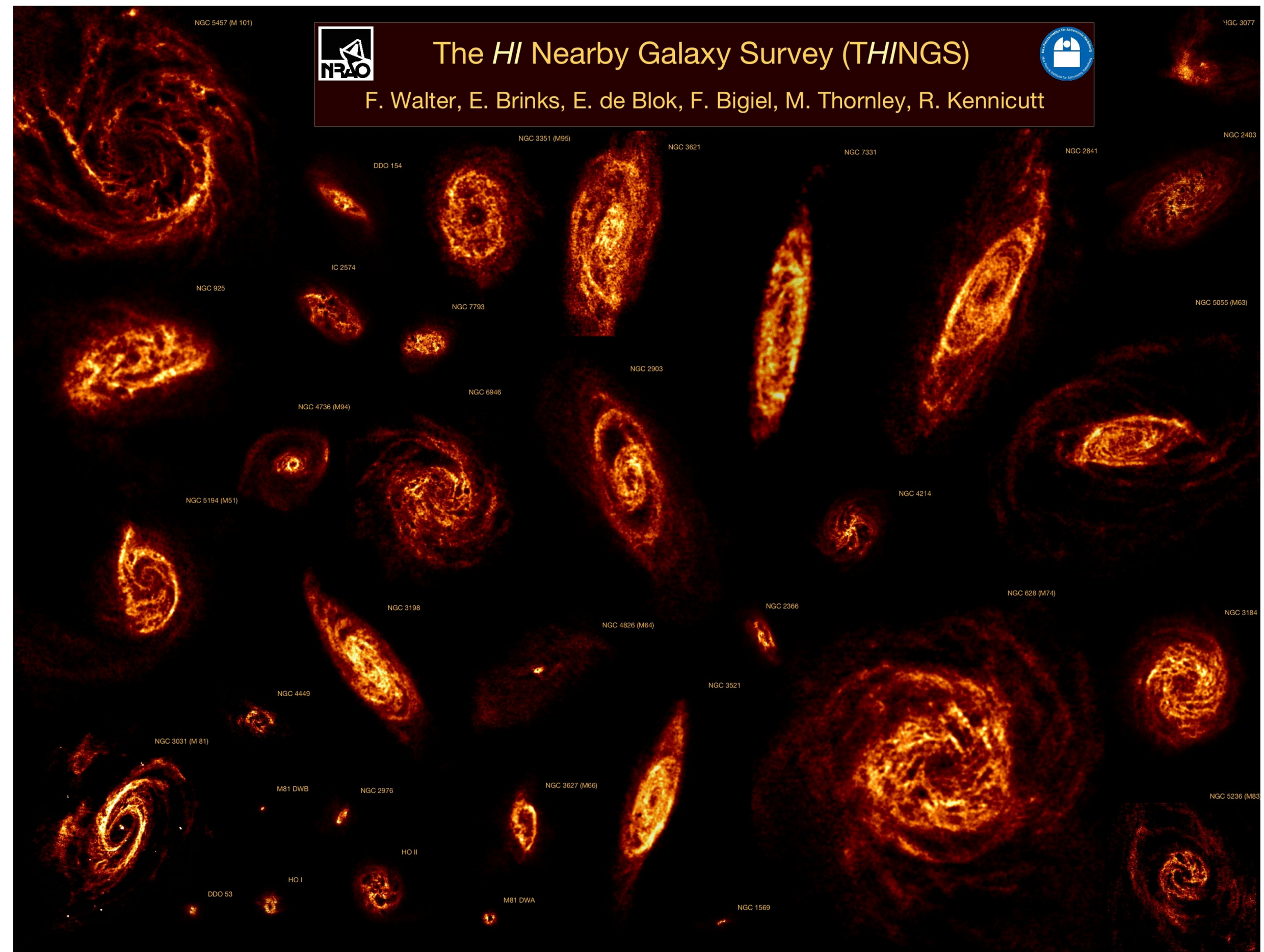
The Arecibo Legacy Fast ALFA Survey



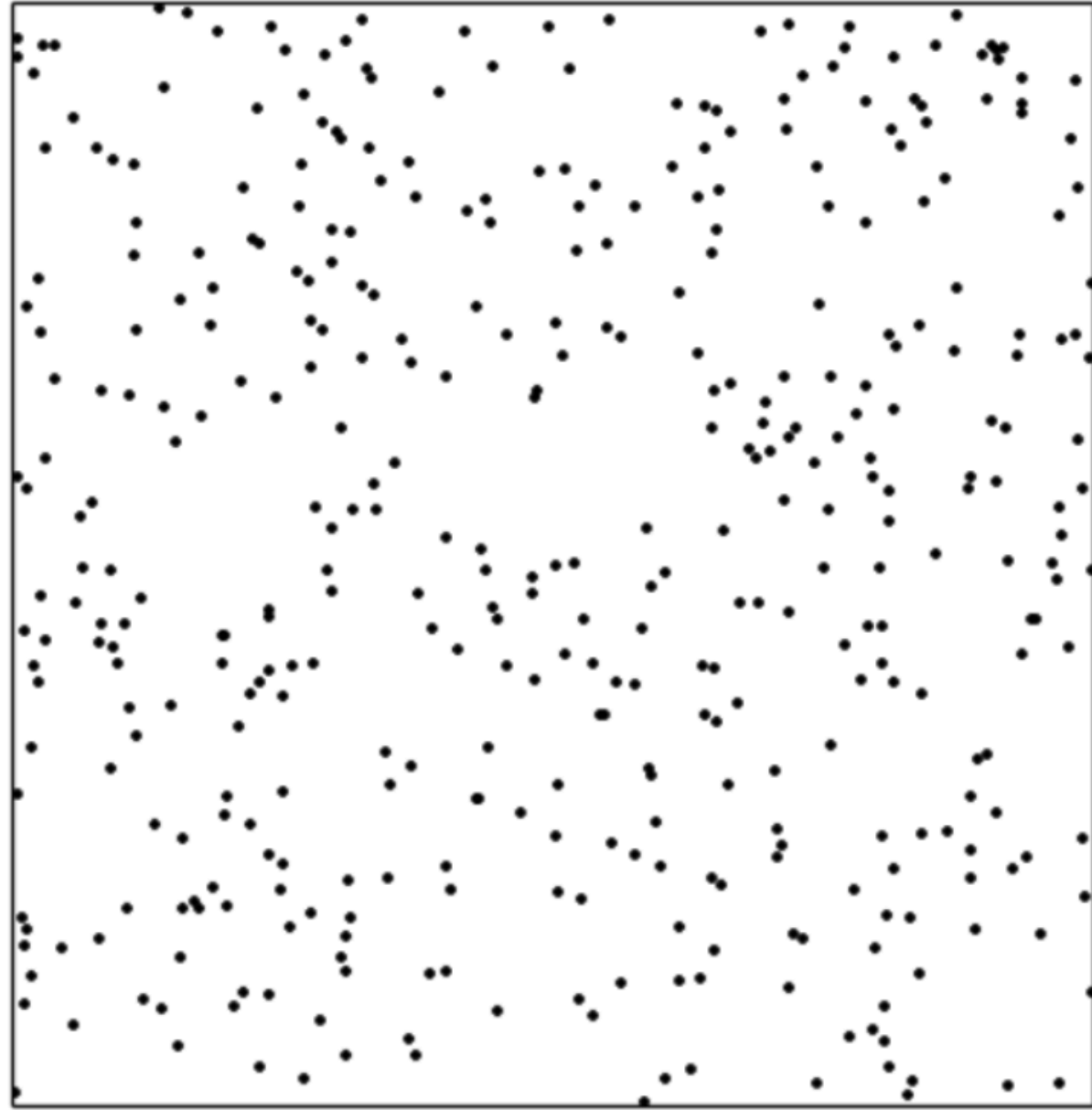
The *HI* Nearby Galaxy Survey (*THINGS*)



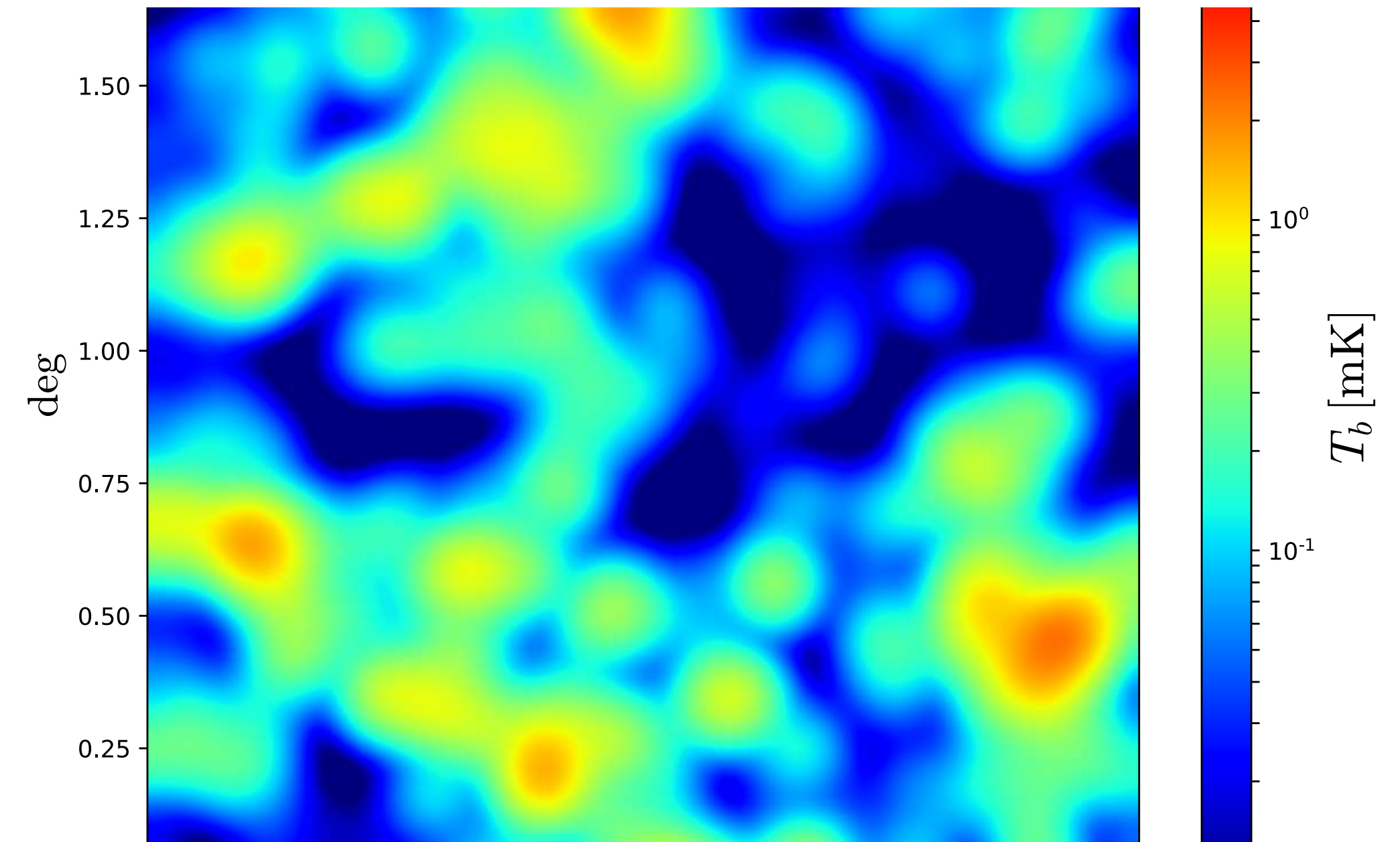
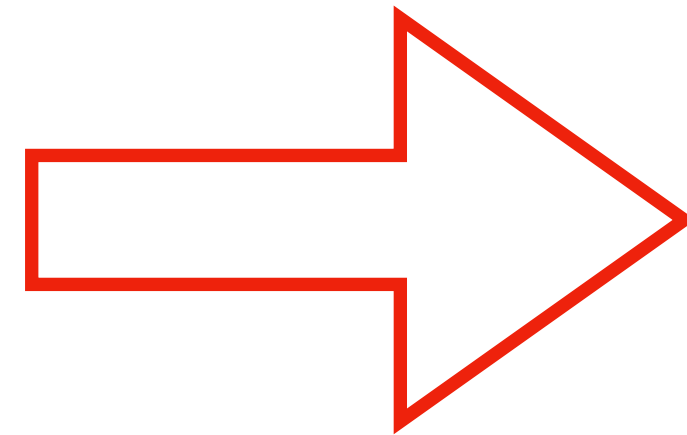
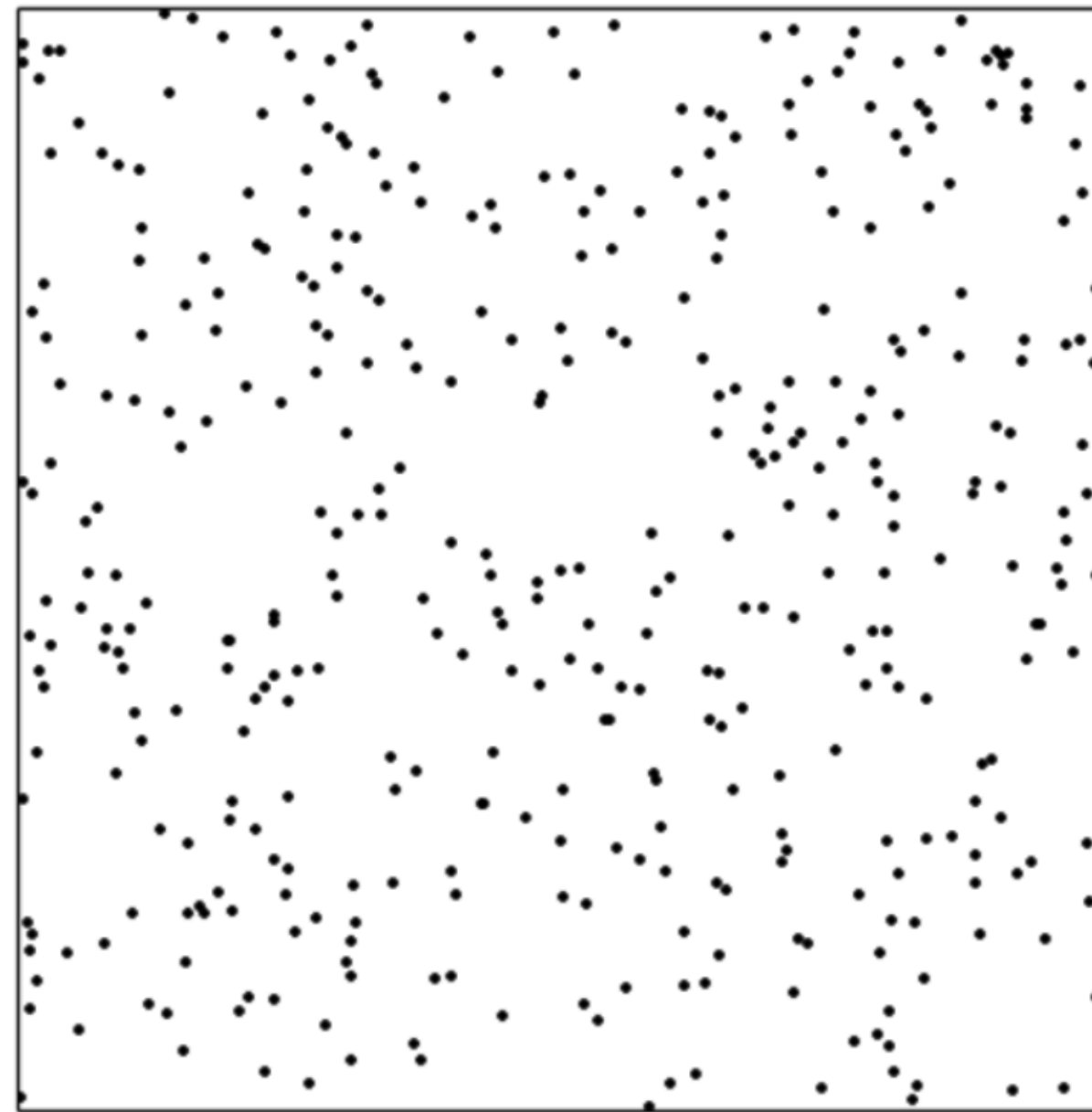
F. Walter, E. Brinks, E. de Blok, F. Bigiel, M. Thornley, R. Kennicutt



Record $z = 0.376$ detection of 21 cm emitting galaxy with 178 hours from VLA [Fernández et al, 2016]

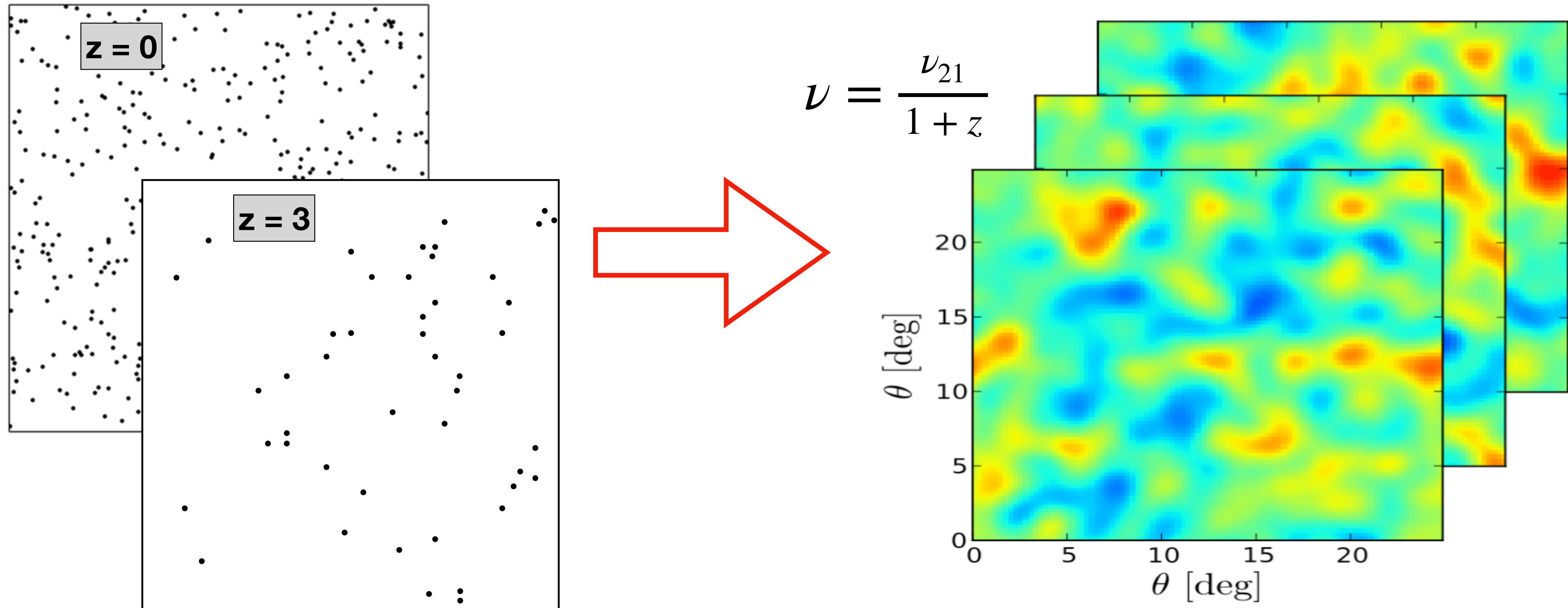


21-cm intensity mapping



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

21-cm intensity mapping



Big volumes (for cheap) and high redshift resolution

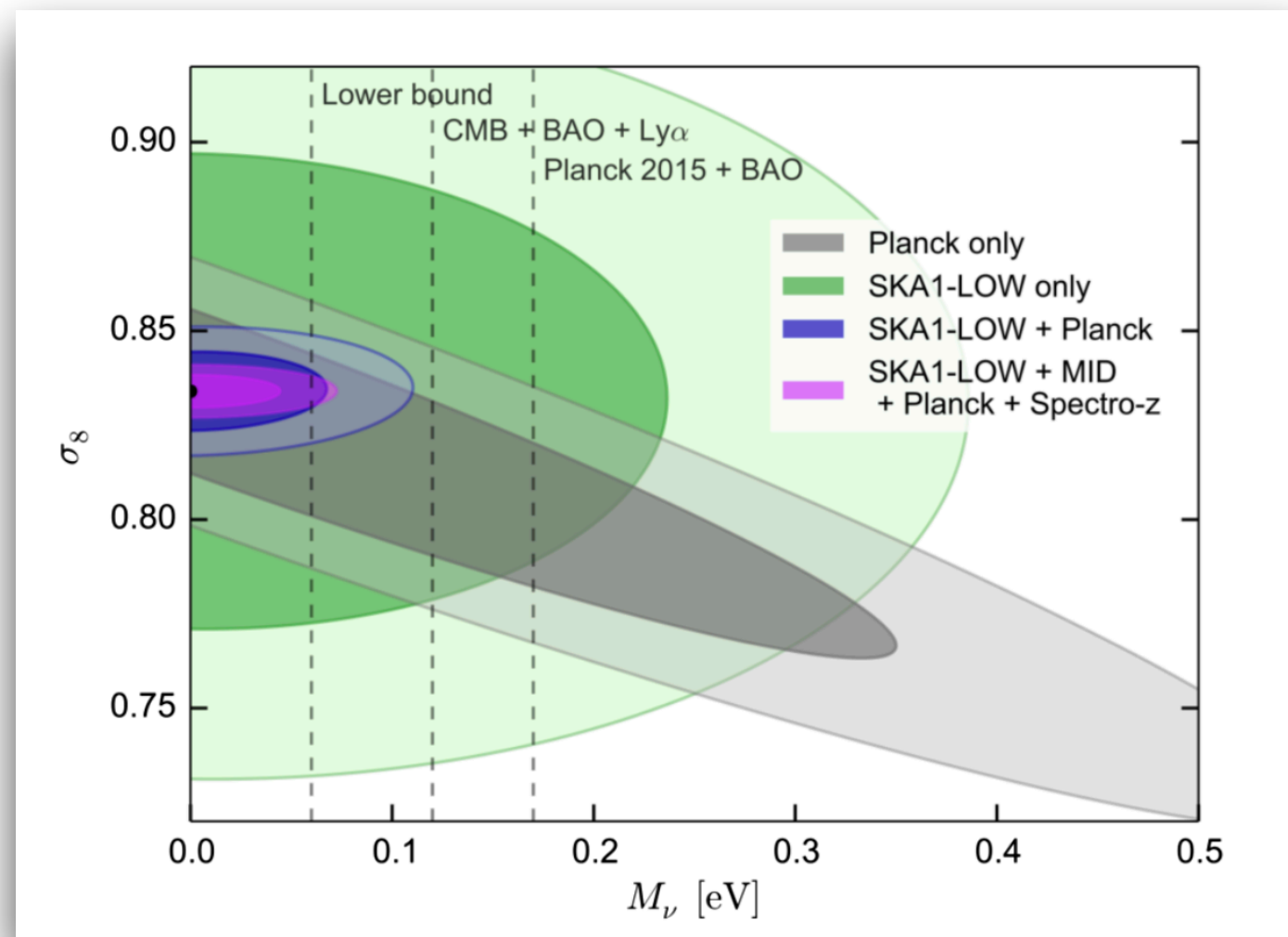
HI intensity mapping with the SKAO

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

Intensity mapping with the SKAO



Proposed SKA1 Cosmology Surveys

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

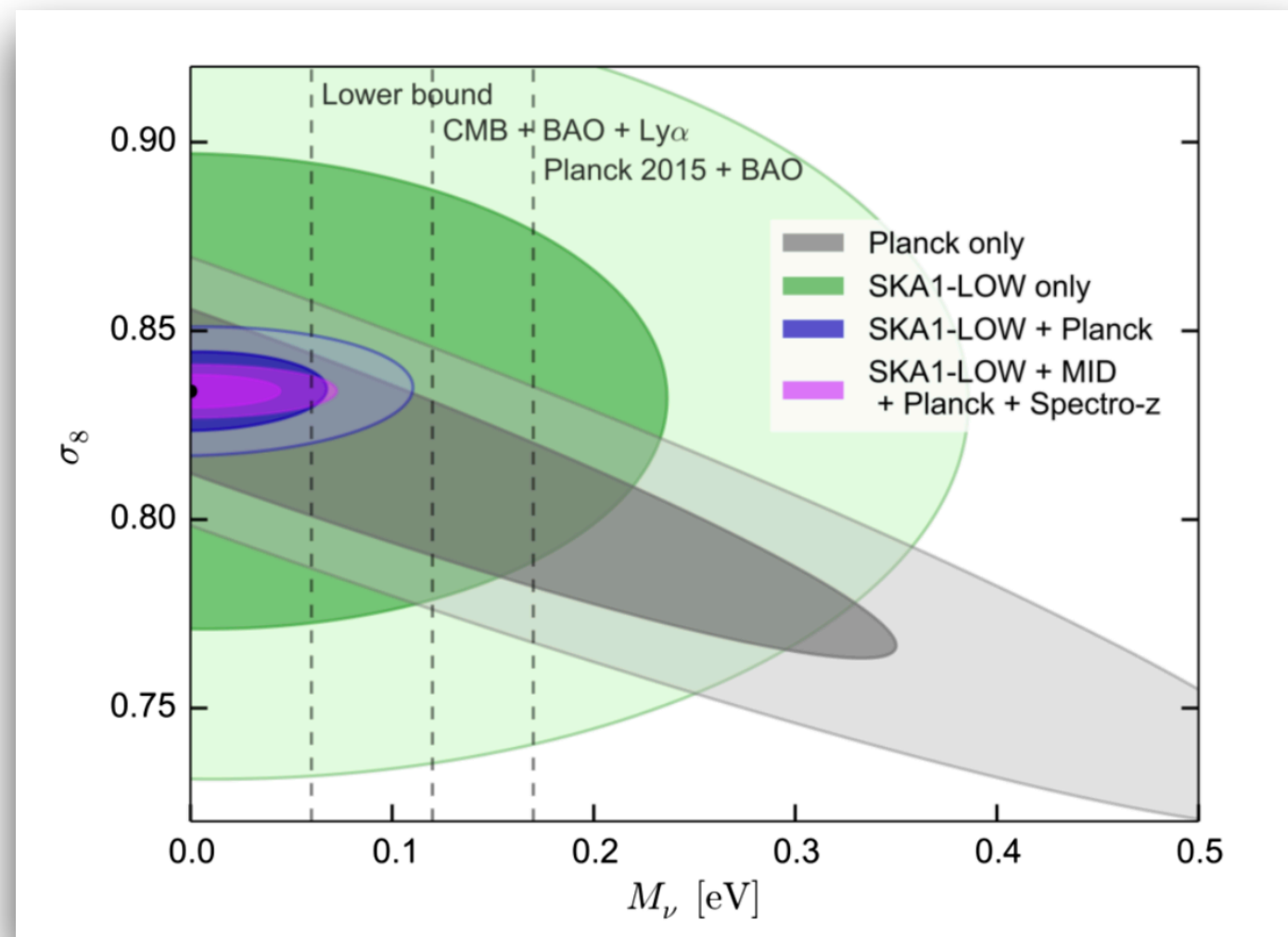
- Continuum galaxy survey with ~60 million objects

- 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

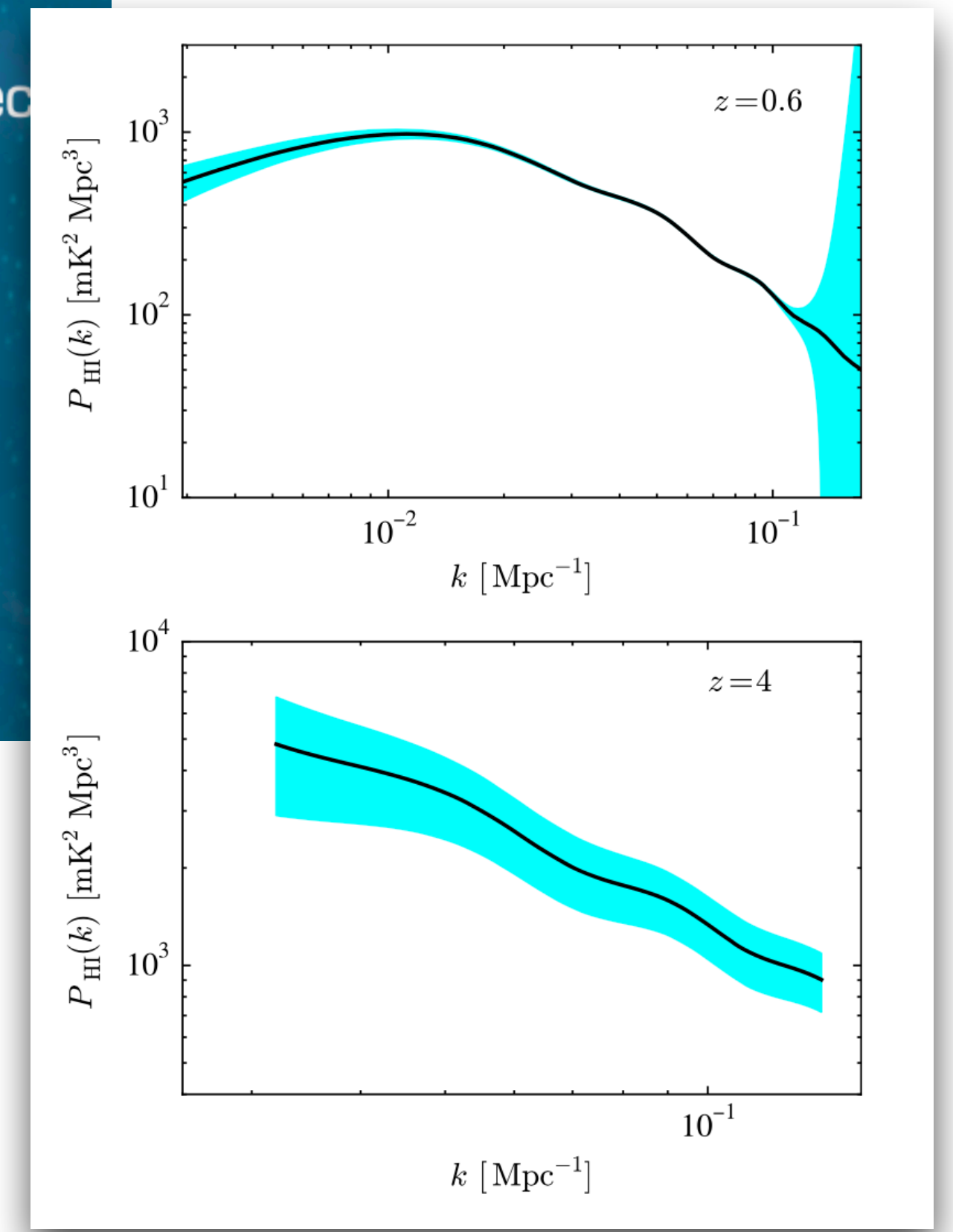
Cosmology with Phase 1 of the Square Kilometre Array **Red Book** 2018:
Technical specifications and performance forecasts

Intensity mapping with the SKAO



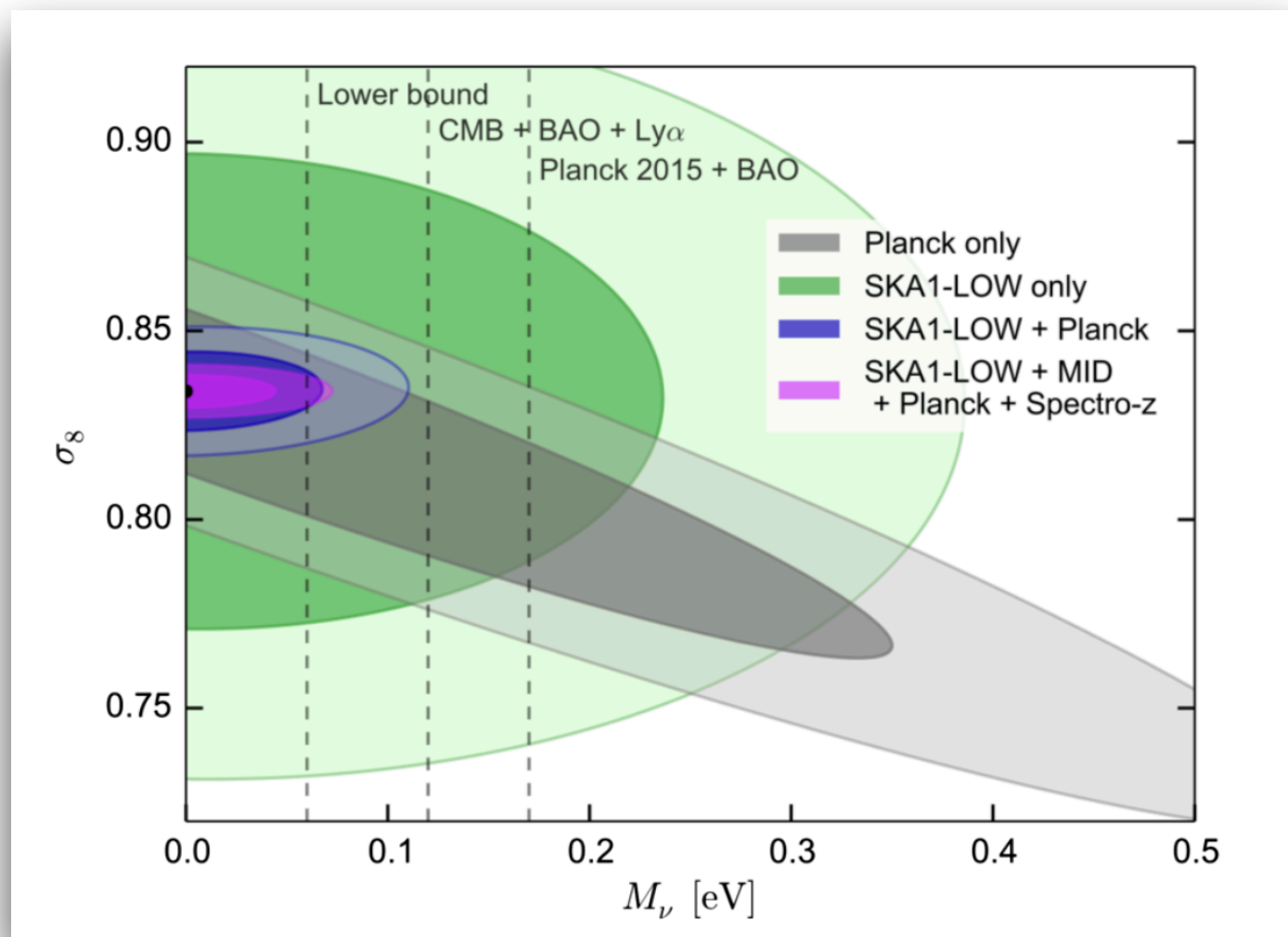
SKA1 Cosmology Surveys

- Wide-Deep Survey of 5,000 deg² at 0.95-1.4 GHz for 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



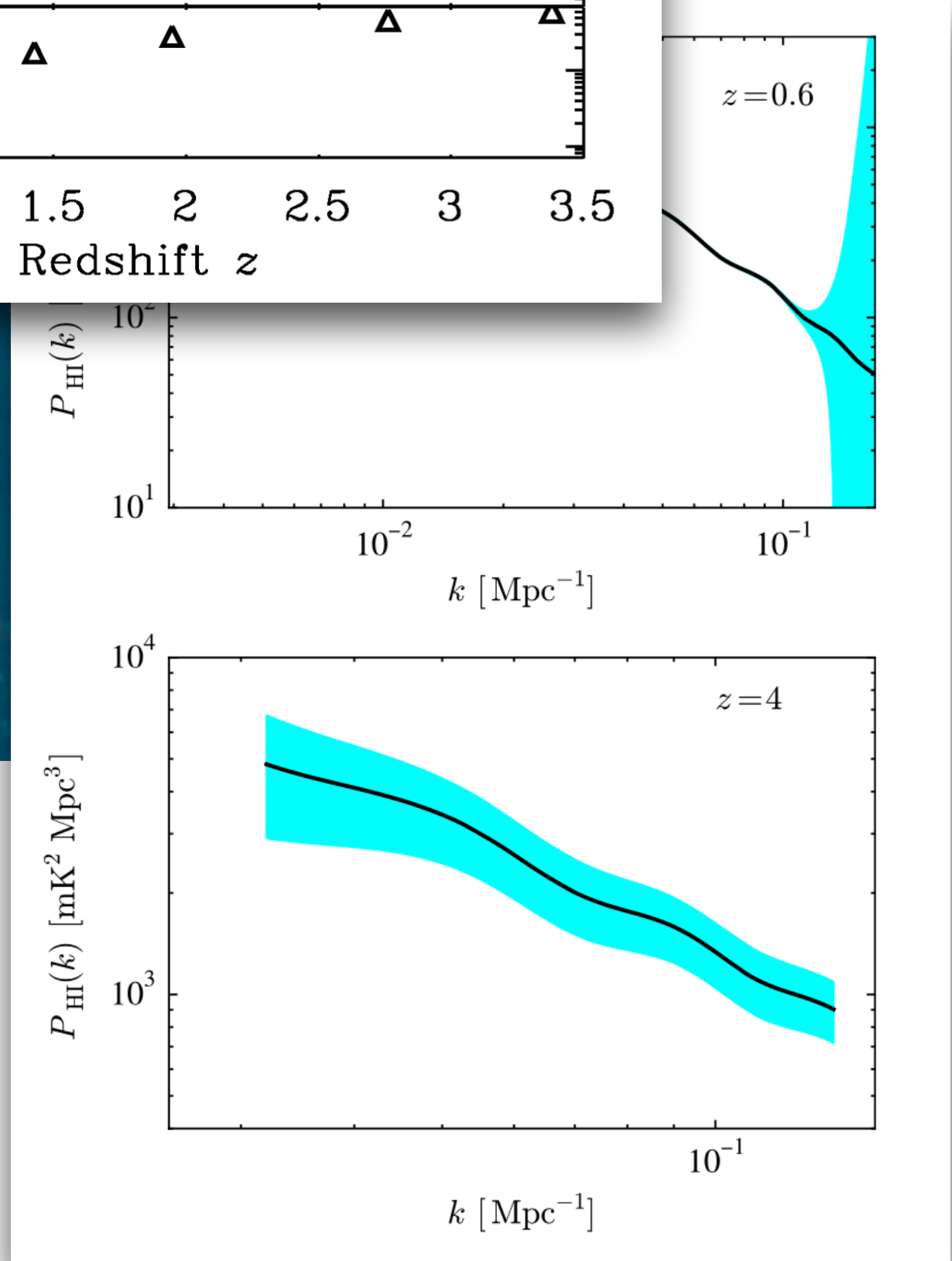
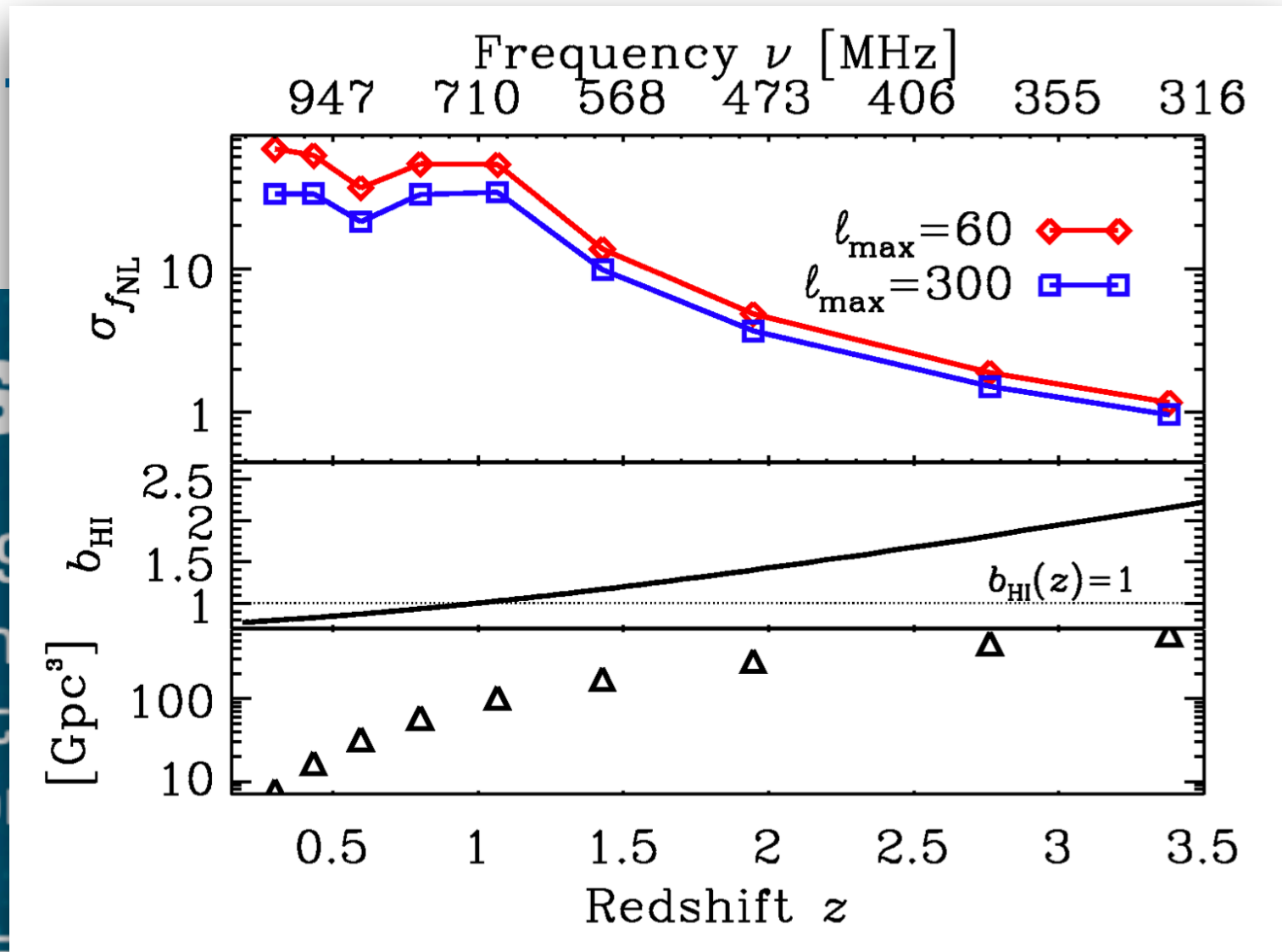
Cosmology with Phase 1 of the Square Kilometre Array **Red Book 2018**:
 Technical specifications and performance forecasts

Intensity mapping with



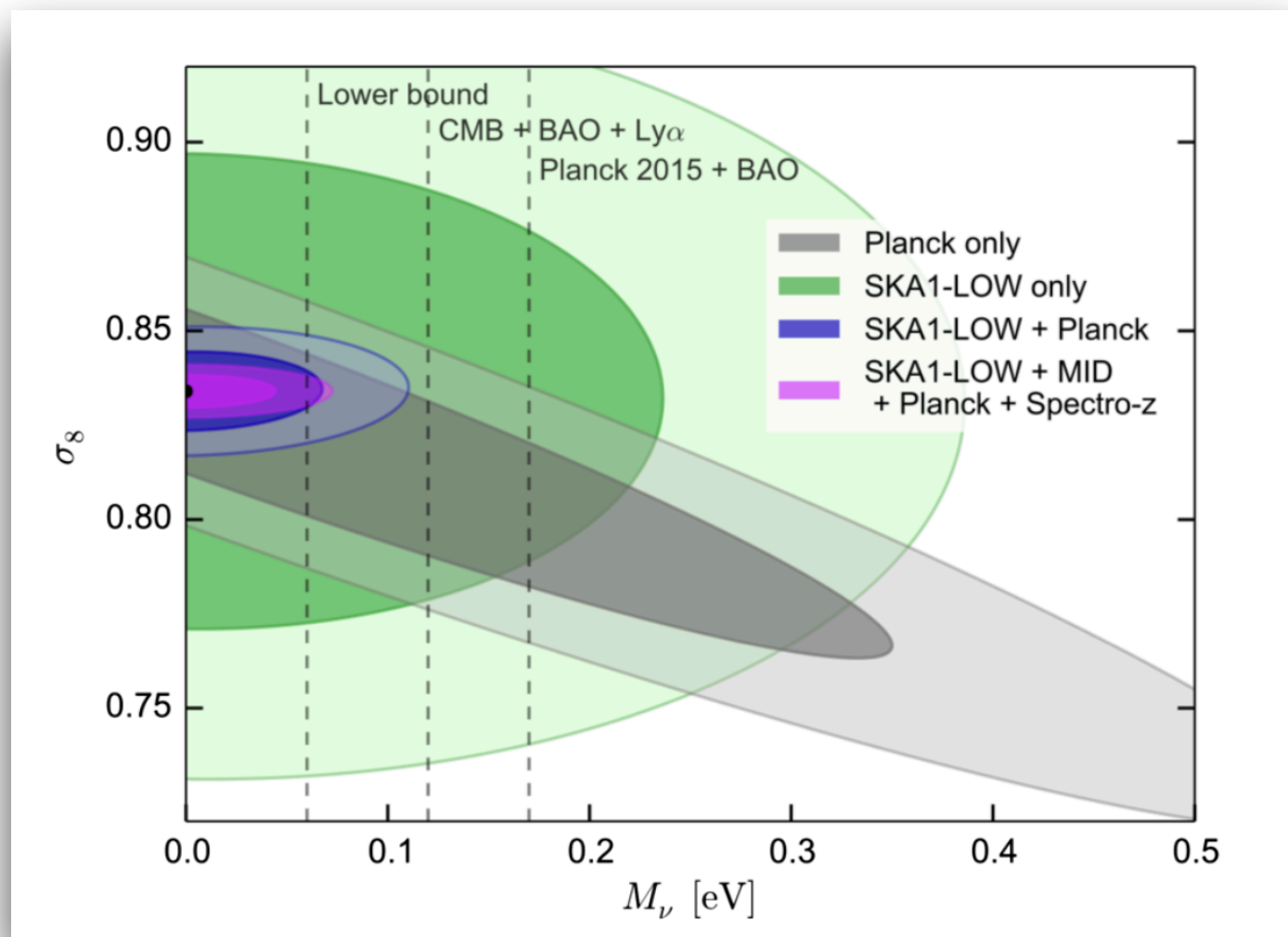
SKA1 Cosmology S

- Continuum galaxy survey with ~60 million objects
- Deep Survey of 5,000 deg² at 0.9-1.6 GHz for galaxy redshift survey with 3.5 million objects
- Weak Lensing shape measurements with 100 million objects
- Continuum galaxy survey with ~60 million objects
- 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
- Deep Survey 100 deg² at 200-350 MHz for
 - HI intensity maps for 3 < z < 6



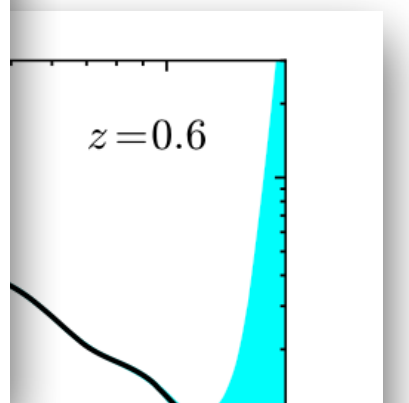
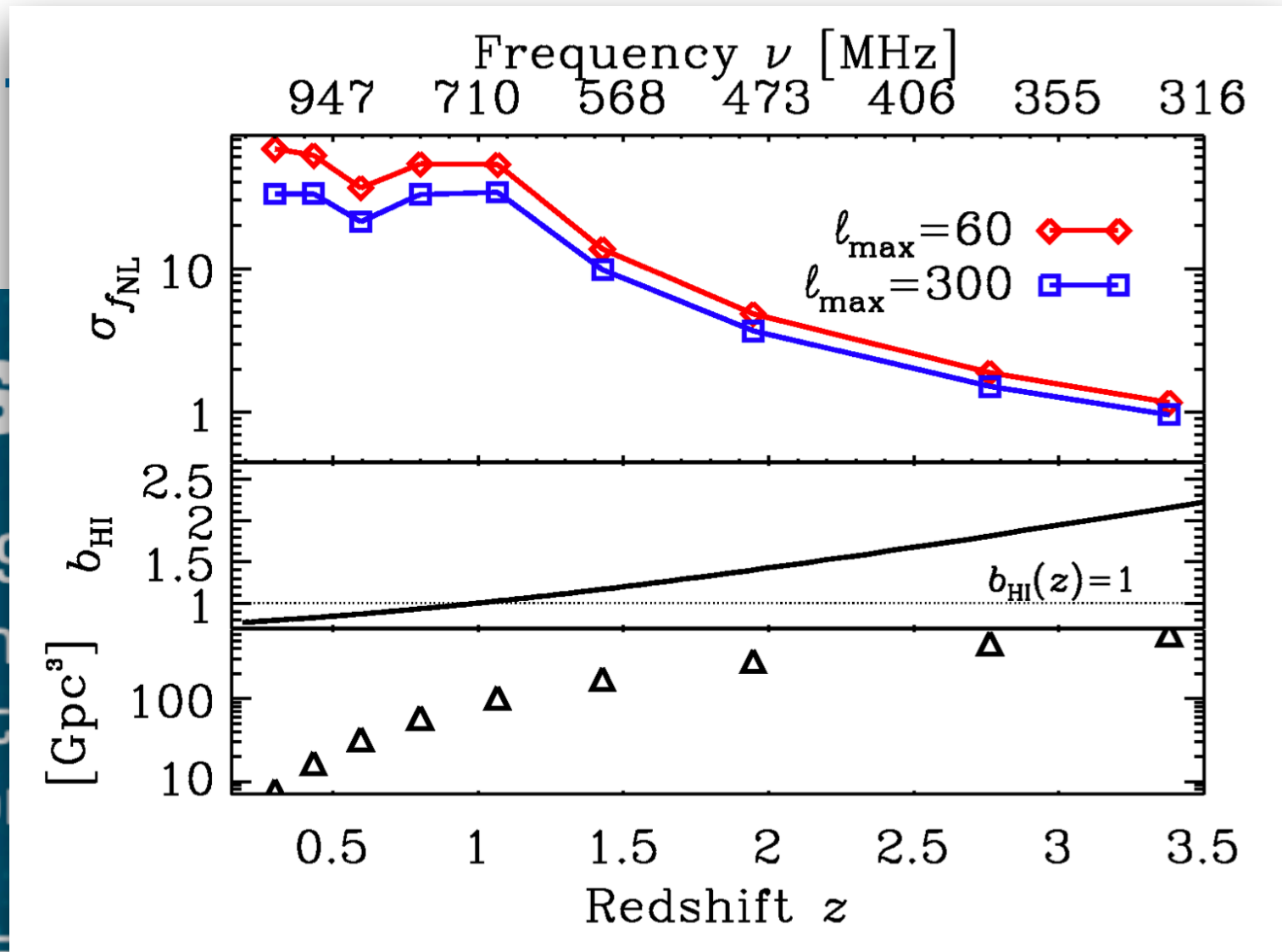
Cosmology with Phase 1 of the Square Kilometre Array **Red Book 2018**:
Technical specifications and performance forecasts

Intensity mapping with



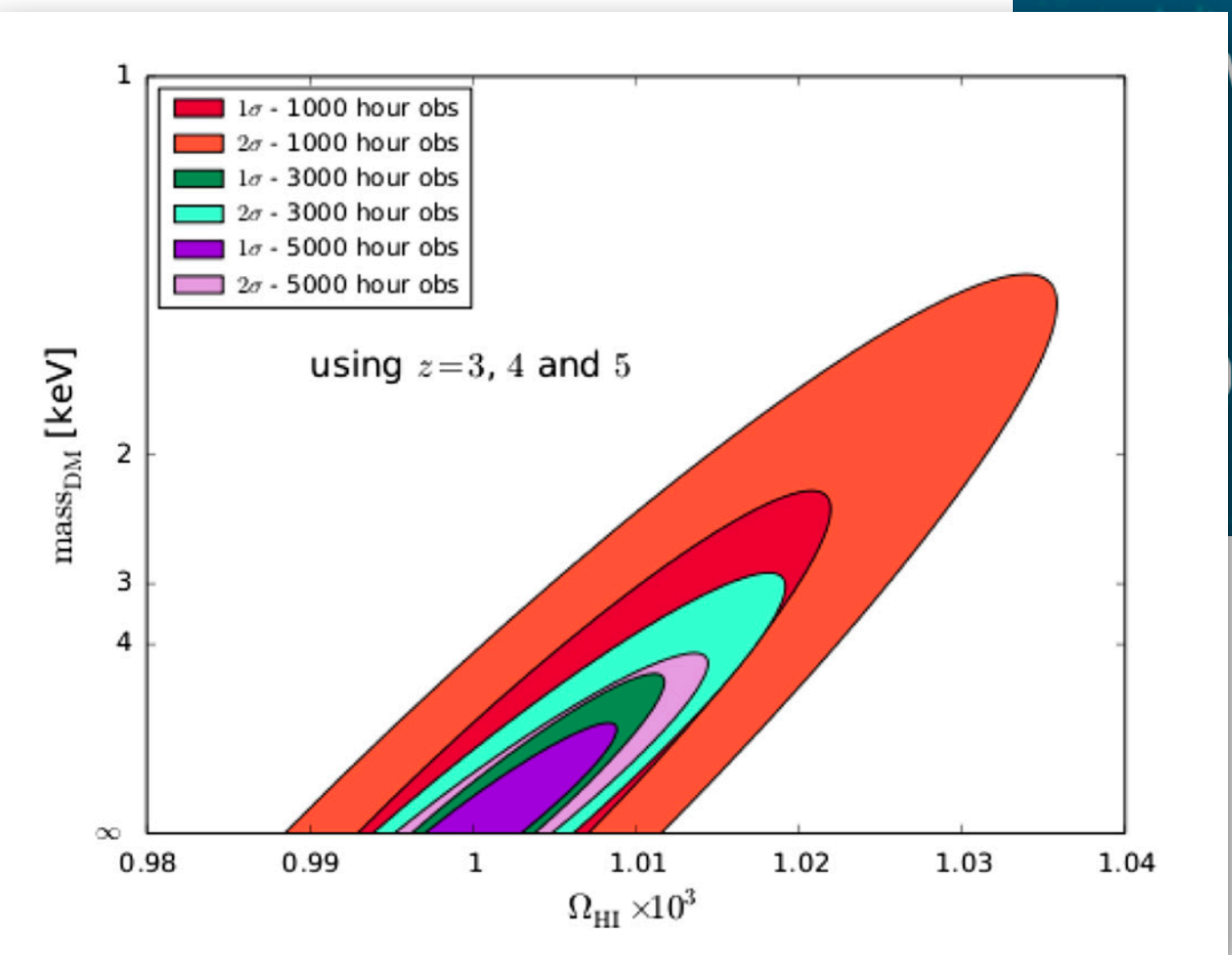
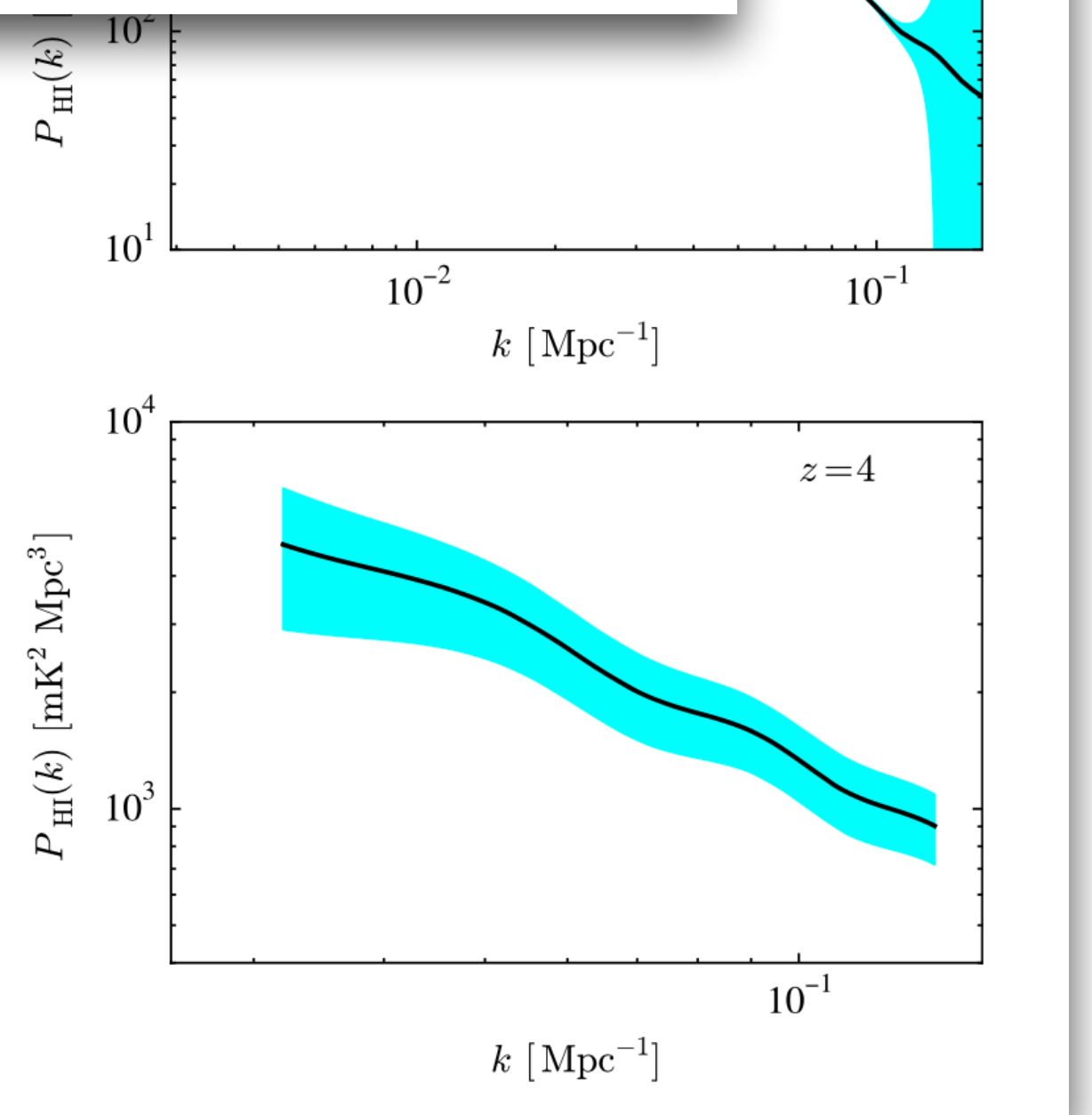
SKA1 Cosmology S

- Deep Survey of 5,000 deg² at 0.9-1.4 GHz for HI intensity maps for 0.35 < z < 3
- Continuum galaxy survey with ~60 million objects



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

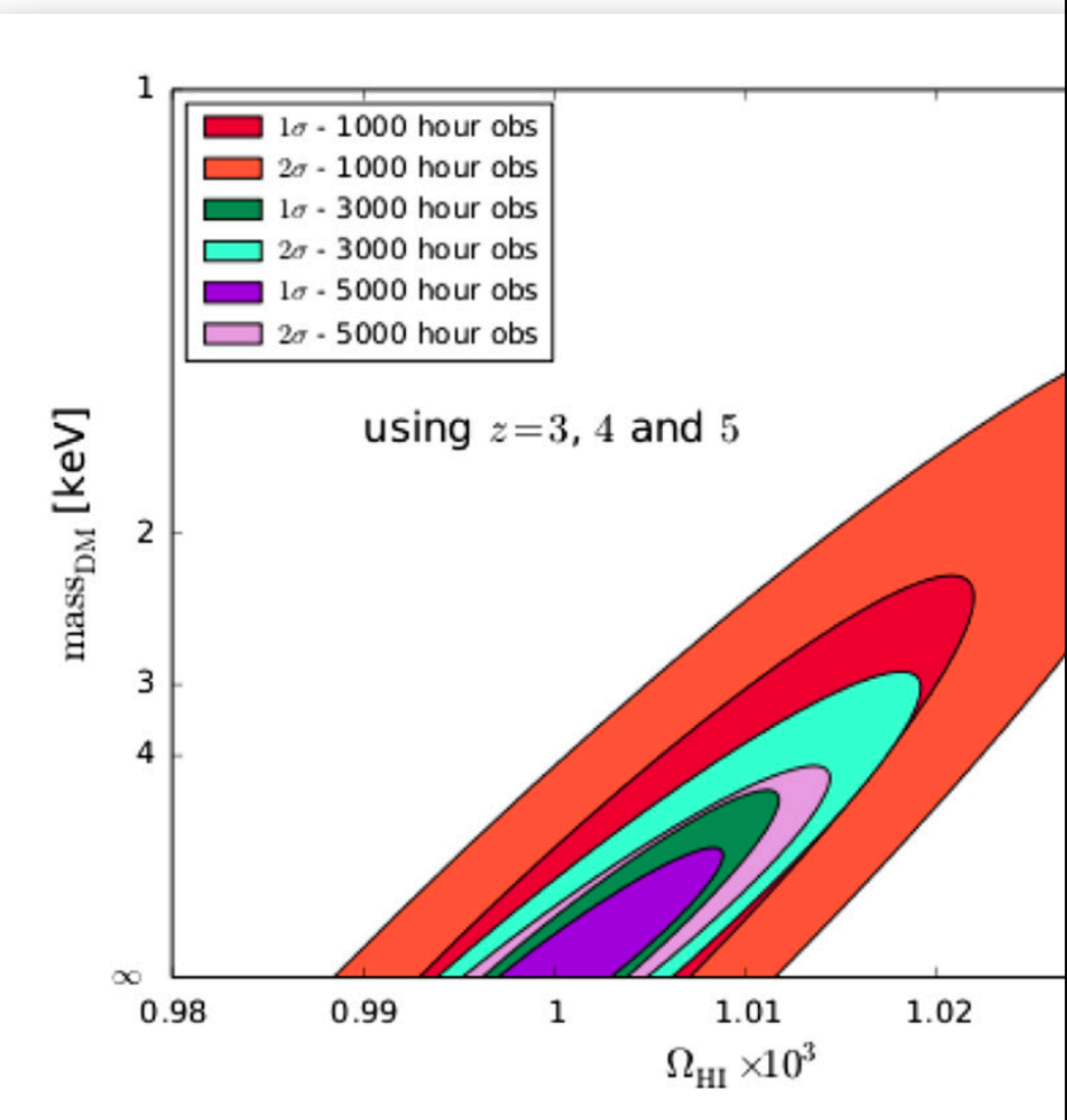
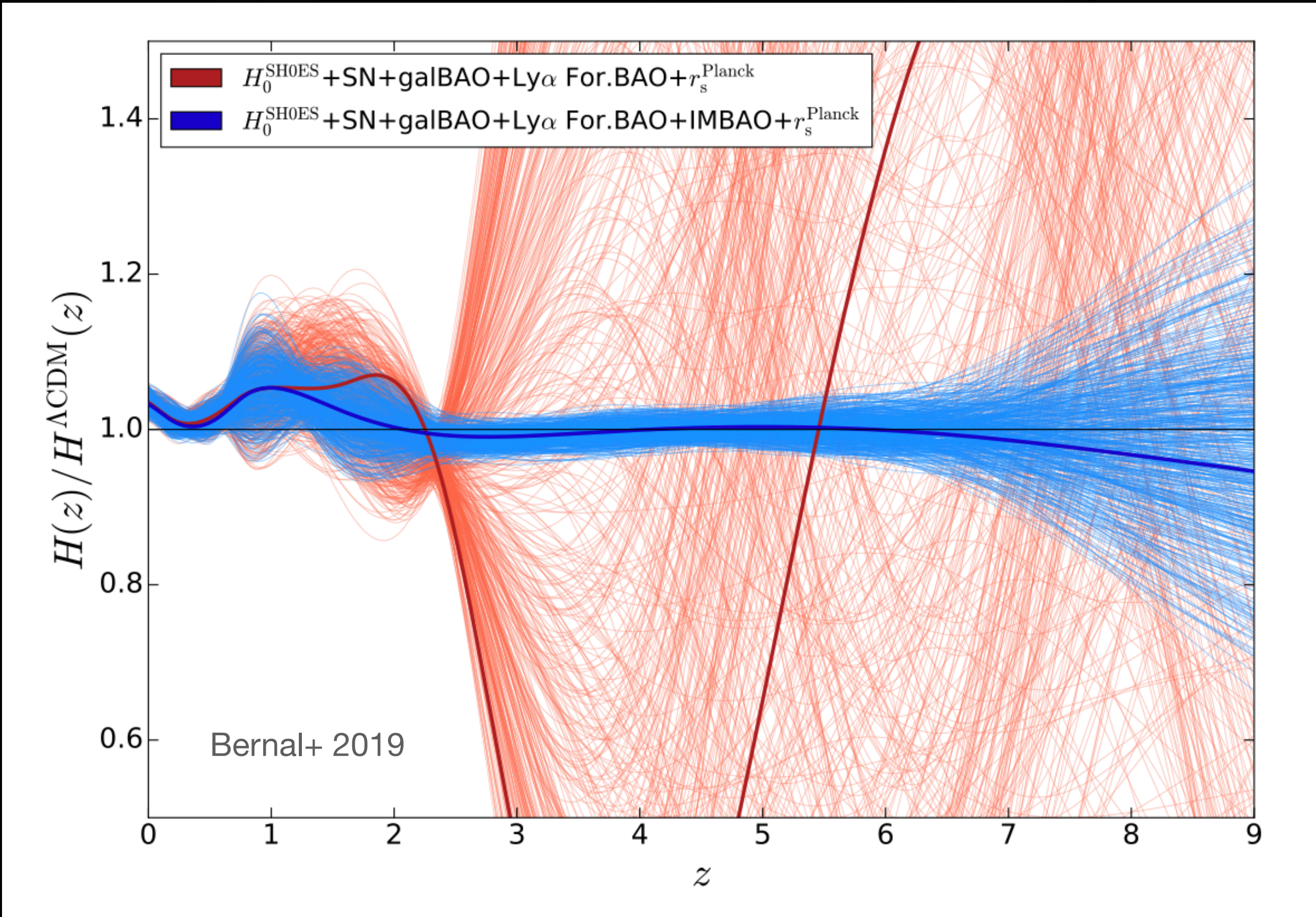
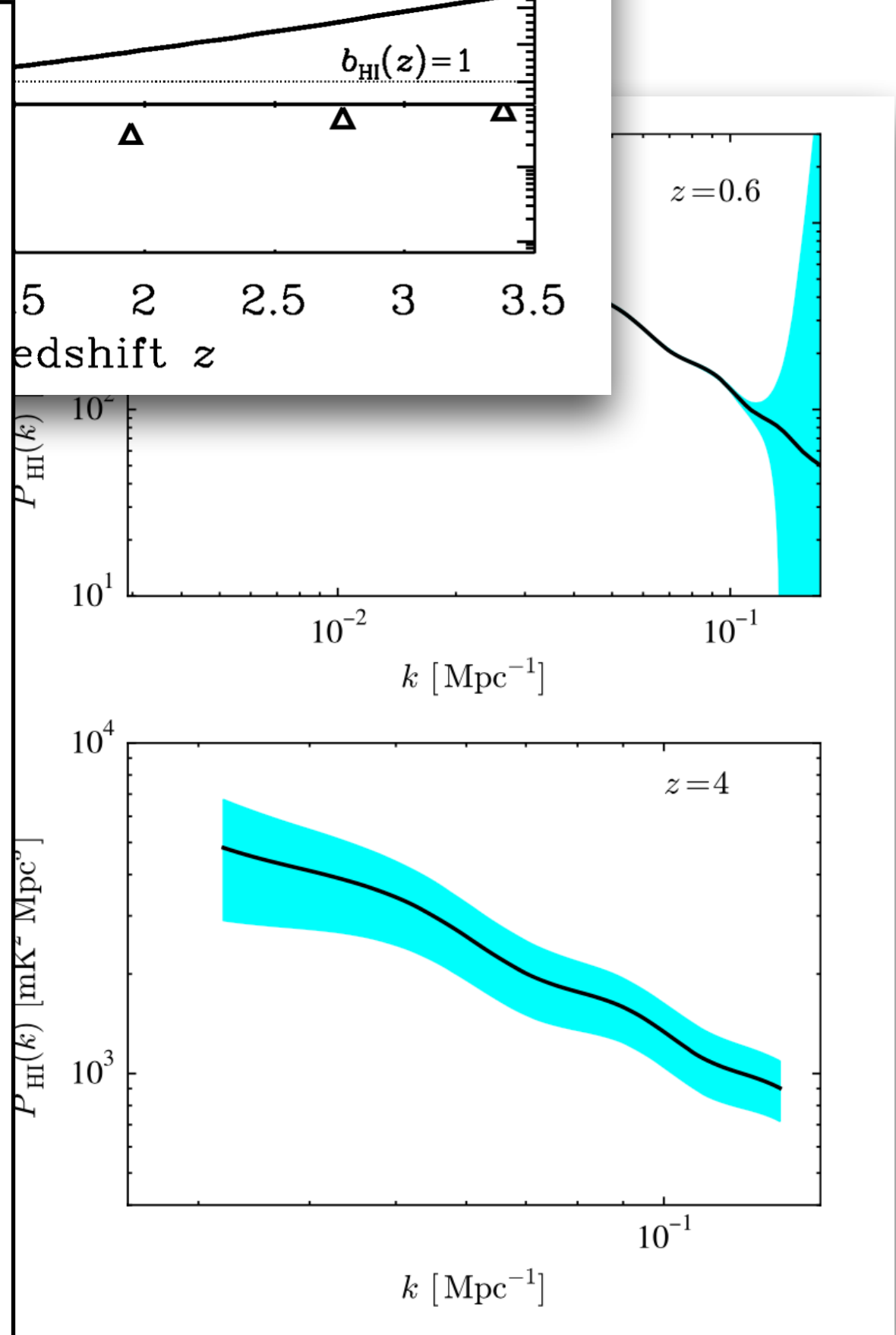
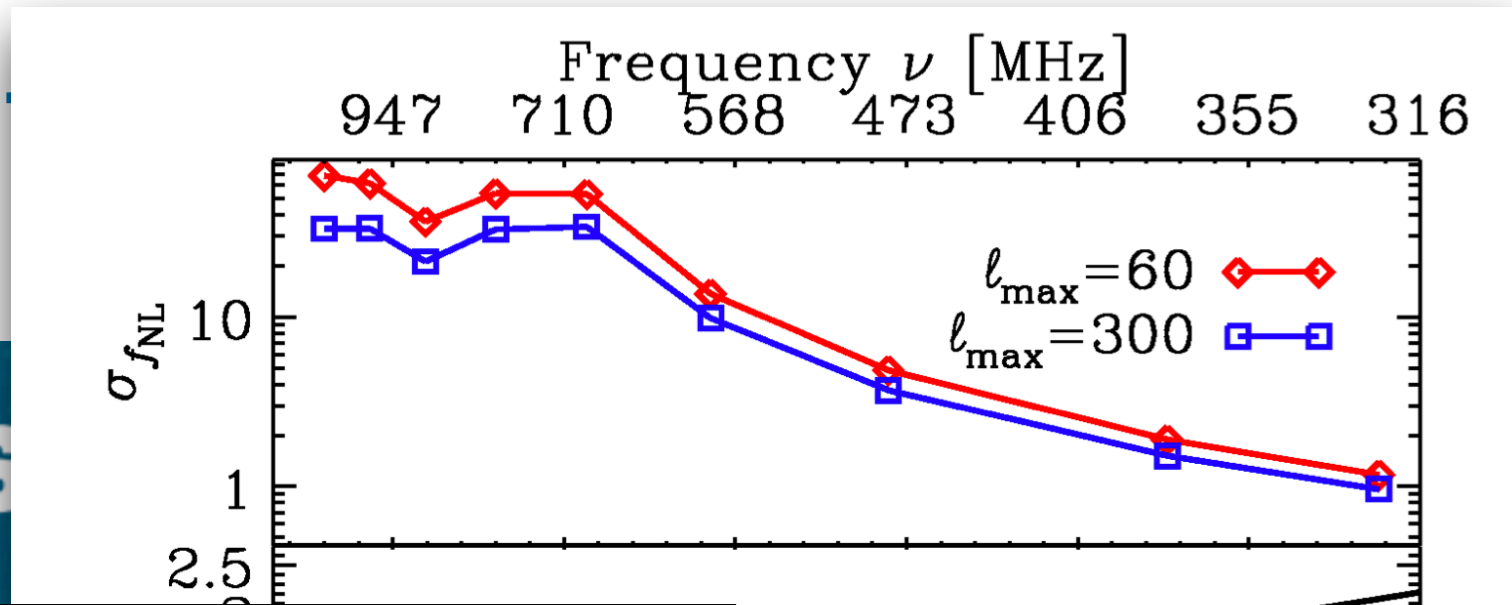
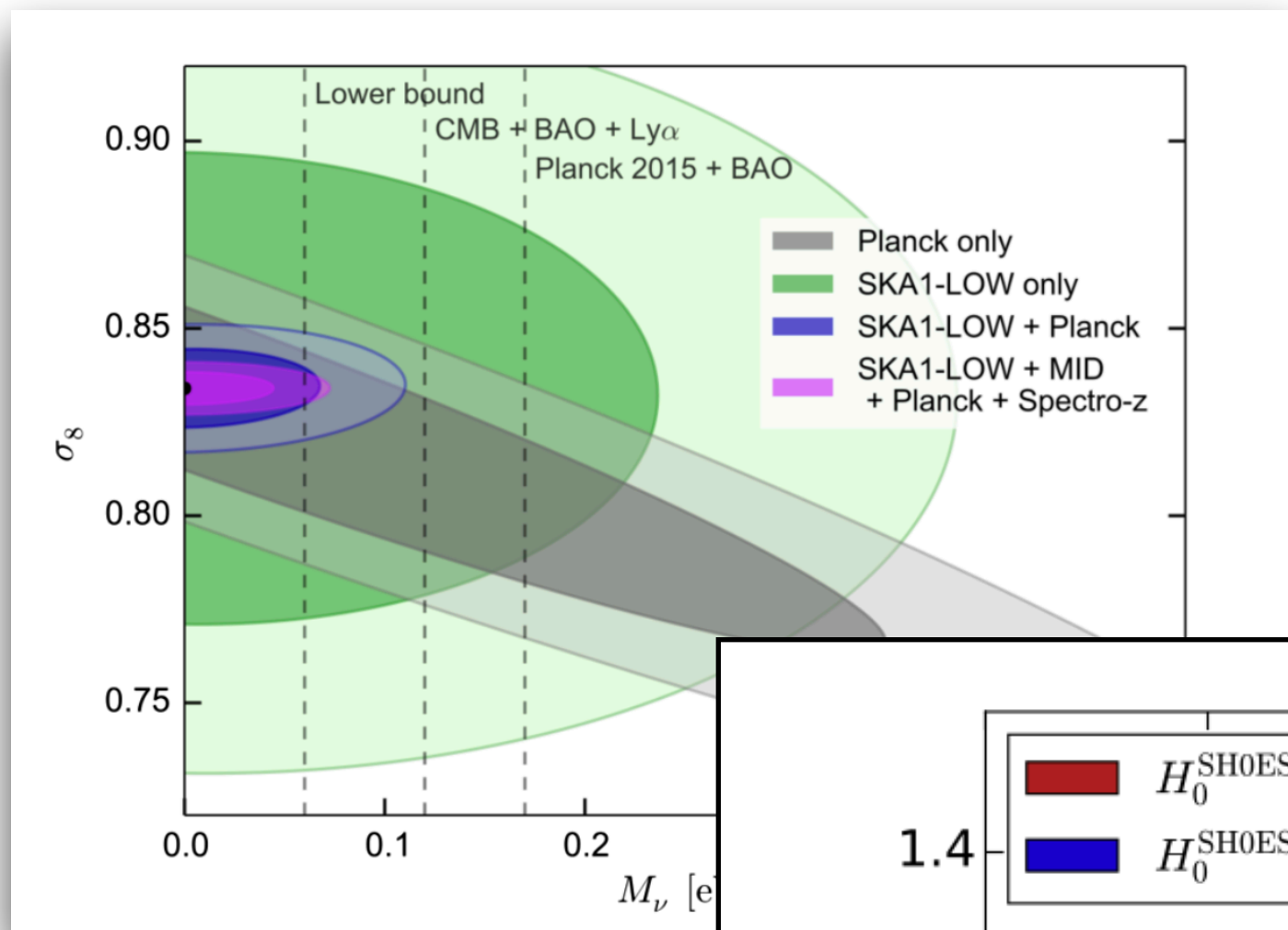
- Deep Survey 100 deg² at 200-350 MHz for HI intensity maps for 3 < z < 6



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

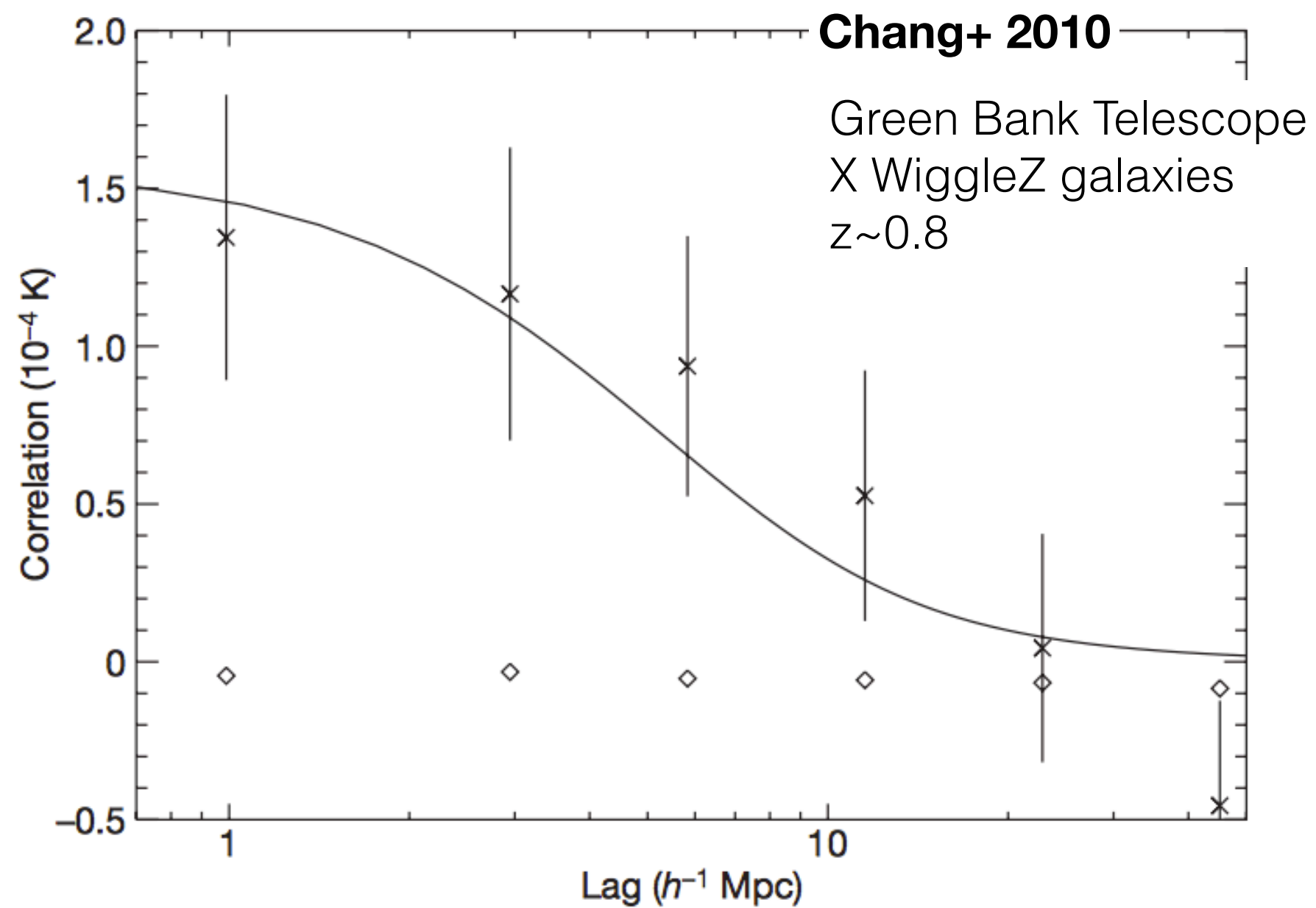
Density mapping with

red SKA1 Cosmology S

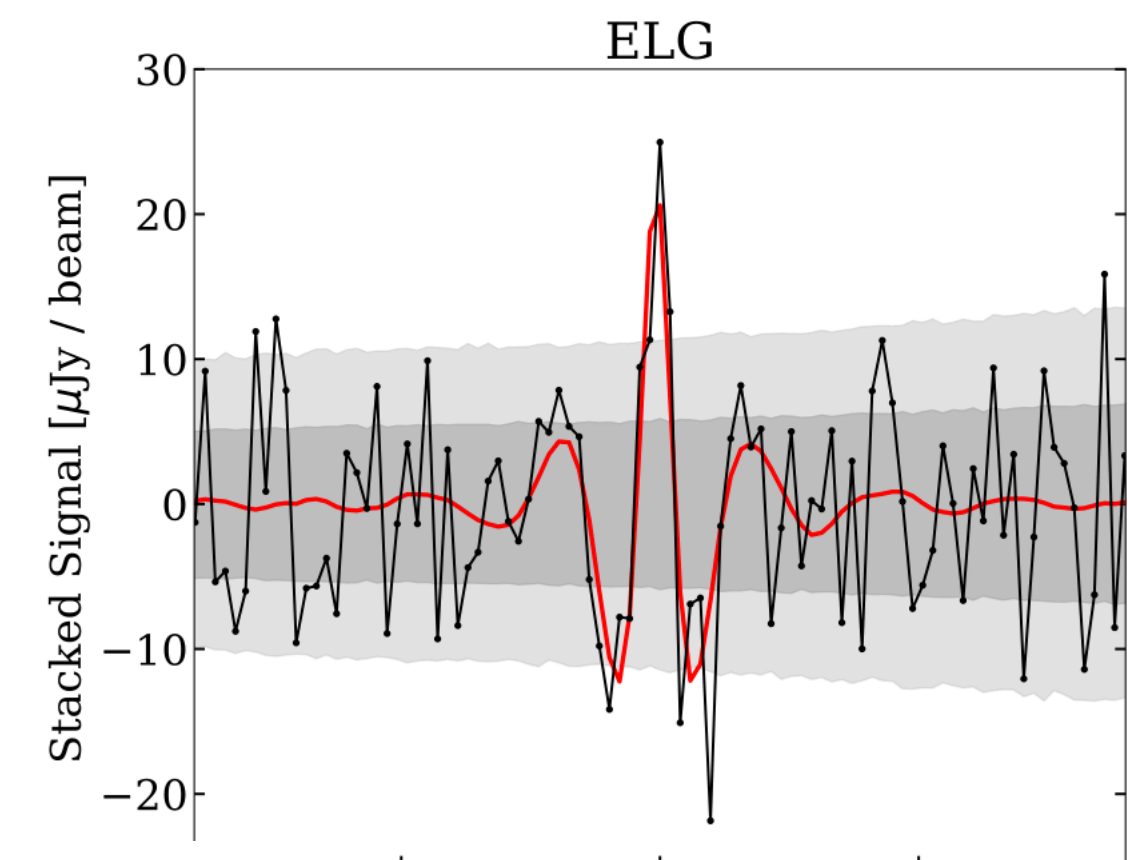
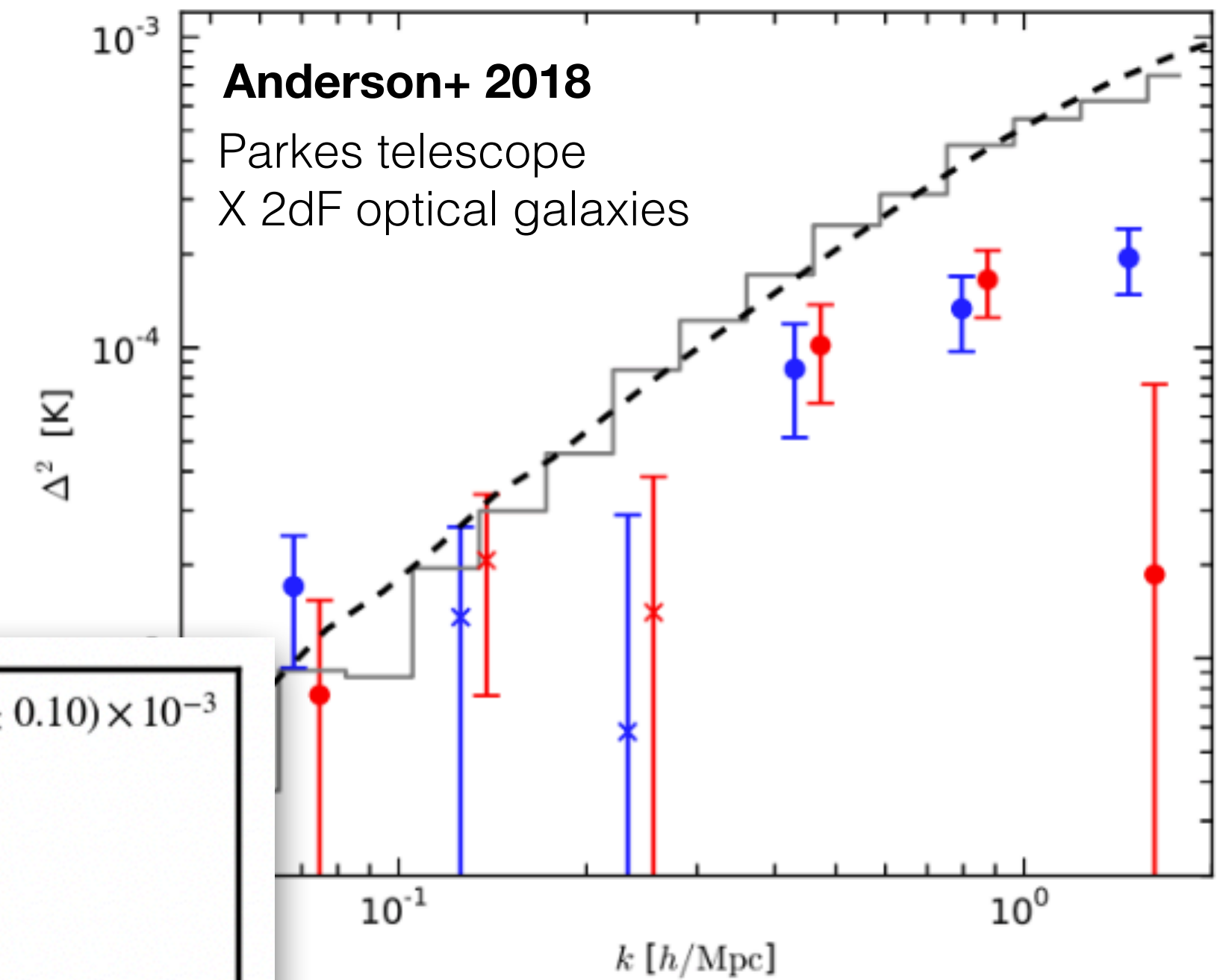
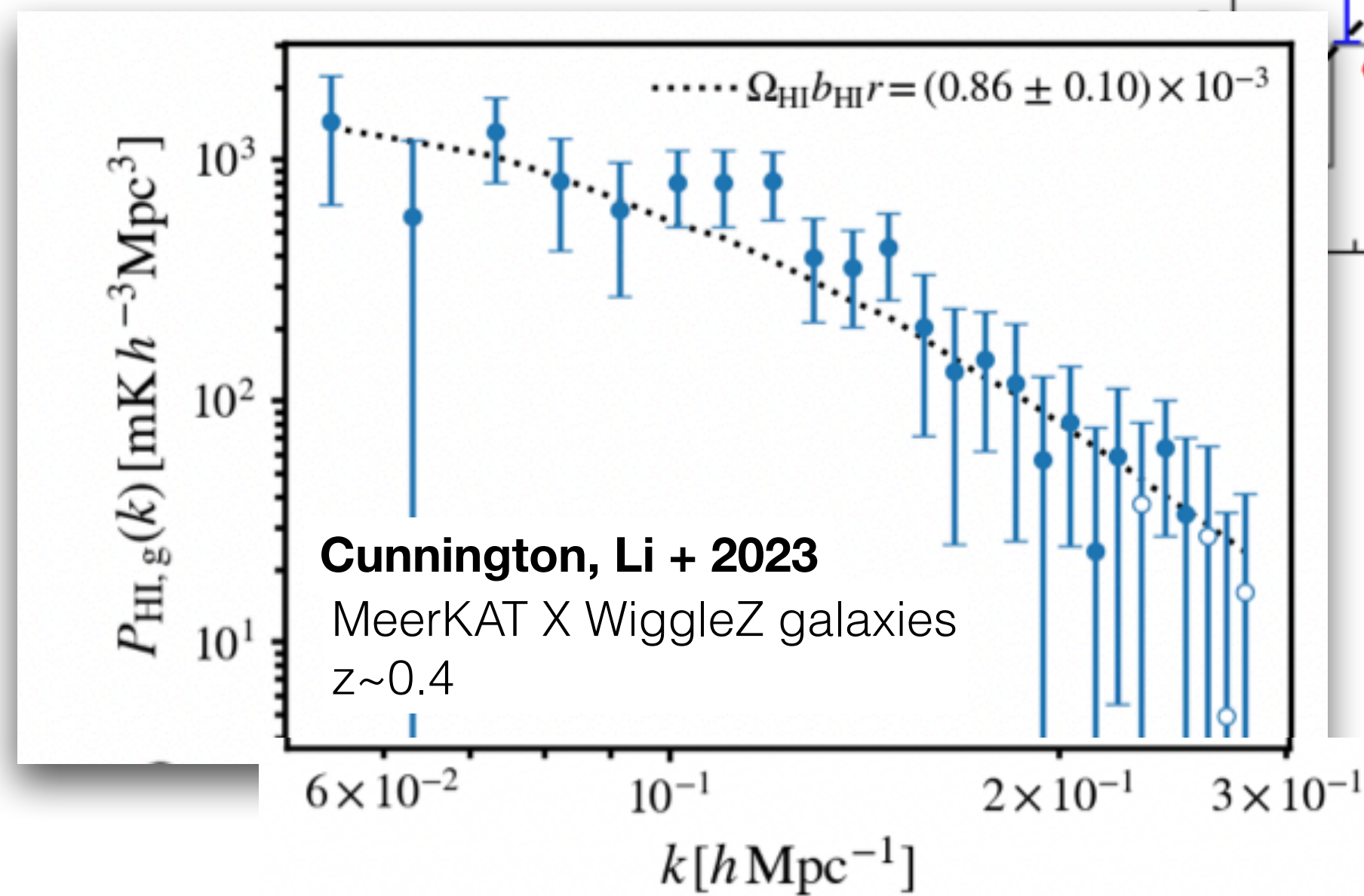


HI intensity mapping

State-of-the-art

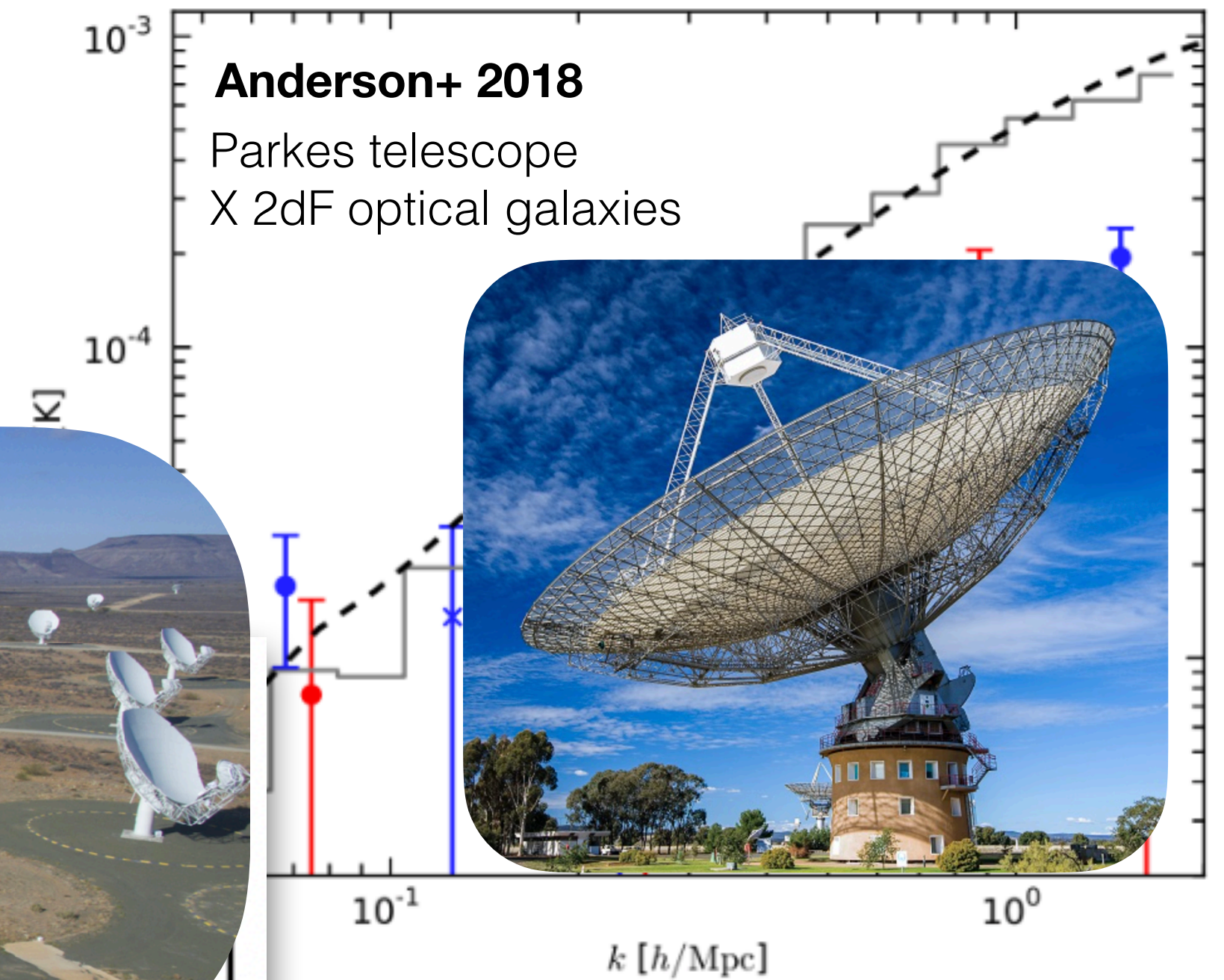
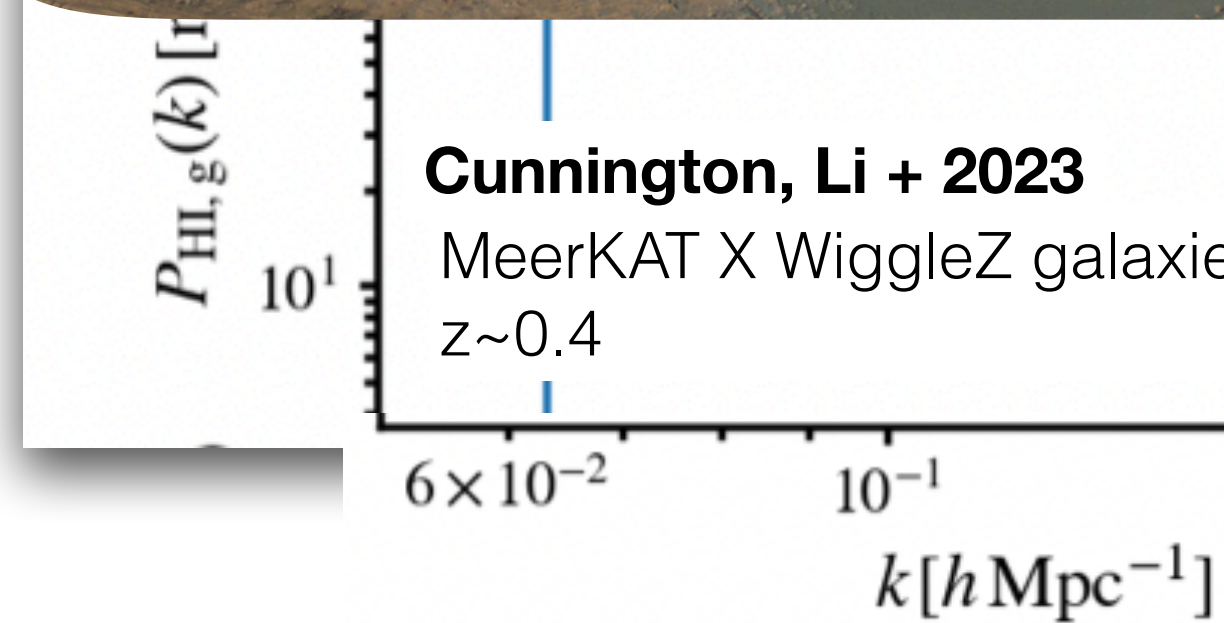
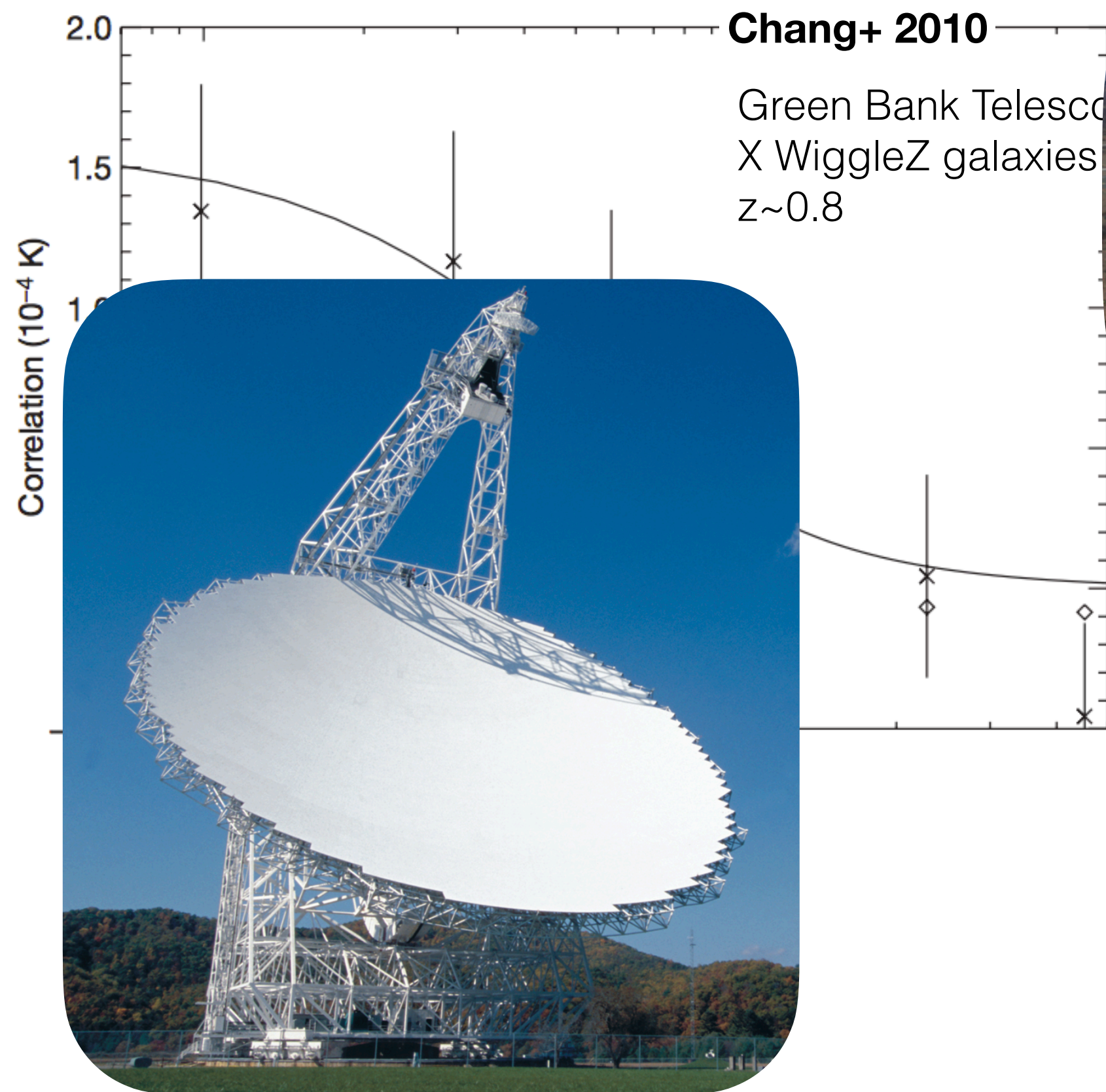


also Masui+ 2013, Switzer+ 2013,
Wolz+ 2017,2022



HI intensity mapping

State-of-the-art

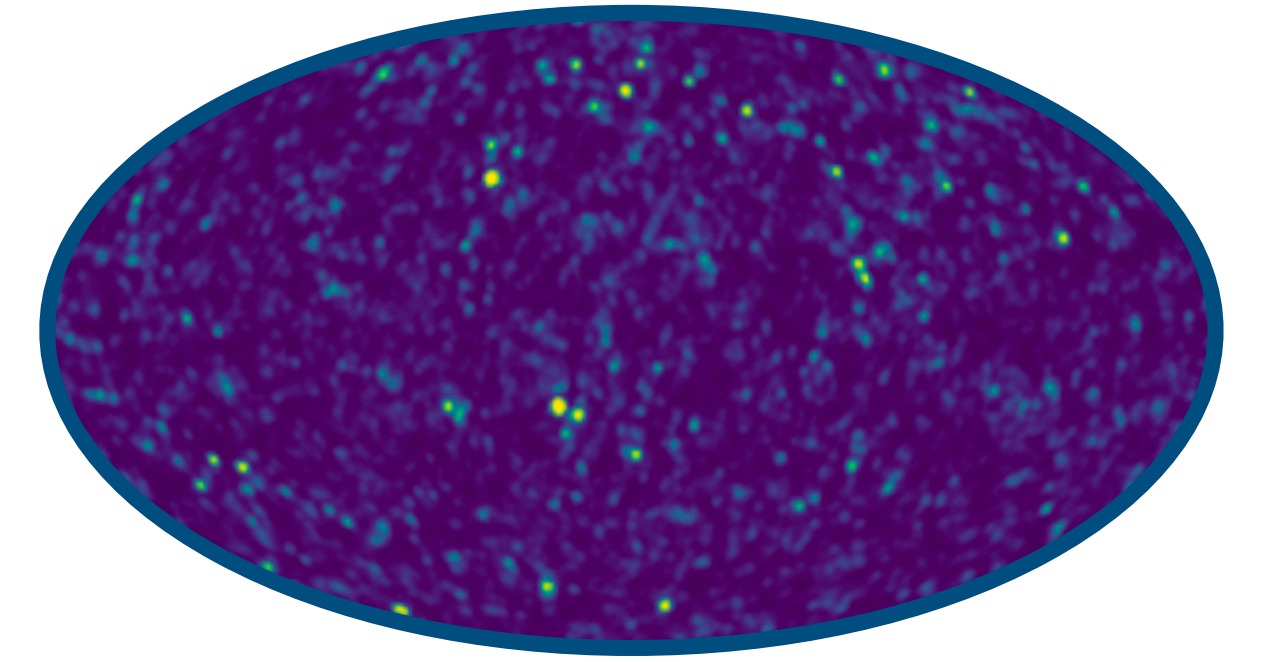


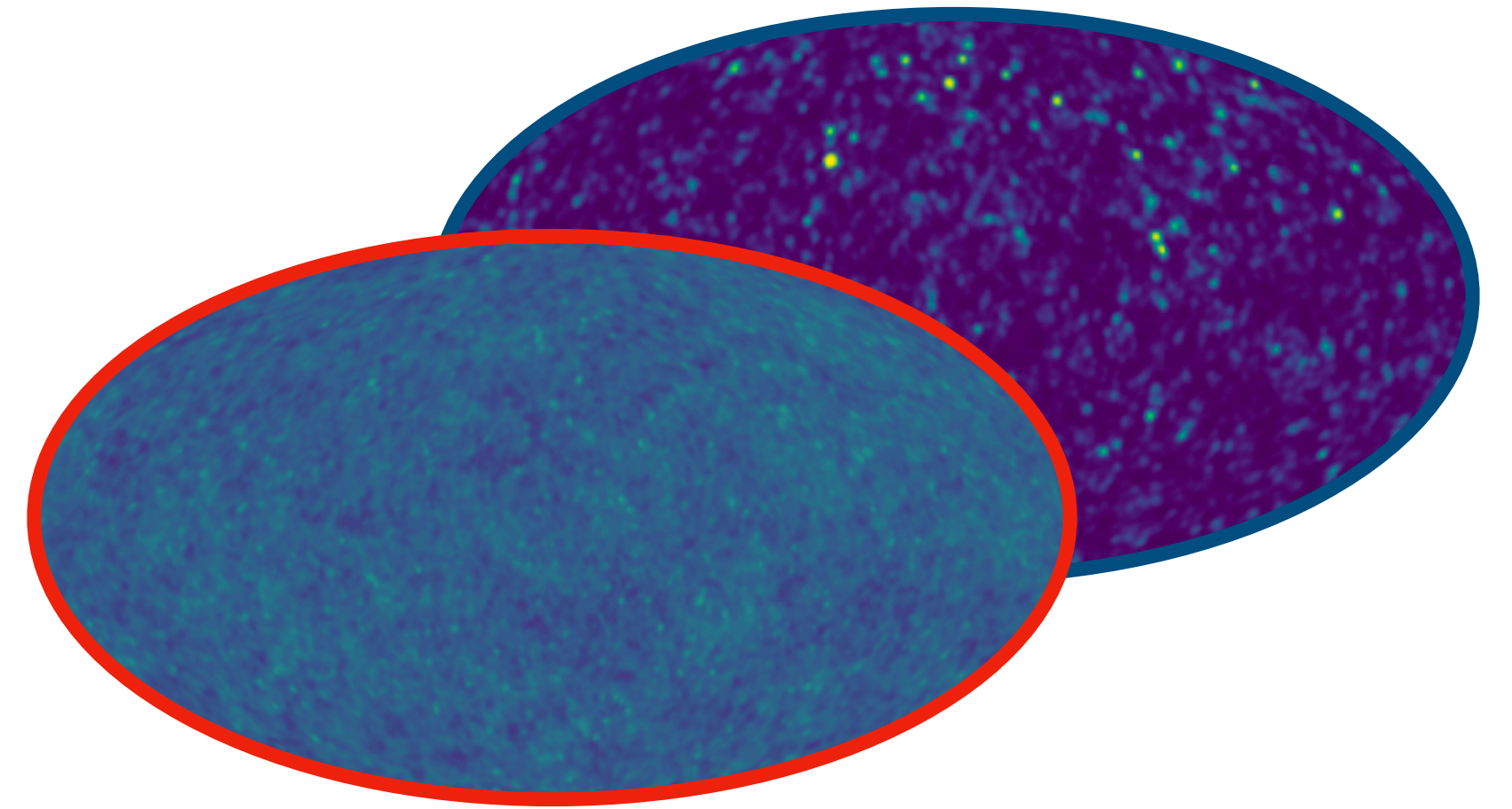
CHIME collaboration, 2022
stacking LRGs, ELG and QSOs from eBOSS
 $0.8 < z < 1.5$

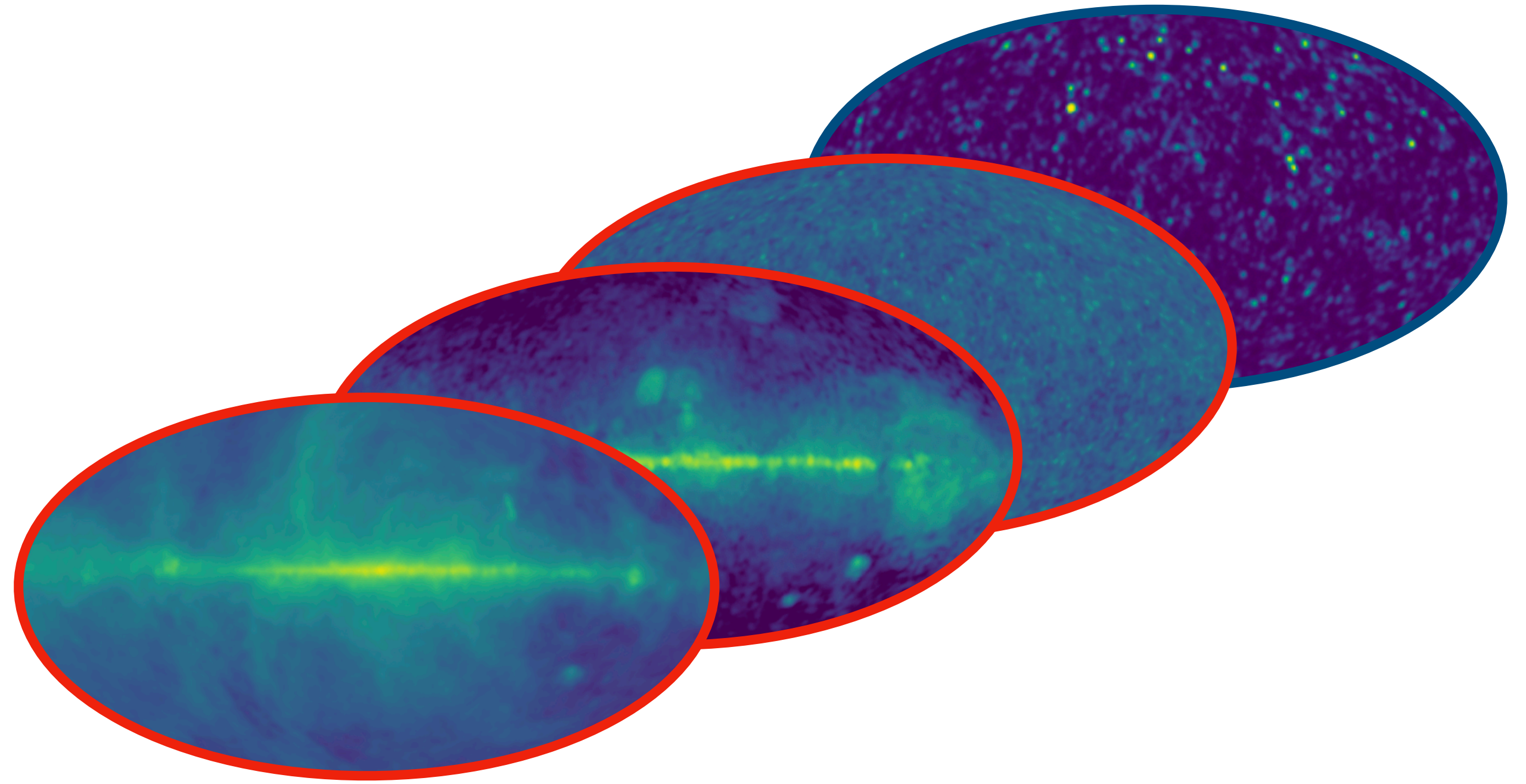
1. **Hydrogen Intensity Mapping (IM):**
what is it and why to do it

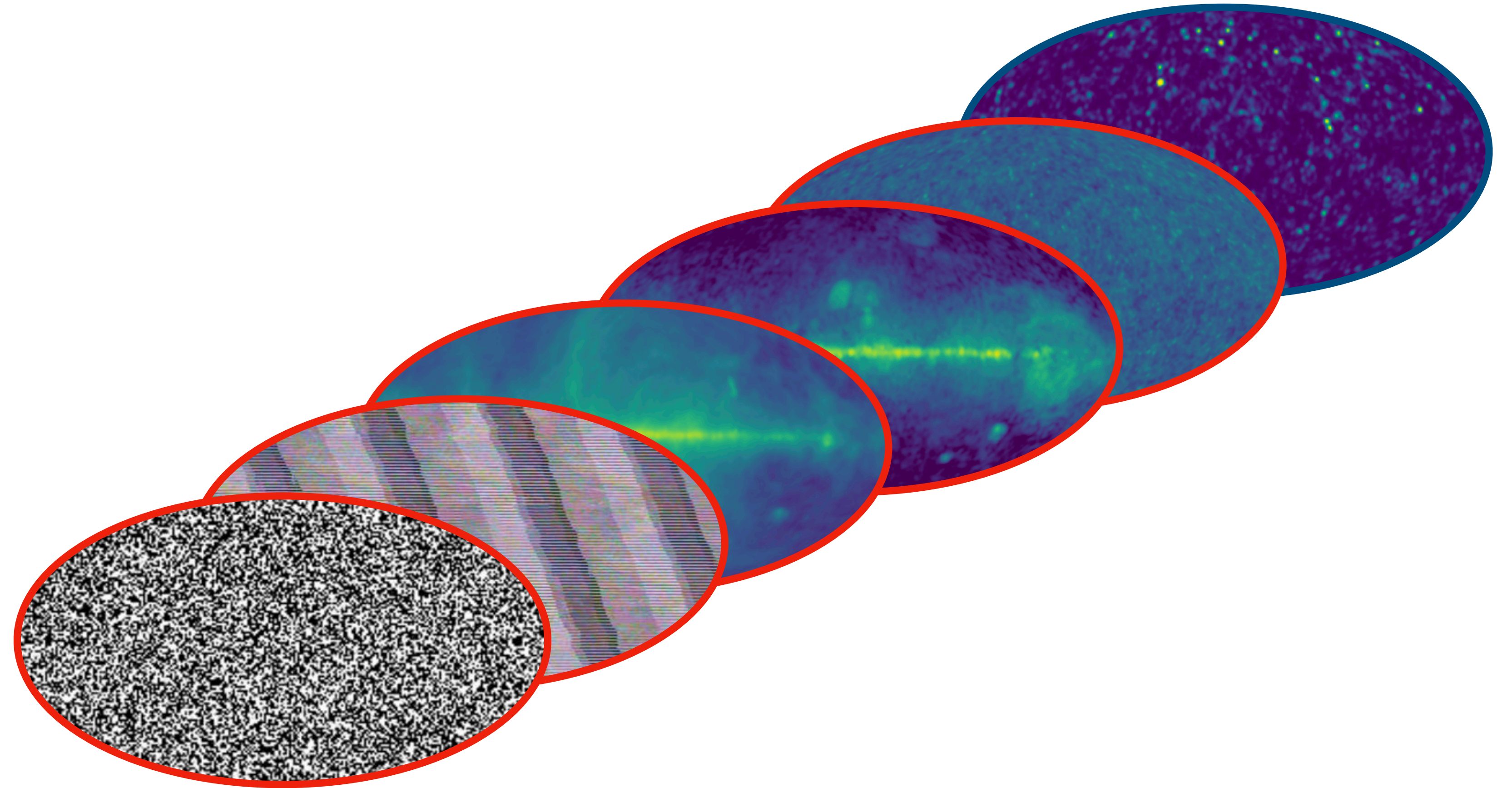
2. **IM is hard...**
(Biggest challenge: weakness of the IM
signal compared to contaminants)

3. **...yet we are getting there!**
(MeerKAT IM observations, analysis ongoing)



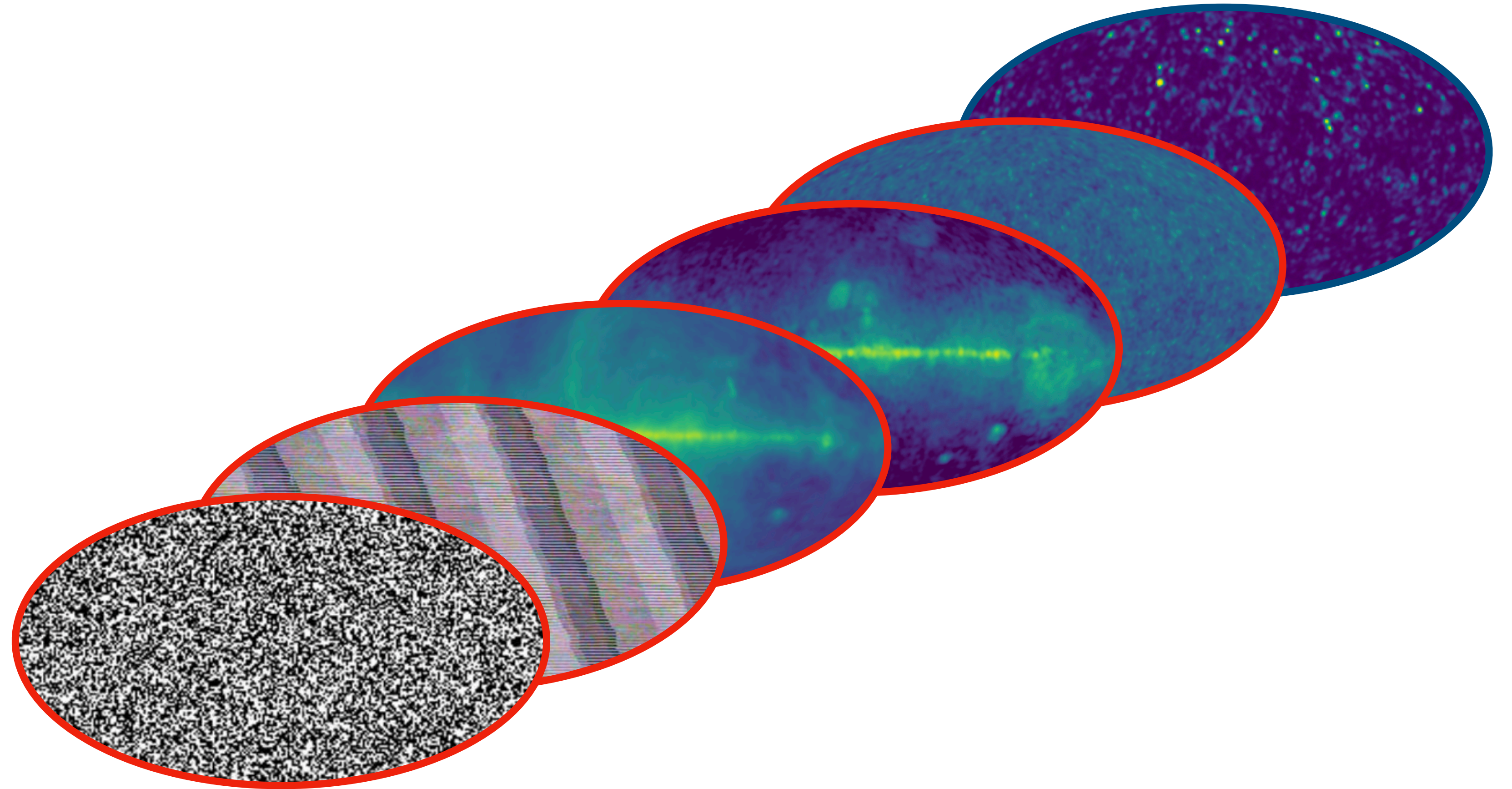








Handwritten blue scribbles

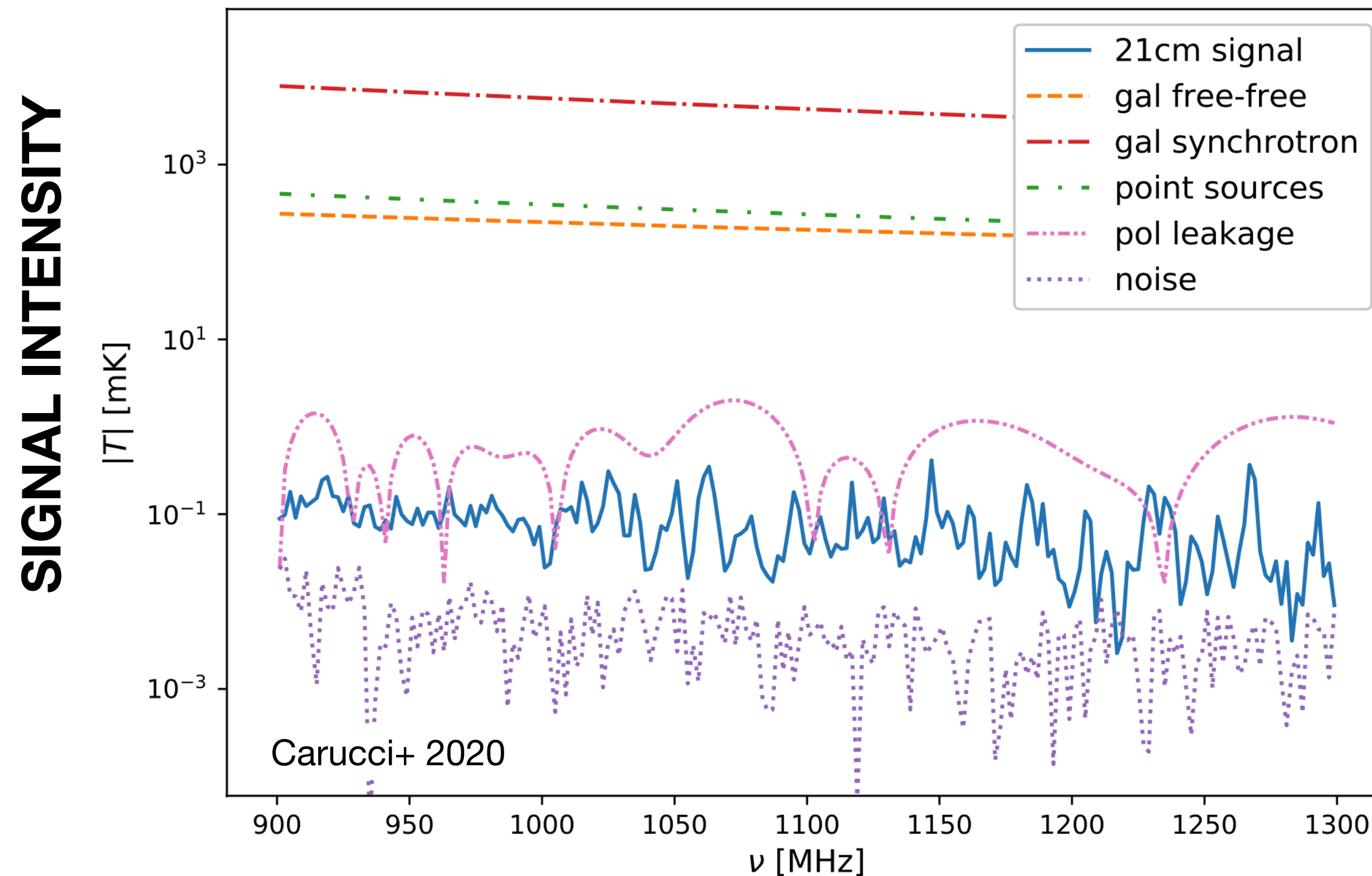


HI intensity mapping

buried under the contaminants

CHALLENGES:

- Foregrounds
- Systematics

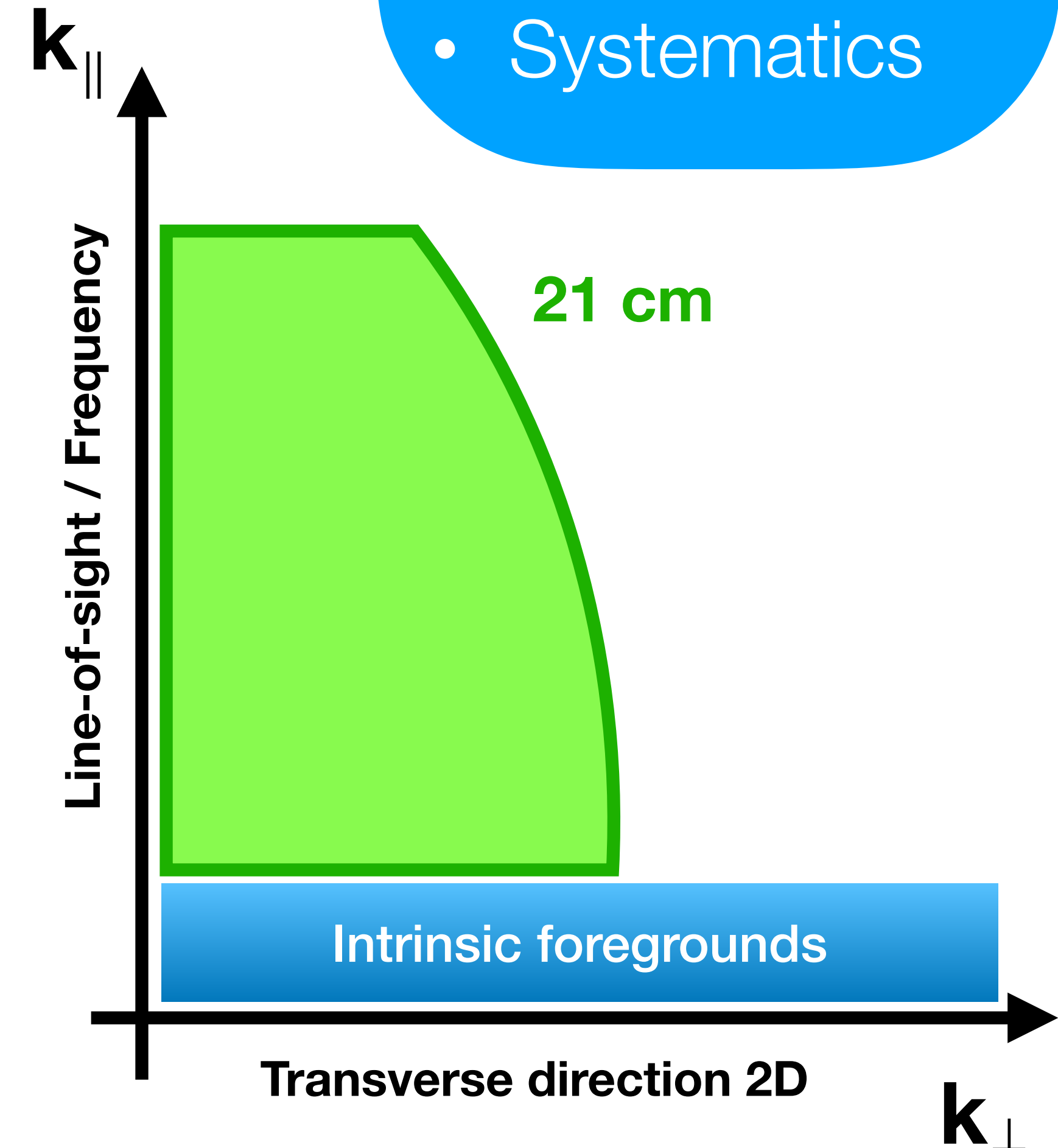
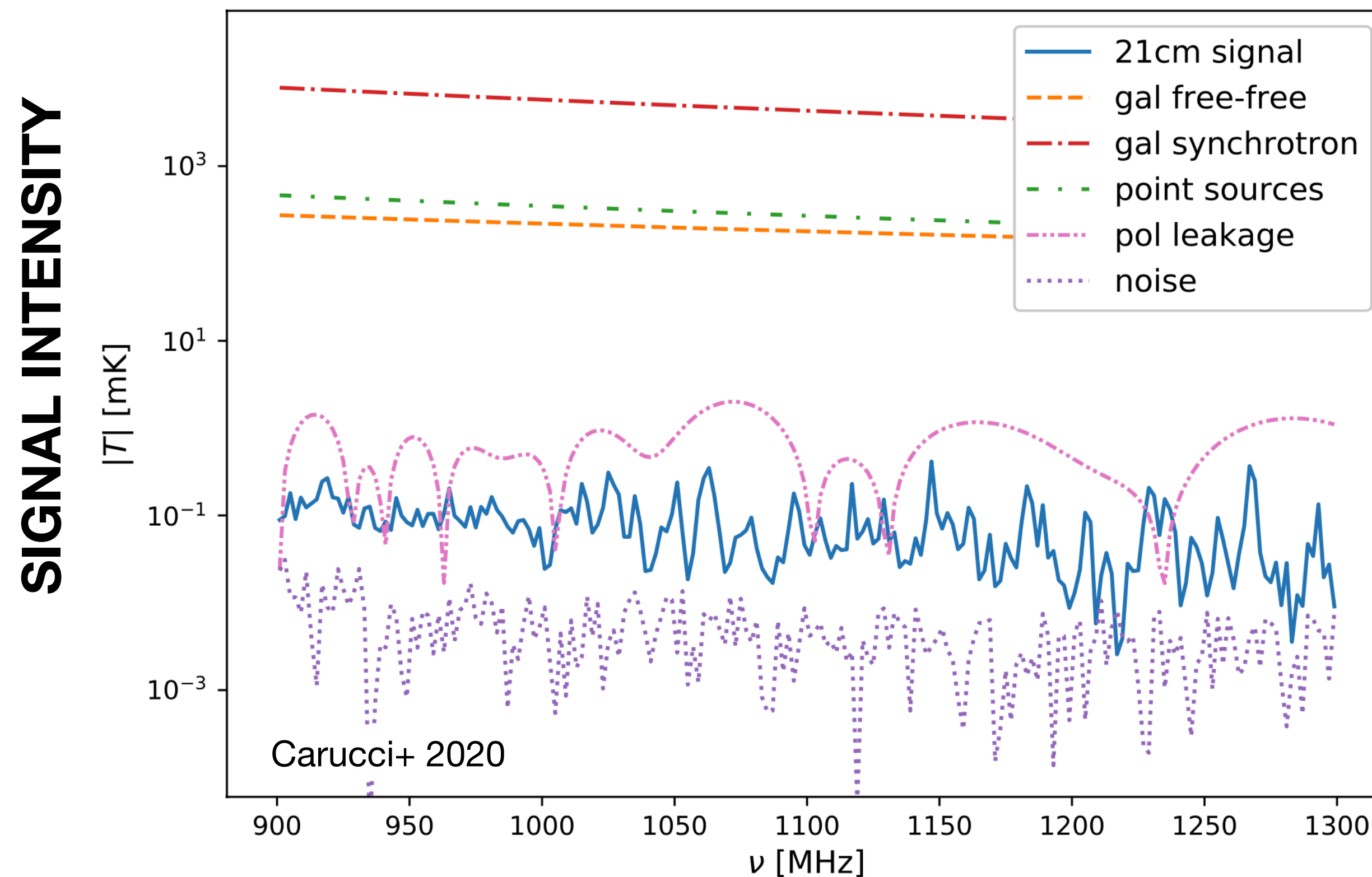


HI intensity mapping

buried under the contaminants

CHALLENGES:

- Foregrounds
- Systematics

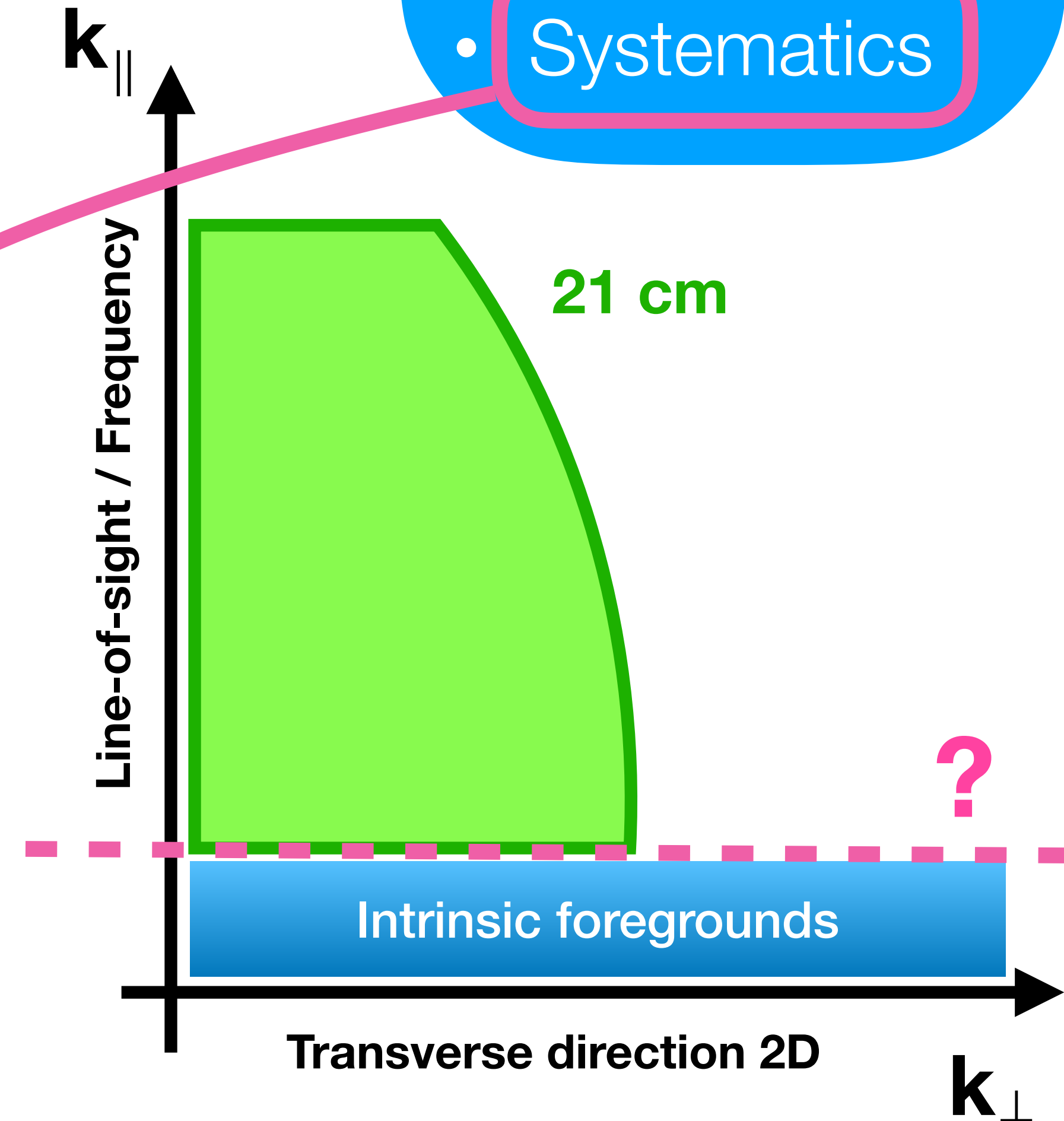
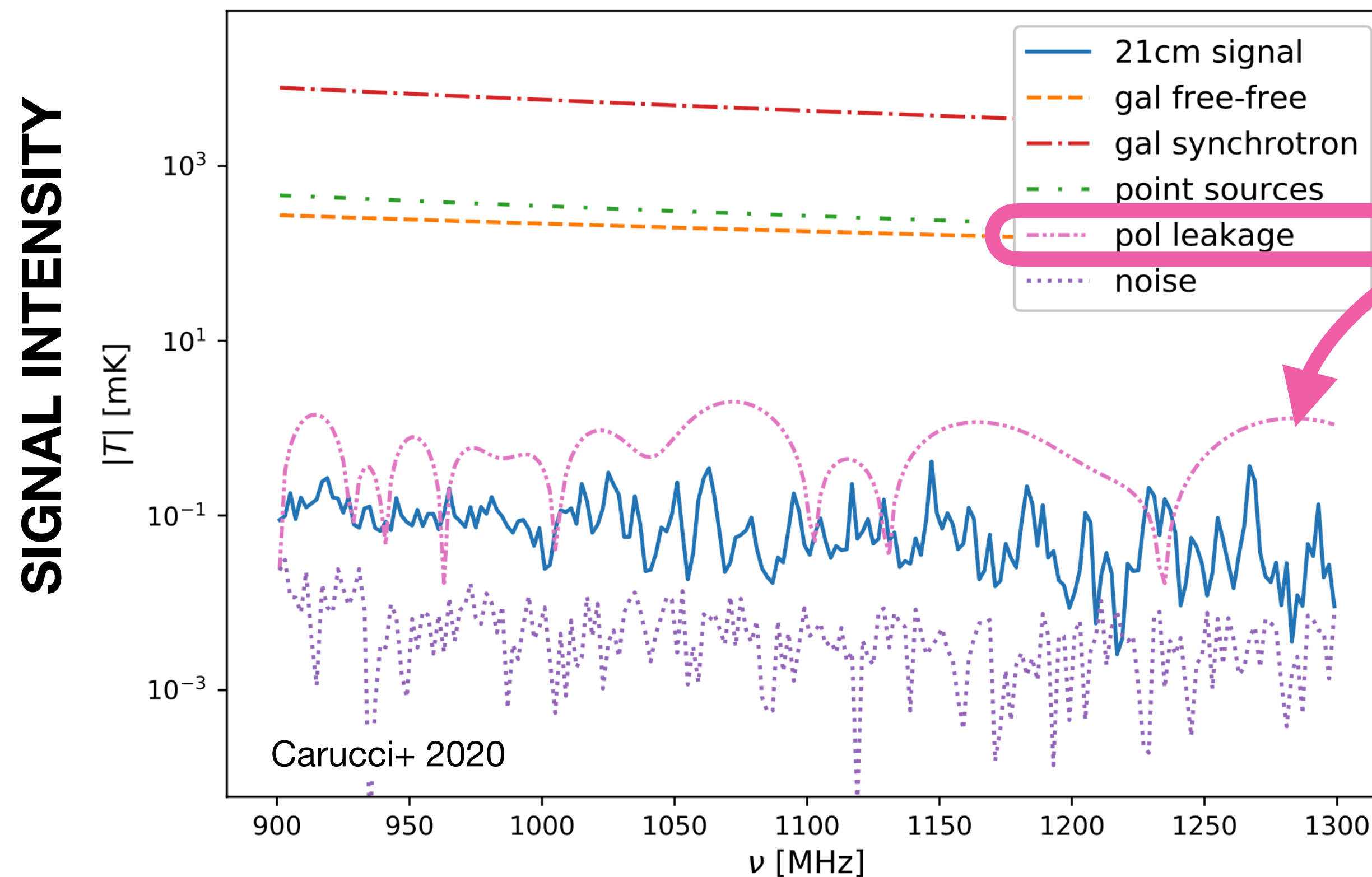


HI intensity mapping

buried under the contaminants

CHALLENGES:

- Foregrounds
- Systematics



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.

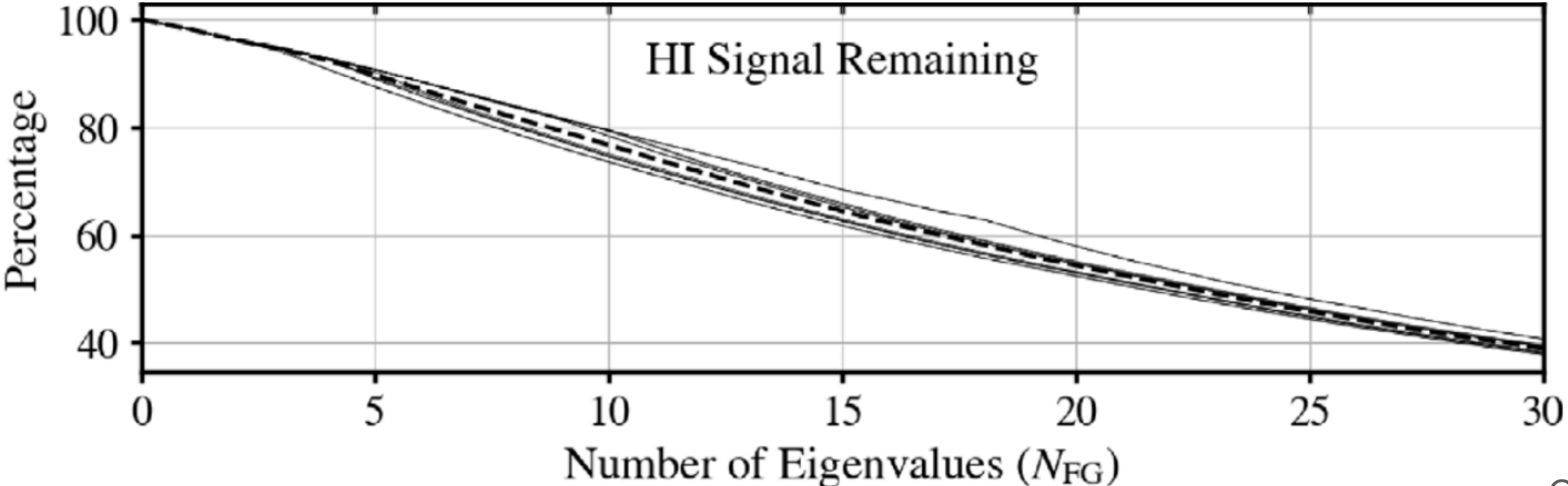
$$\begin{array}{c} \mathbf{X} \\ \text{signal} \\ (f,p) \end{array} = \begin{array}{c} \text{mixing} \\ \text{matrix } (f,n) \\ \mathbf{A} \end{array} \begin{array}{c} \mathbf{S} \\ \text{sources} \\ (n,p) \end{array} + \begin{array}{c} \mathbf{N} \\ \text{HI signal!} \end{array}$$

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.

Need to set number n of sources!

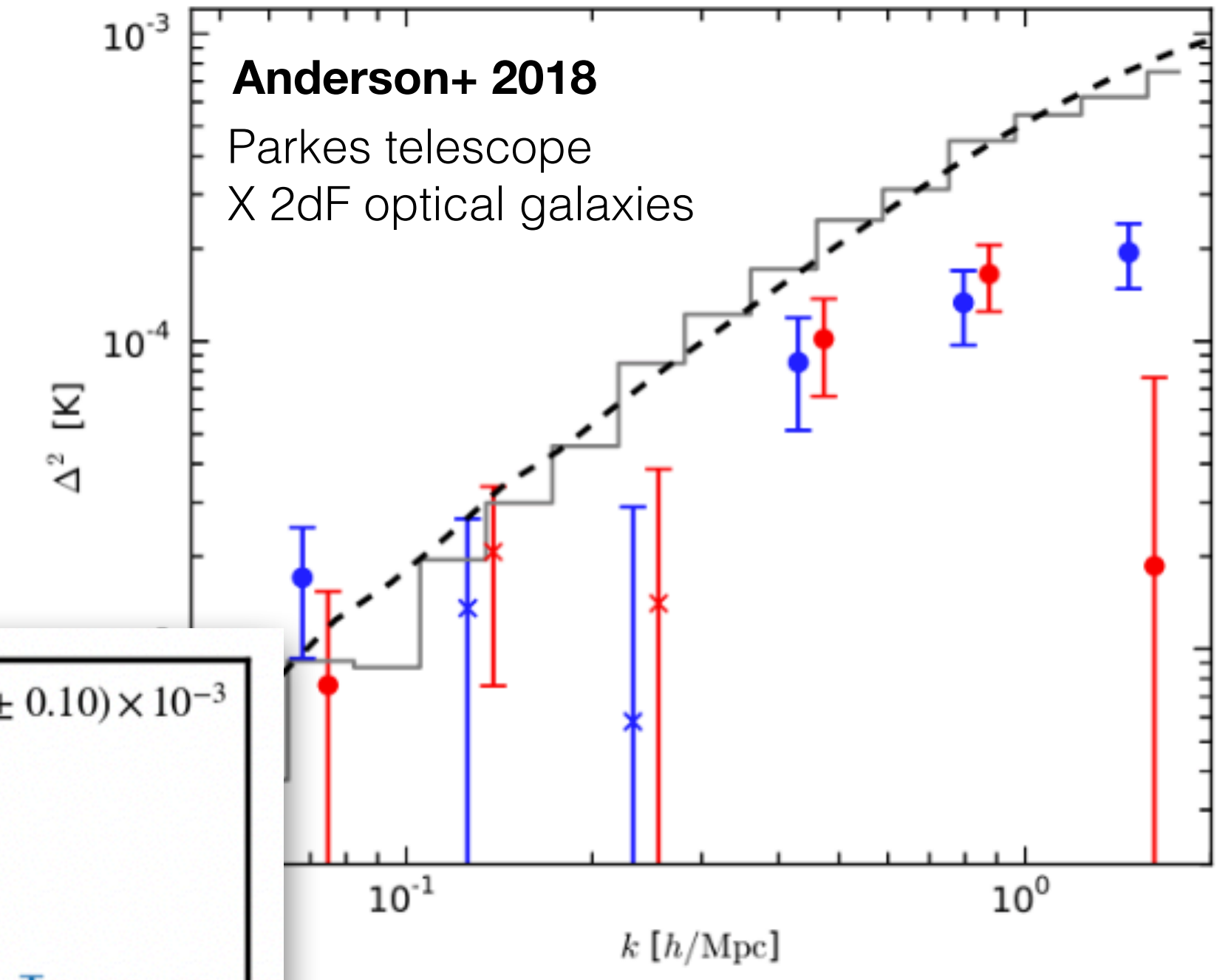
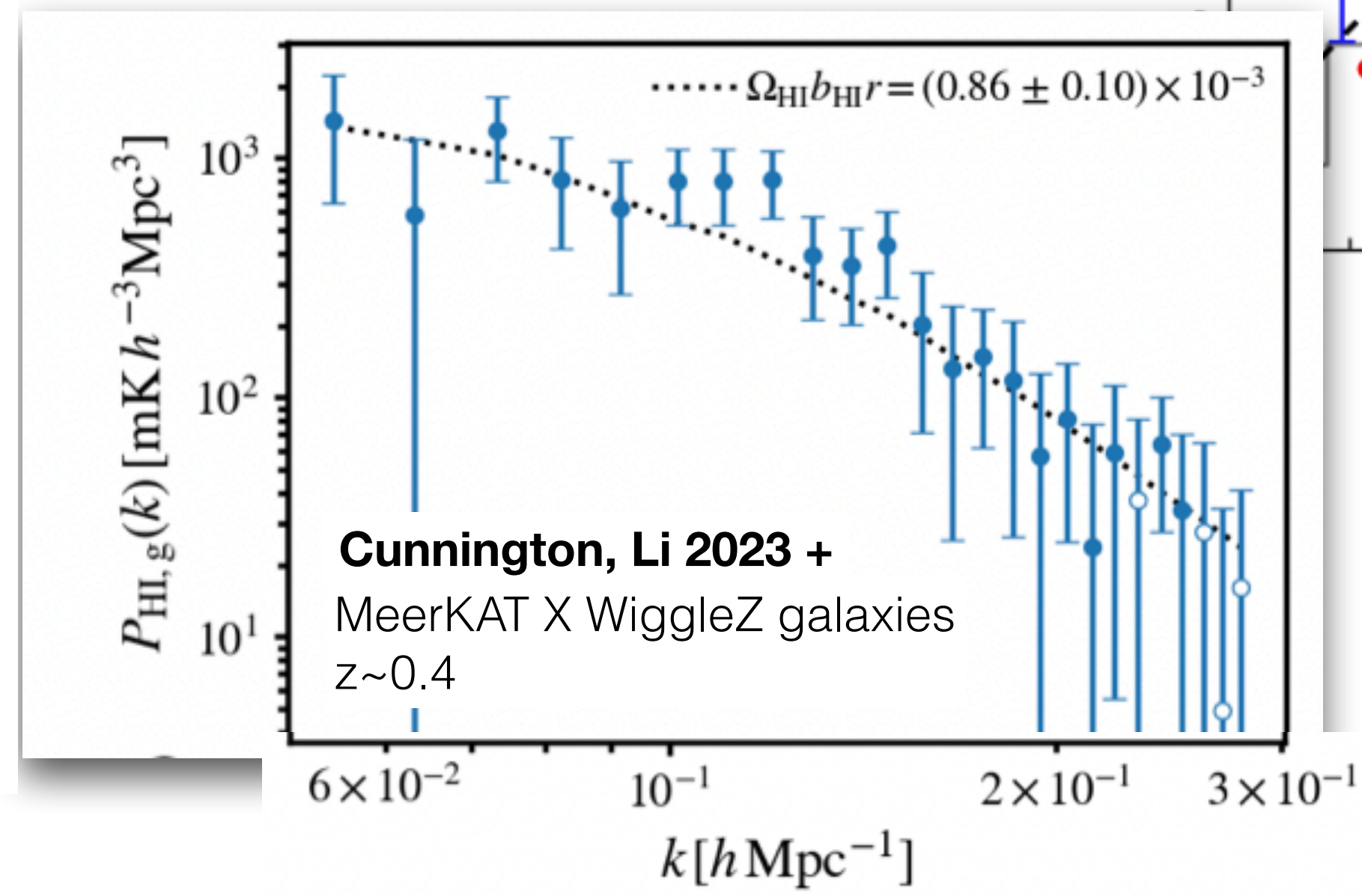
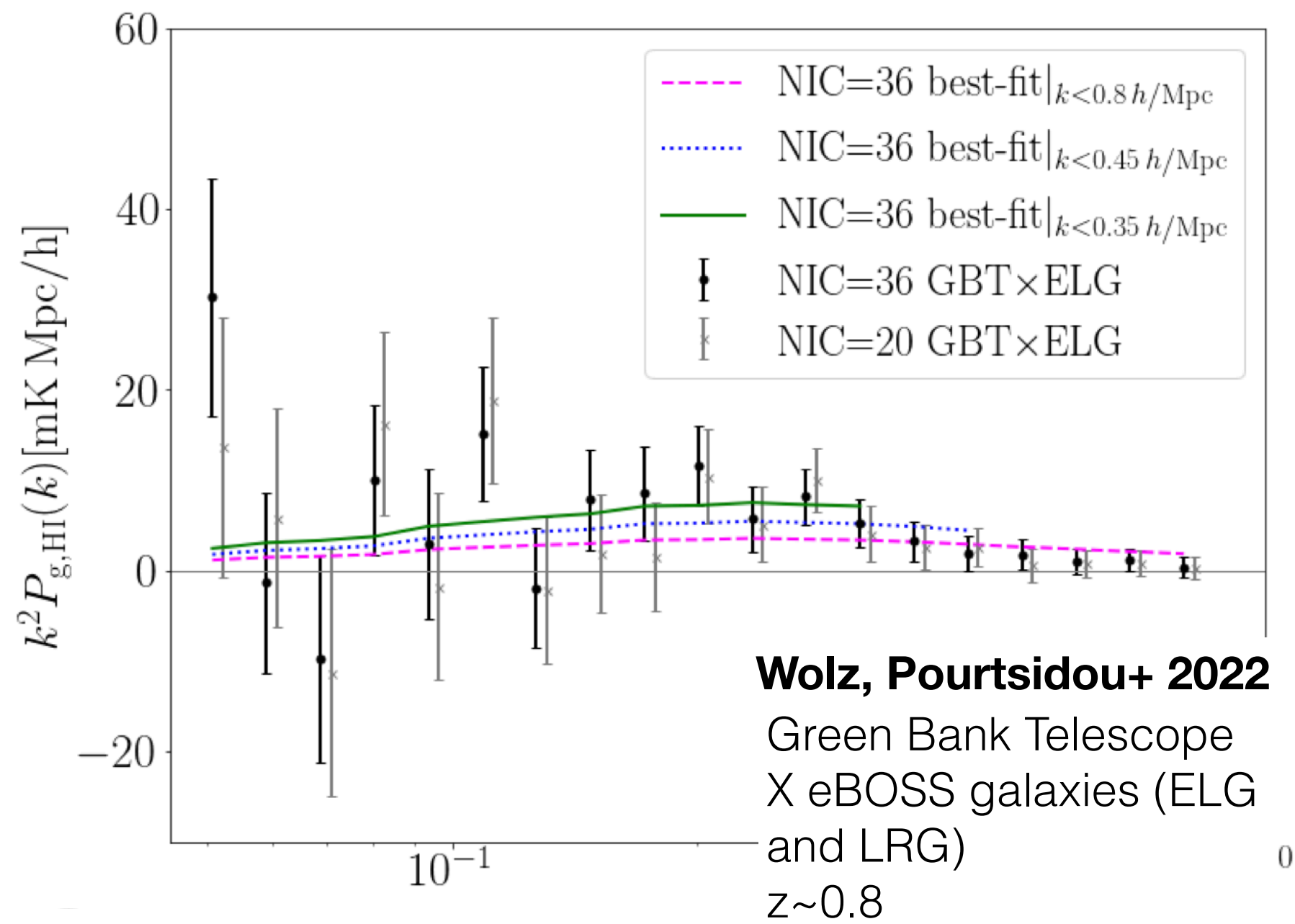
$$\begin{array}{c}
 \text{signal} \\
 \mathbf{X} \\
 (f,p)
 \end{array}
 =
 \begin{array}{c}
 \text{mixing} \\
 \text{matrix} \\
 \mathbf{A} \\
 (f,n)
 \end{array}
 \begin{array}{c}
 \mathbf{S} \\
 \text{sources} \\
 (n,p)
 \end{array}
 +
 \begin{array}{c}
 \mathbf{N} \\
 \text{HI signal!}
 \end{array}$$



Cunnington+ 2021

Blind Source Separation algorithms

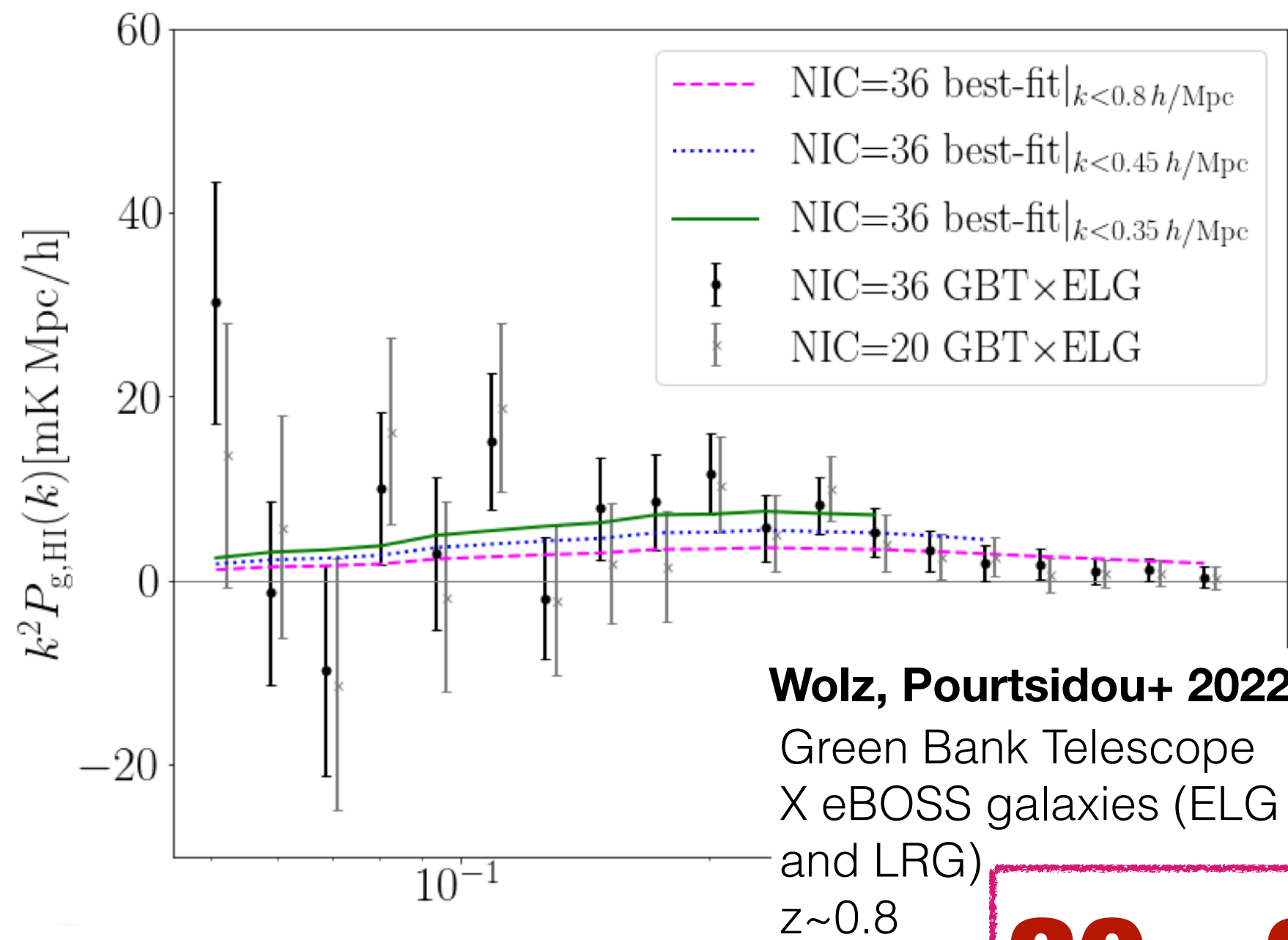
State-of-the-art



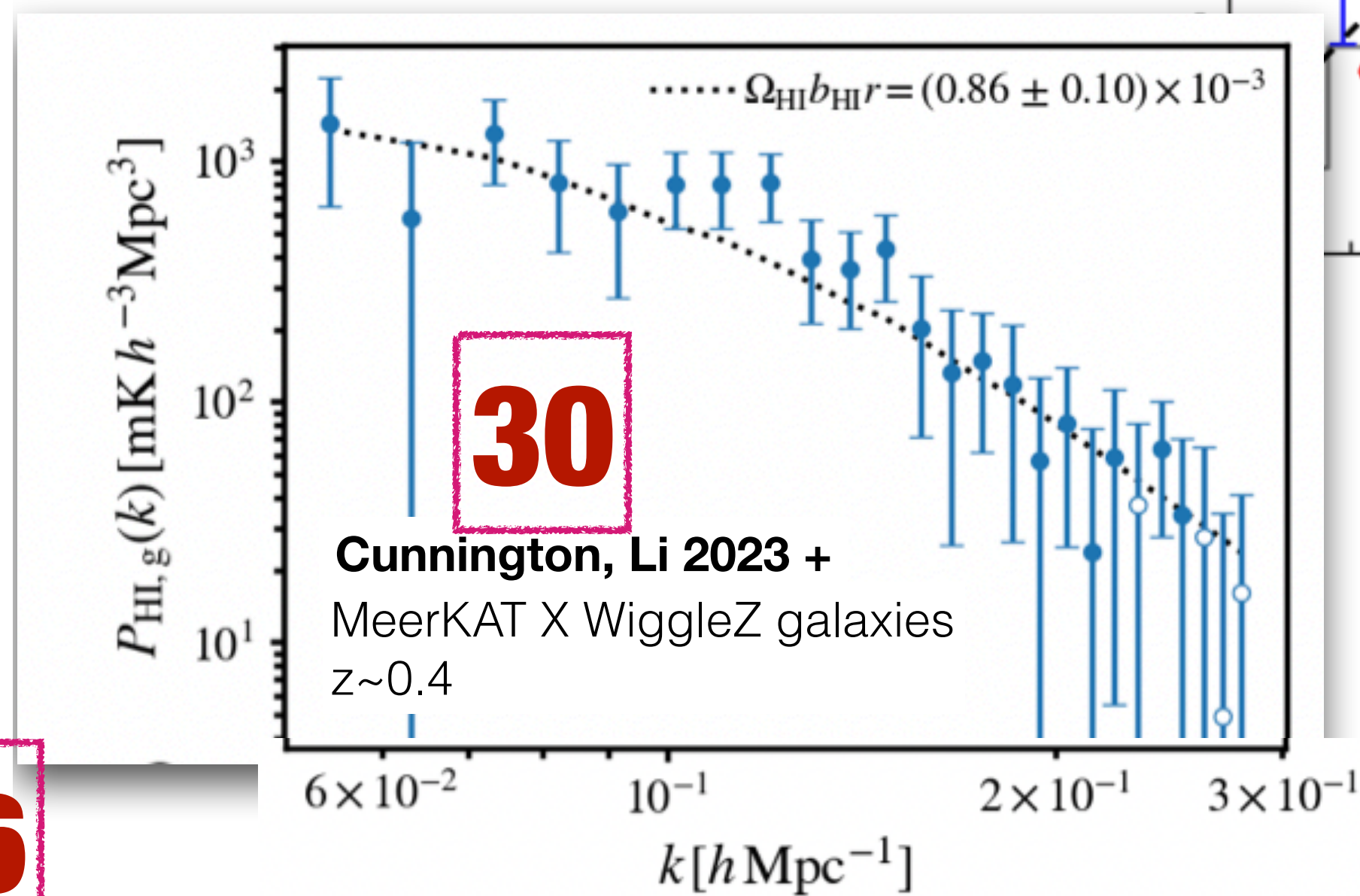
See also Masui+ 2013, Switzer+ 2013, Wolz+ 2017

Blind Source Separation algorithms

State-of-the-art

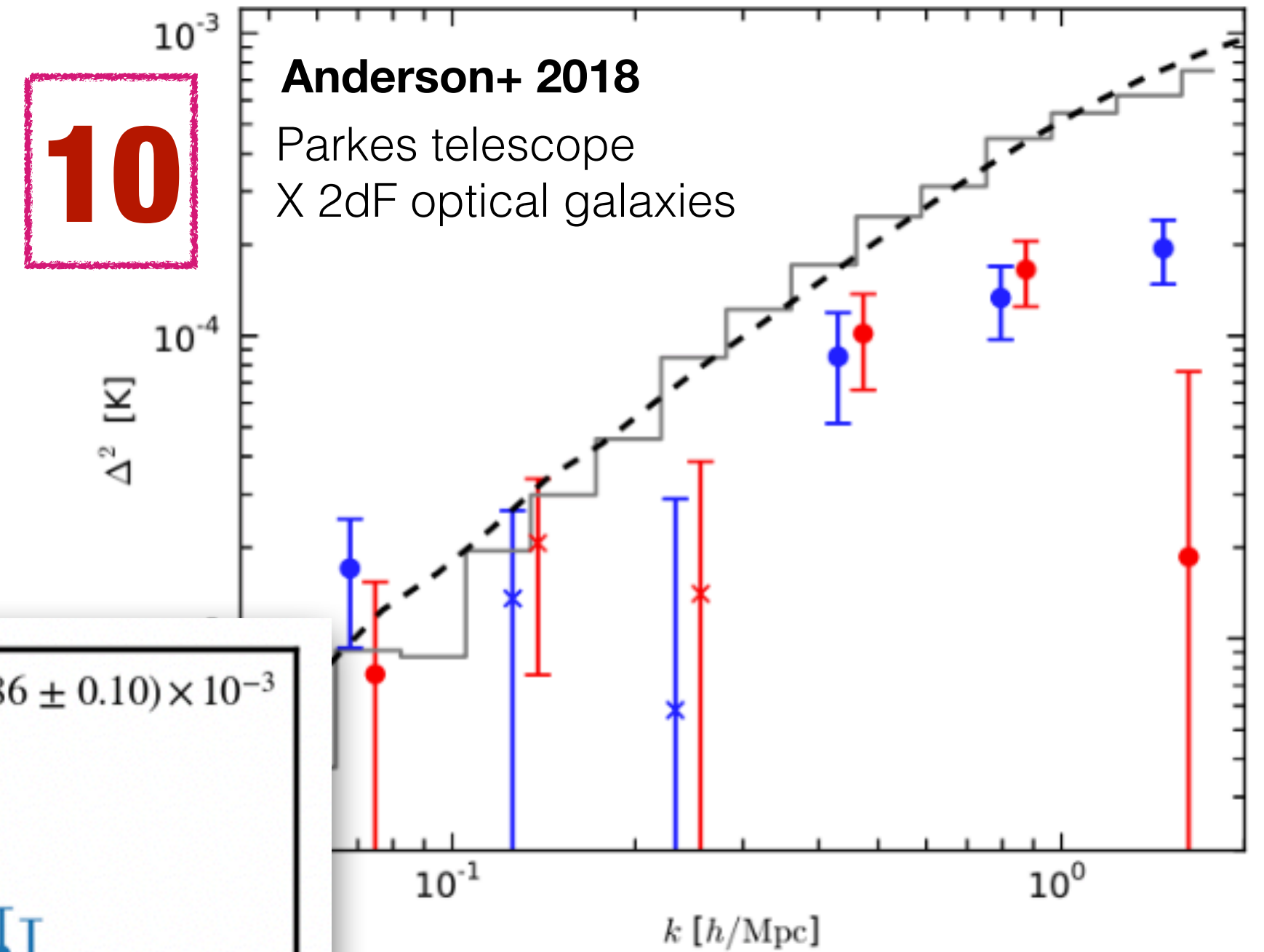


20 - 36



30

10

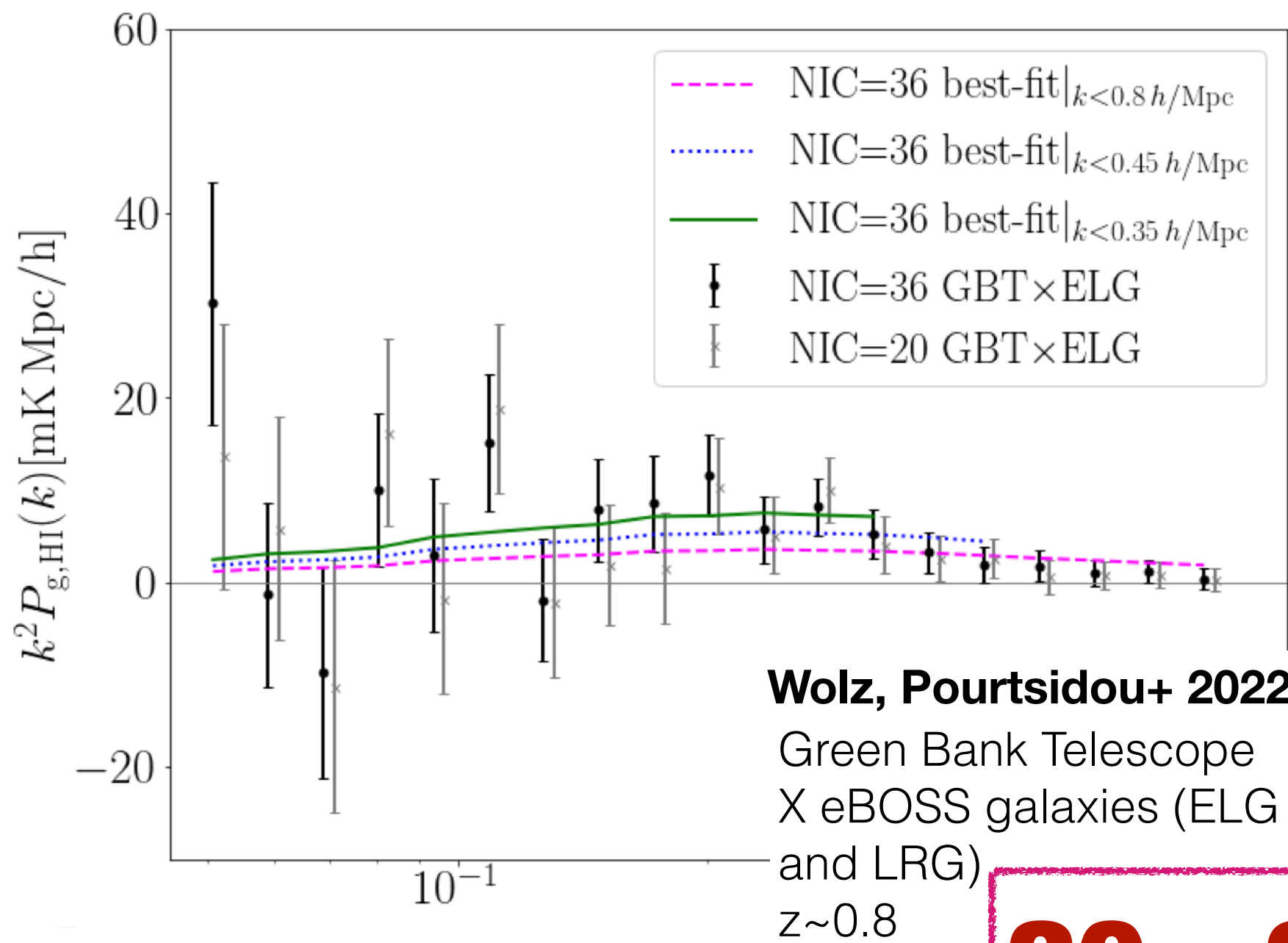


See also Masui+ 2013, Switzer+ 2013, Wolz+ 2017

10 - 20

Blind Source Separation algorithms

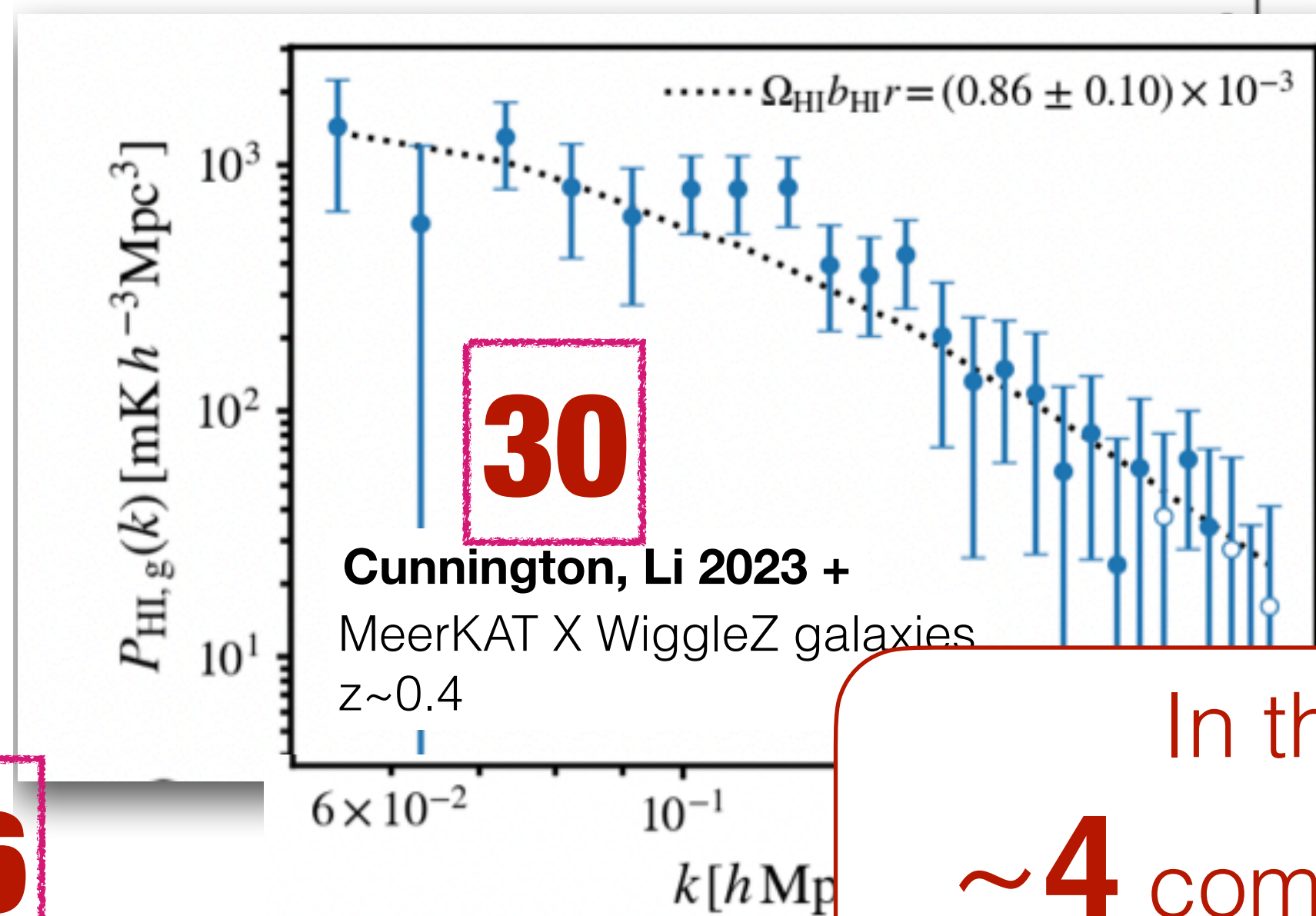
State-of-the-art



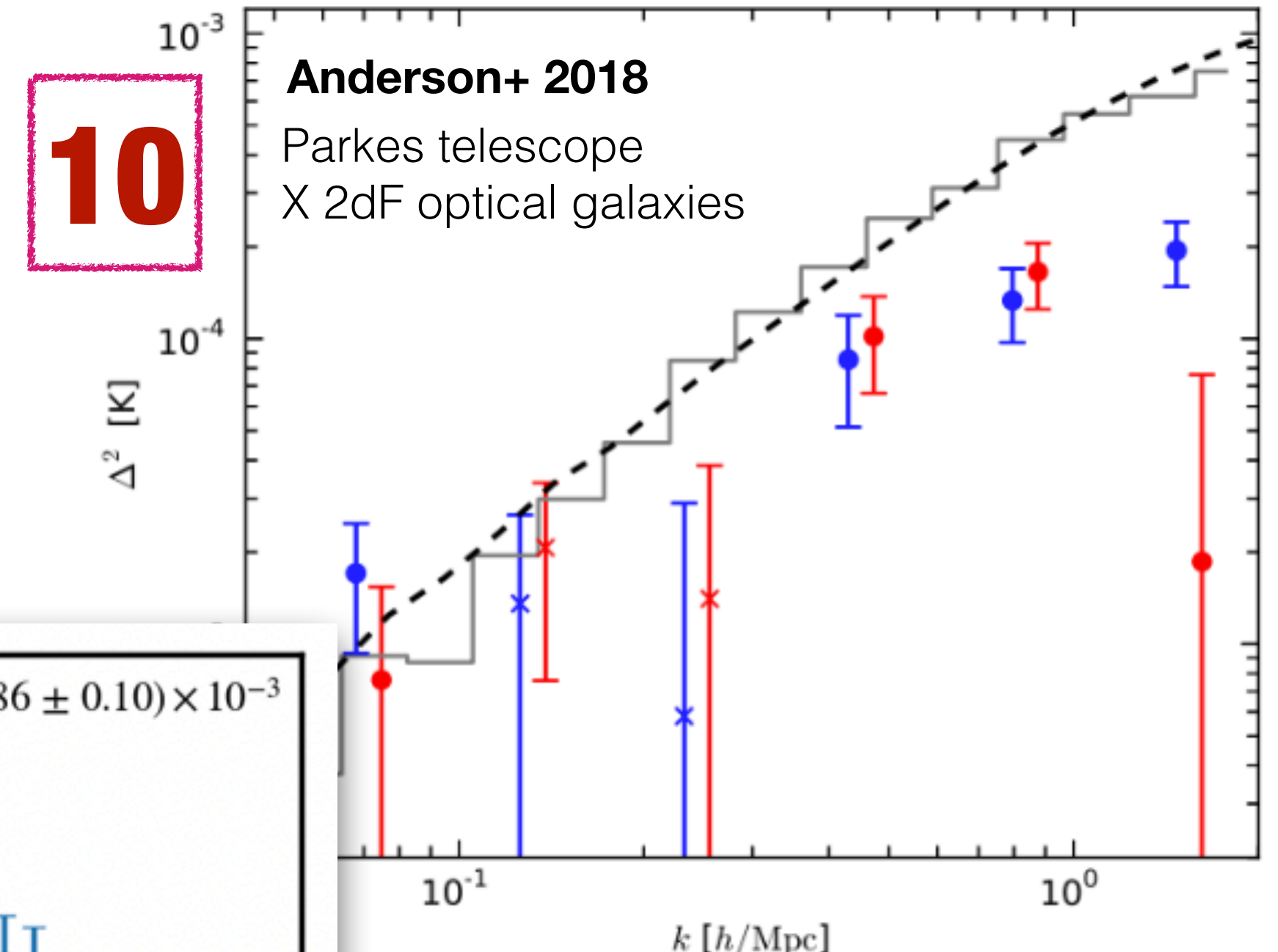
20 - 36

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

10 - 20

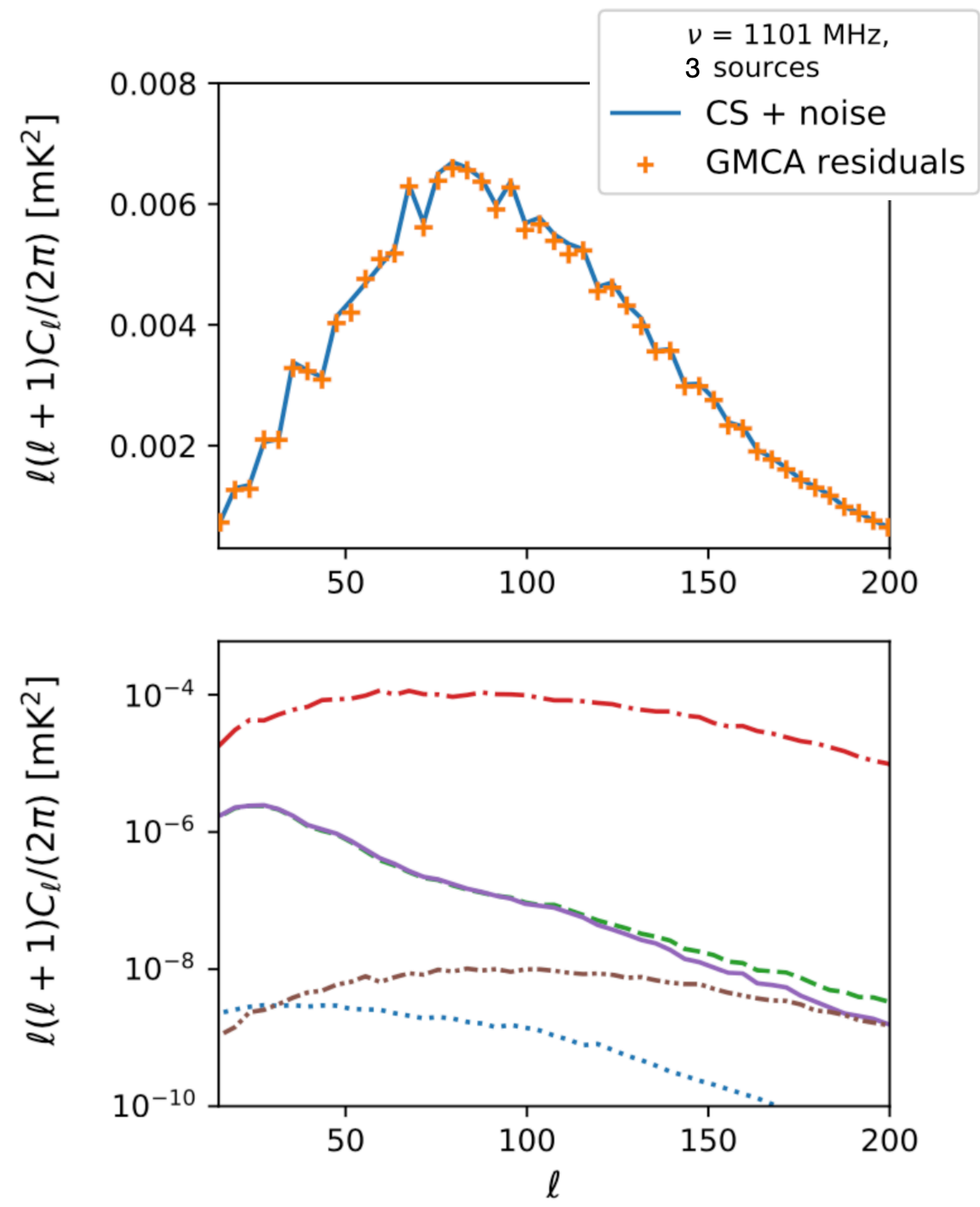


30

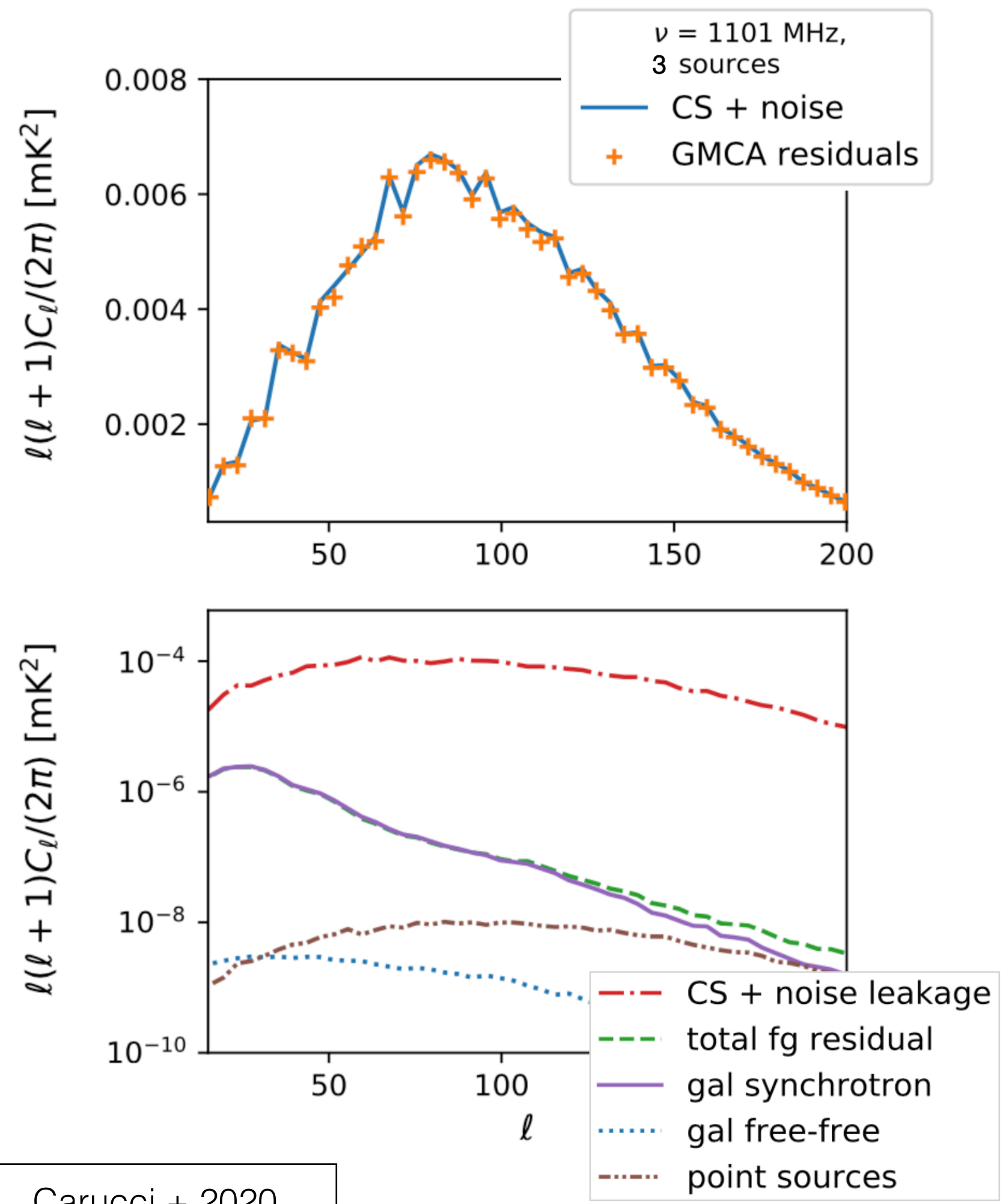


10

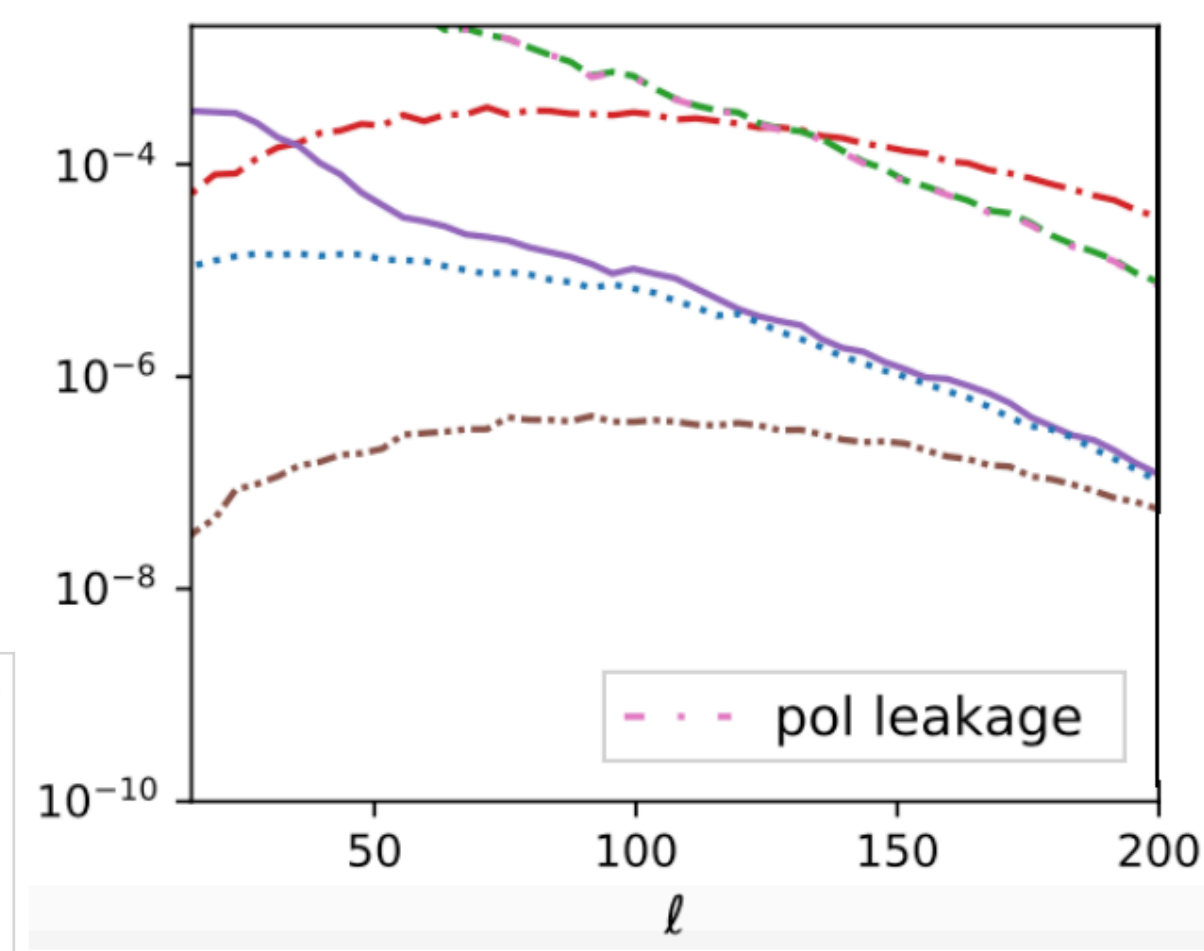
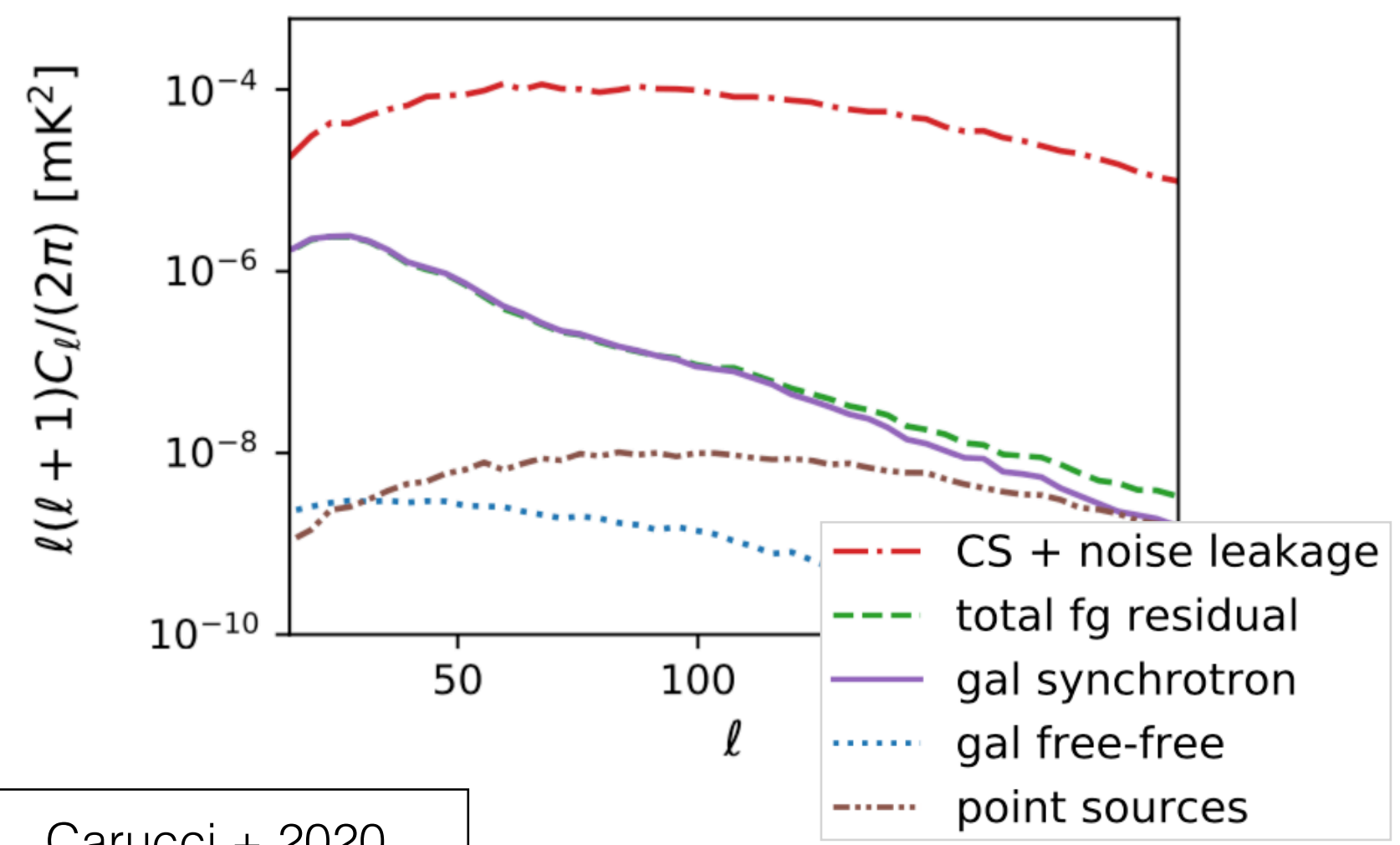
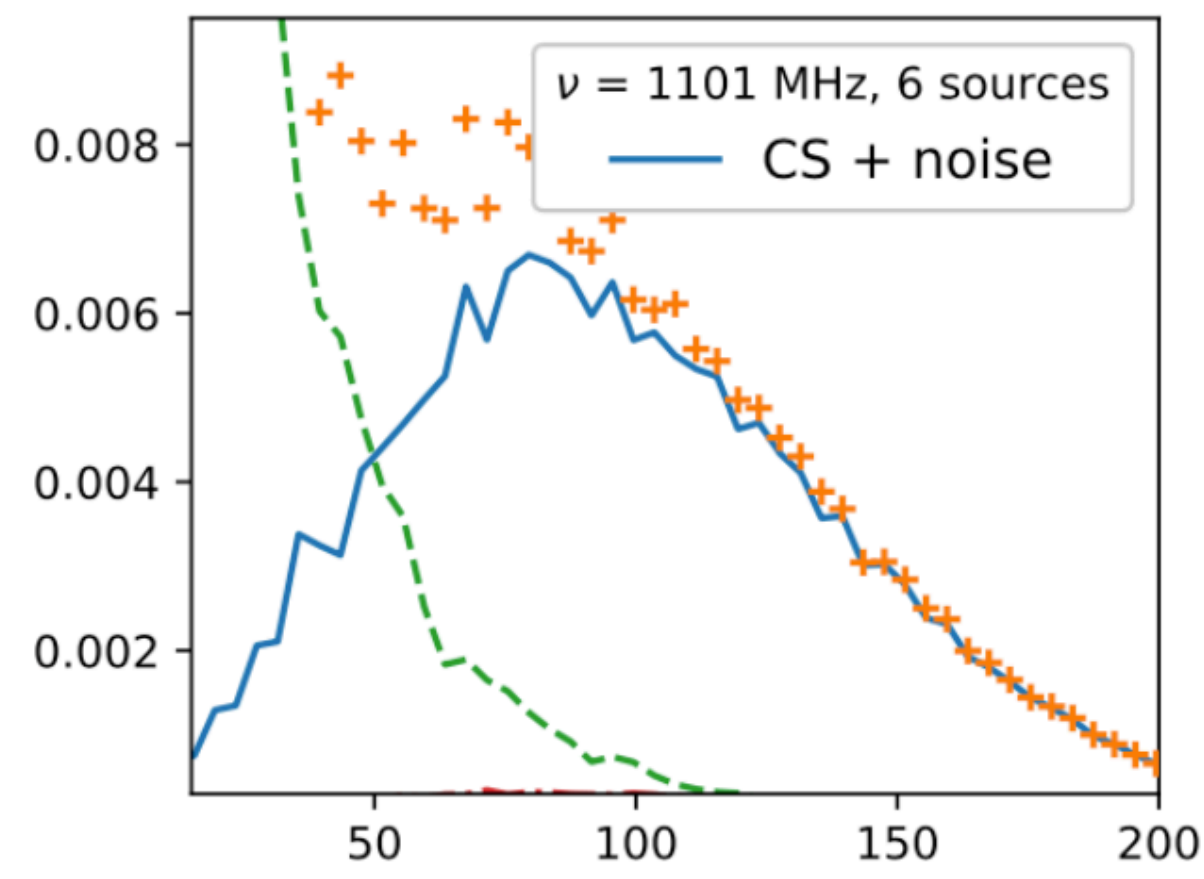
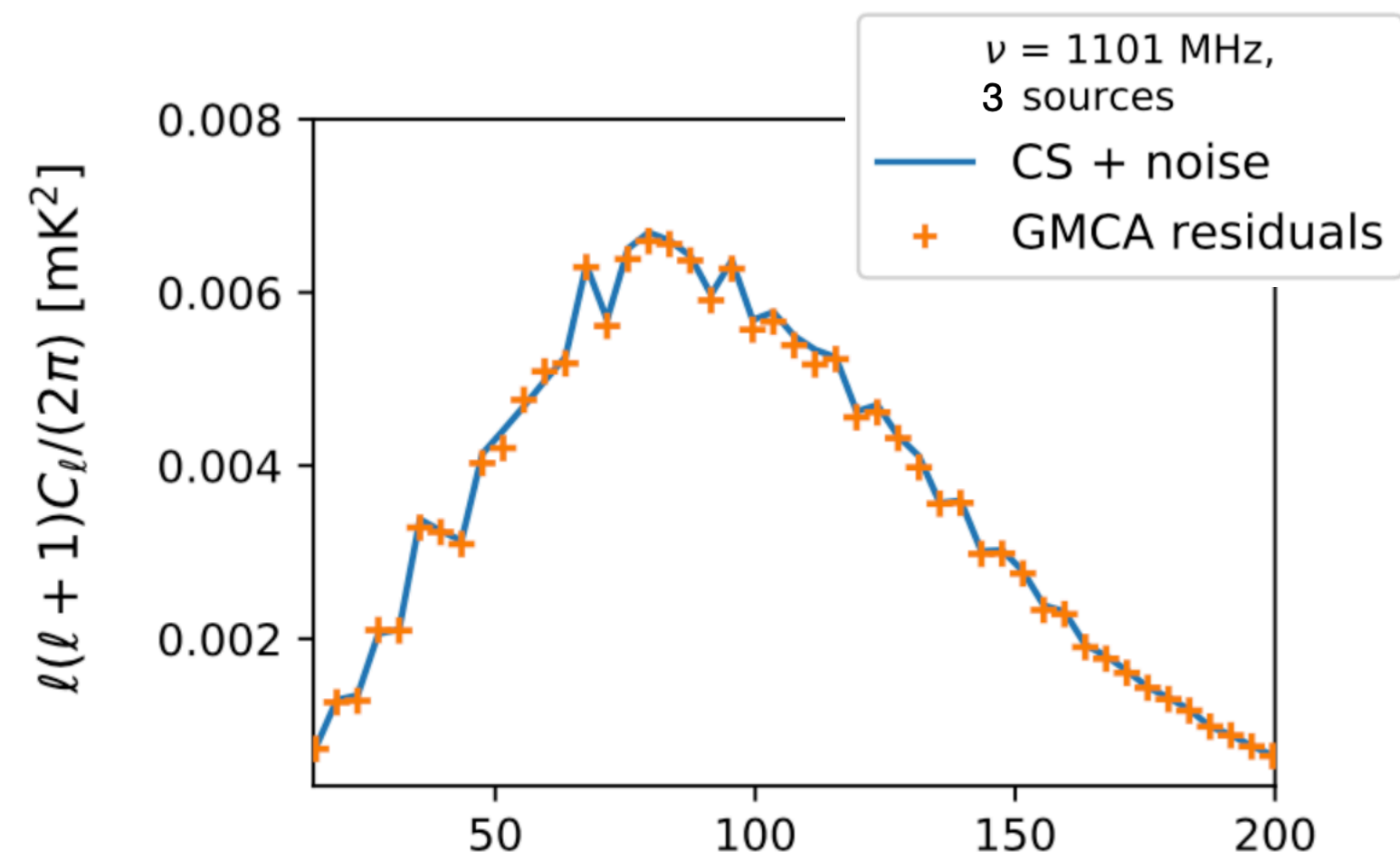
In theoretical works:
~4 components removed are
 enough
 (e.g., Wolz+ 2014, Alonso+ 2015, Cunnington+ 2019, ...)



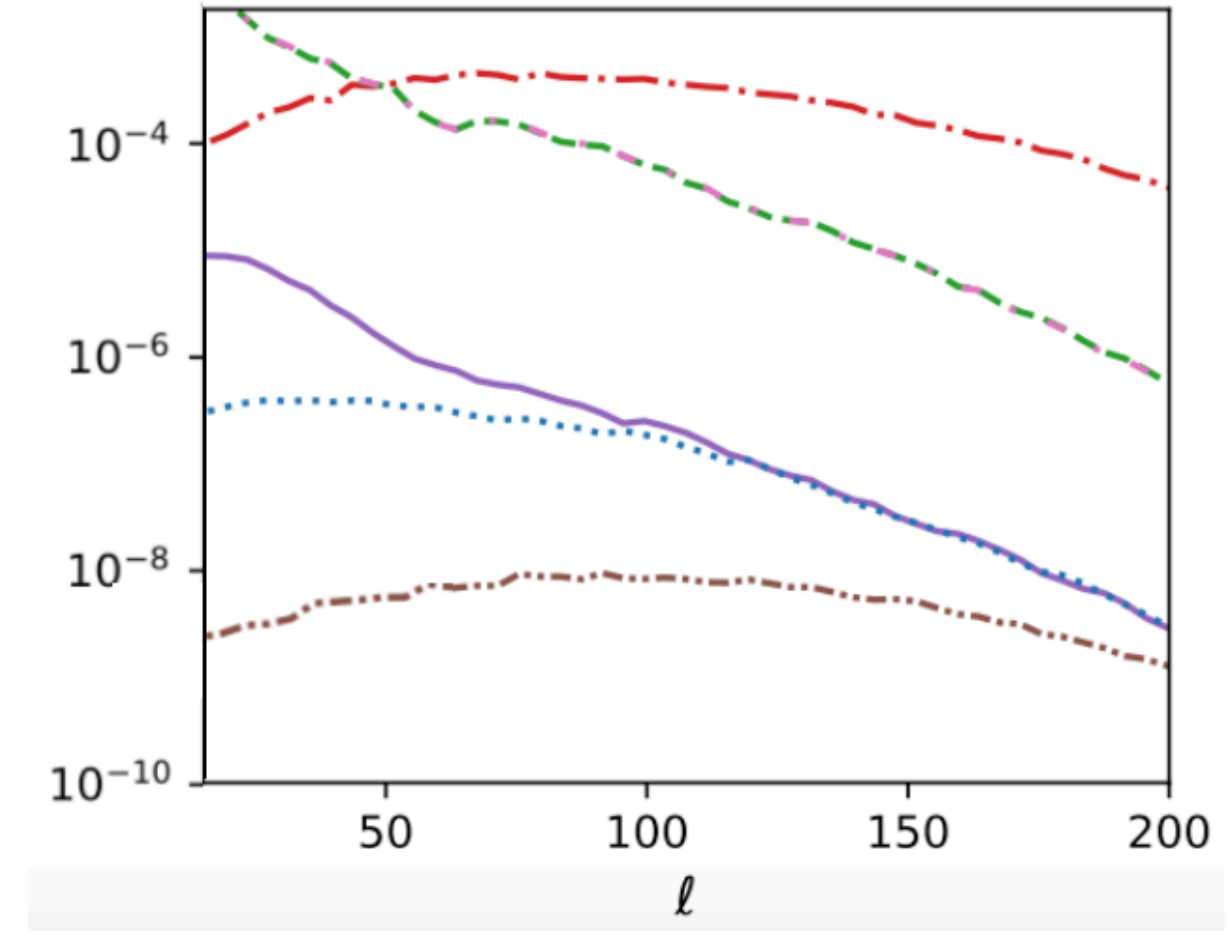
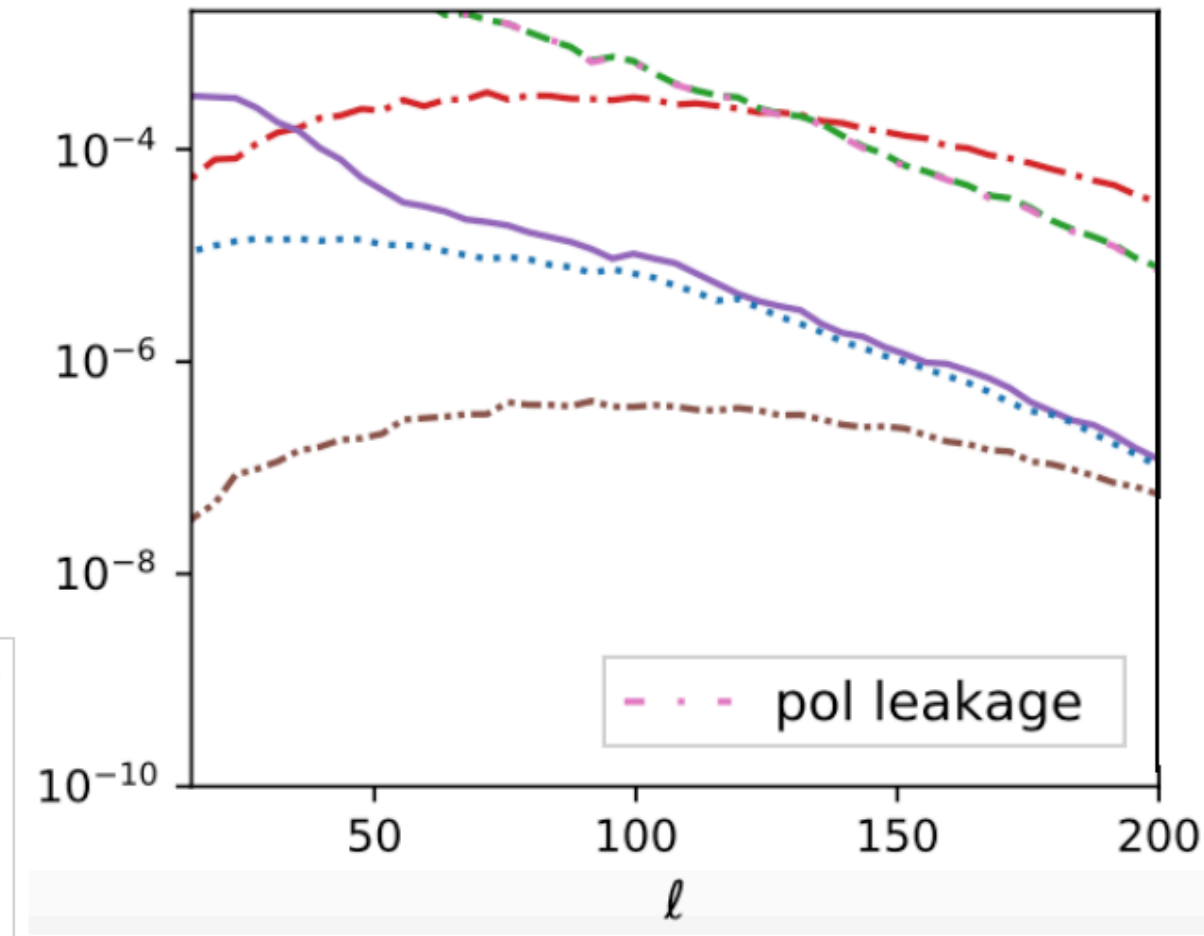
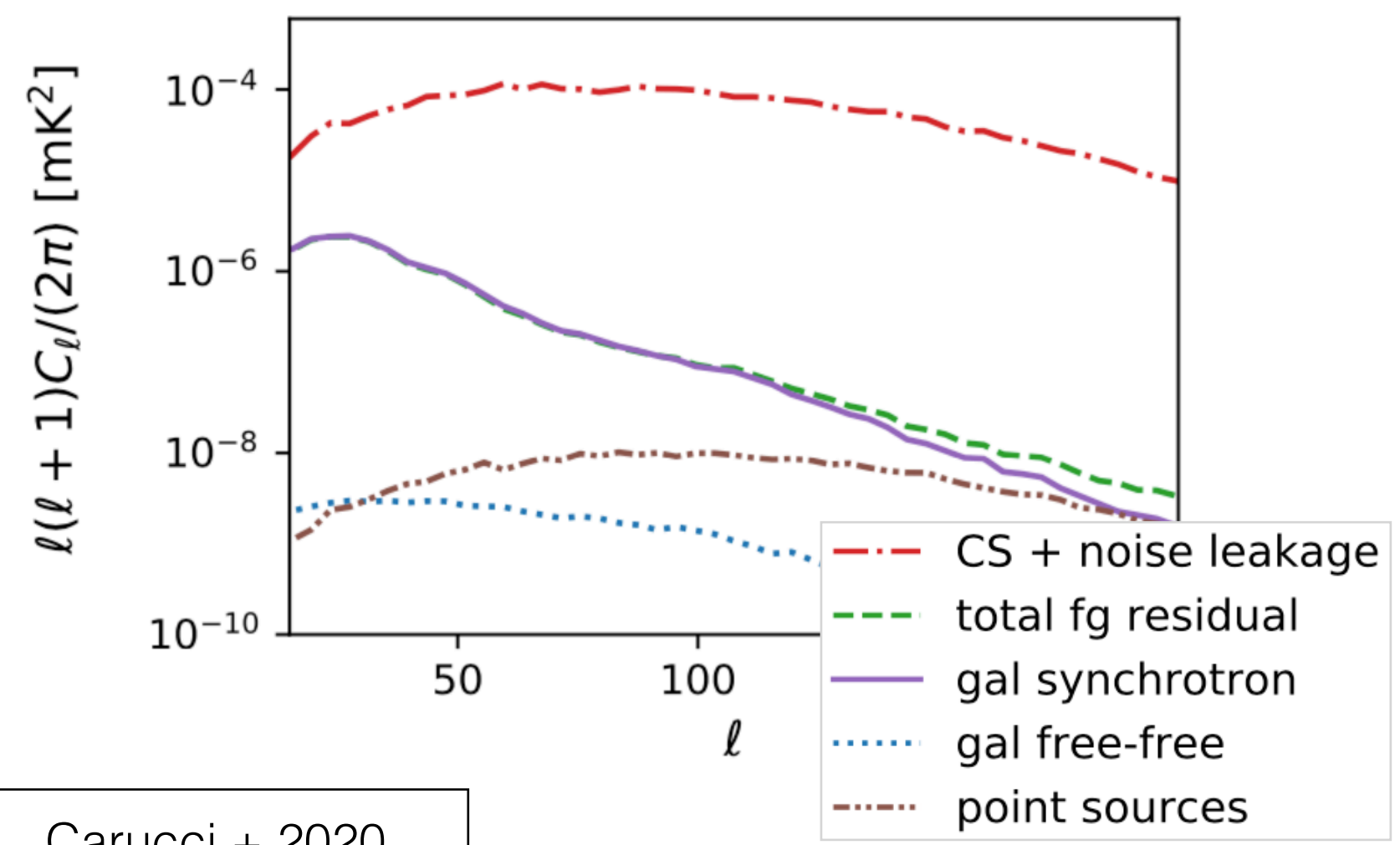
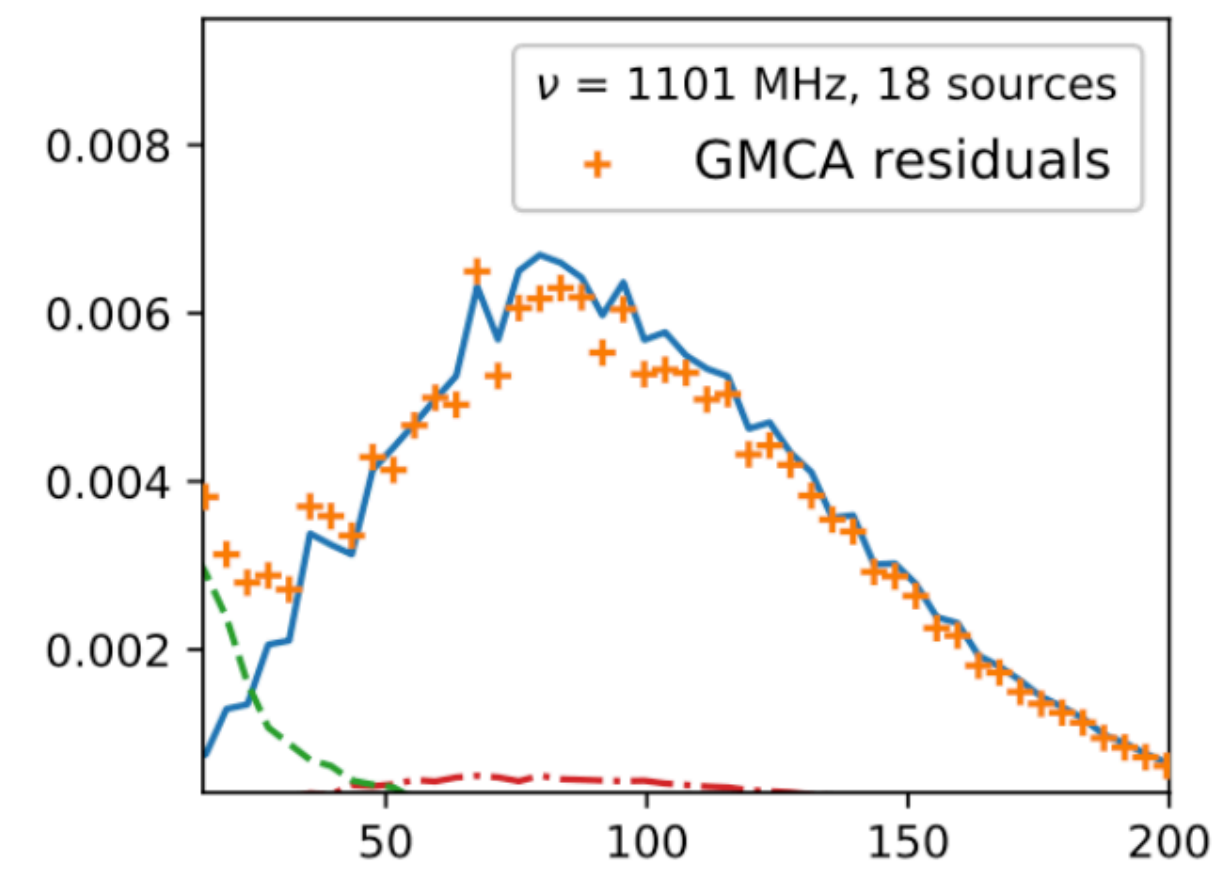
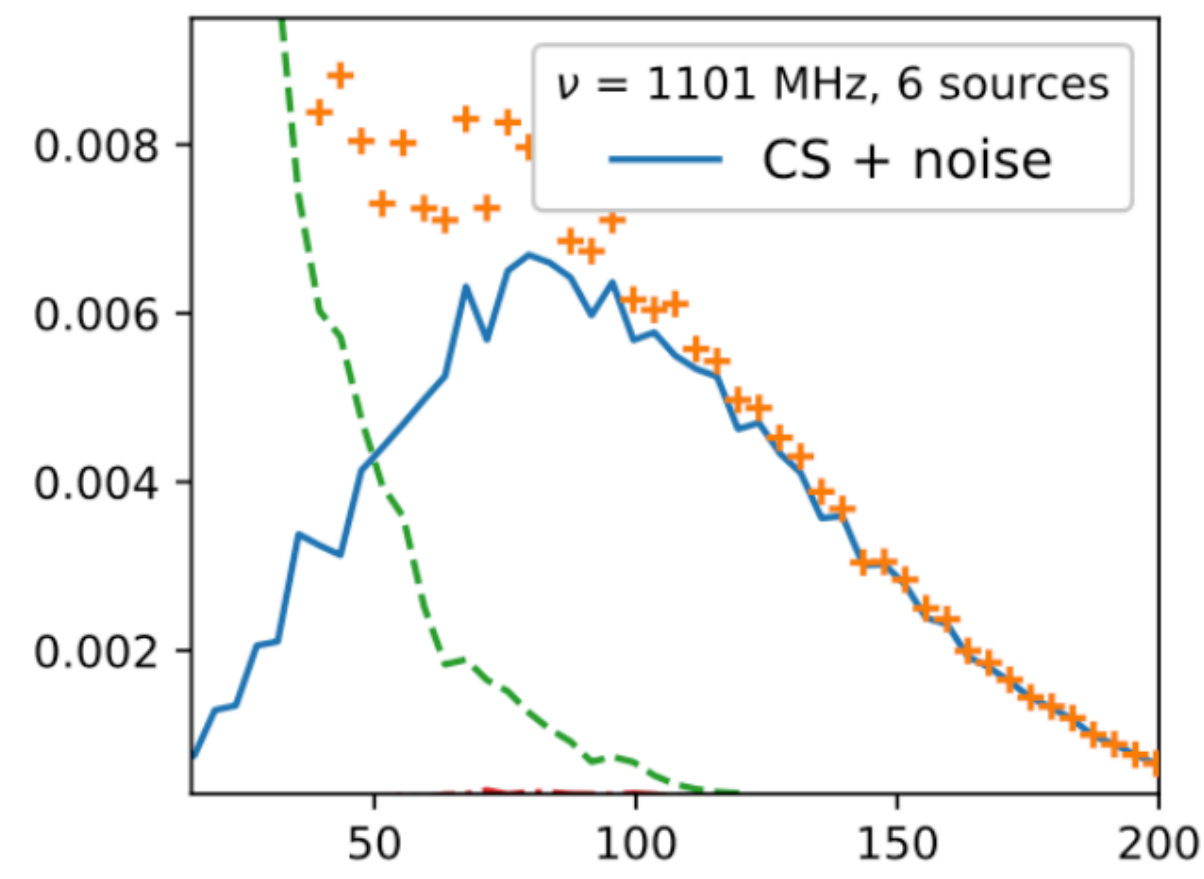
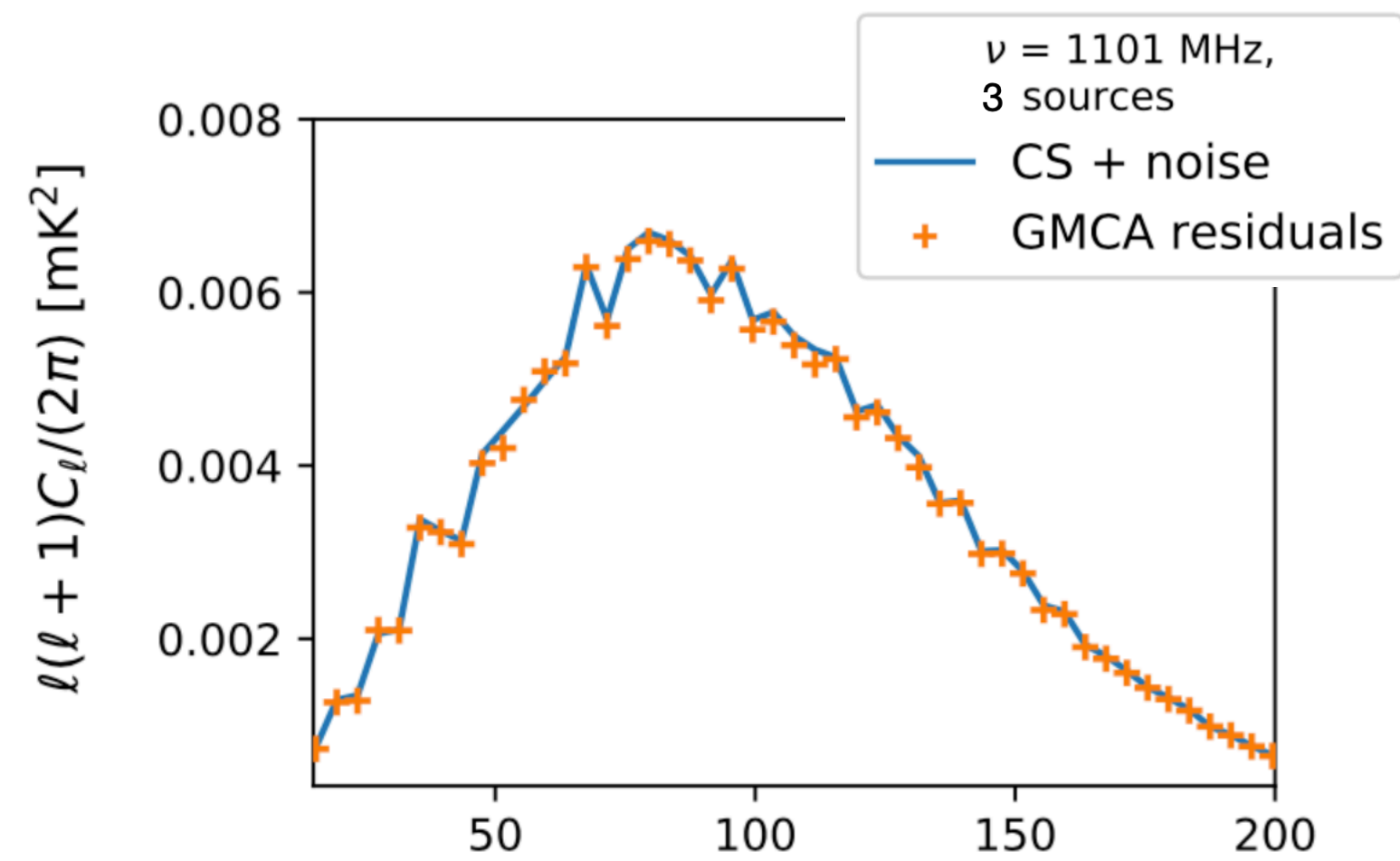
Carucci + 2020



Carucci + 2020

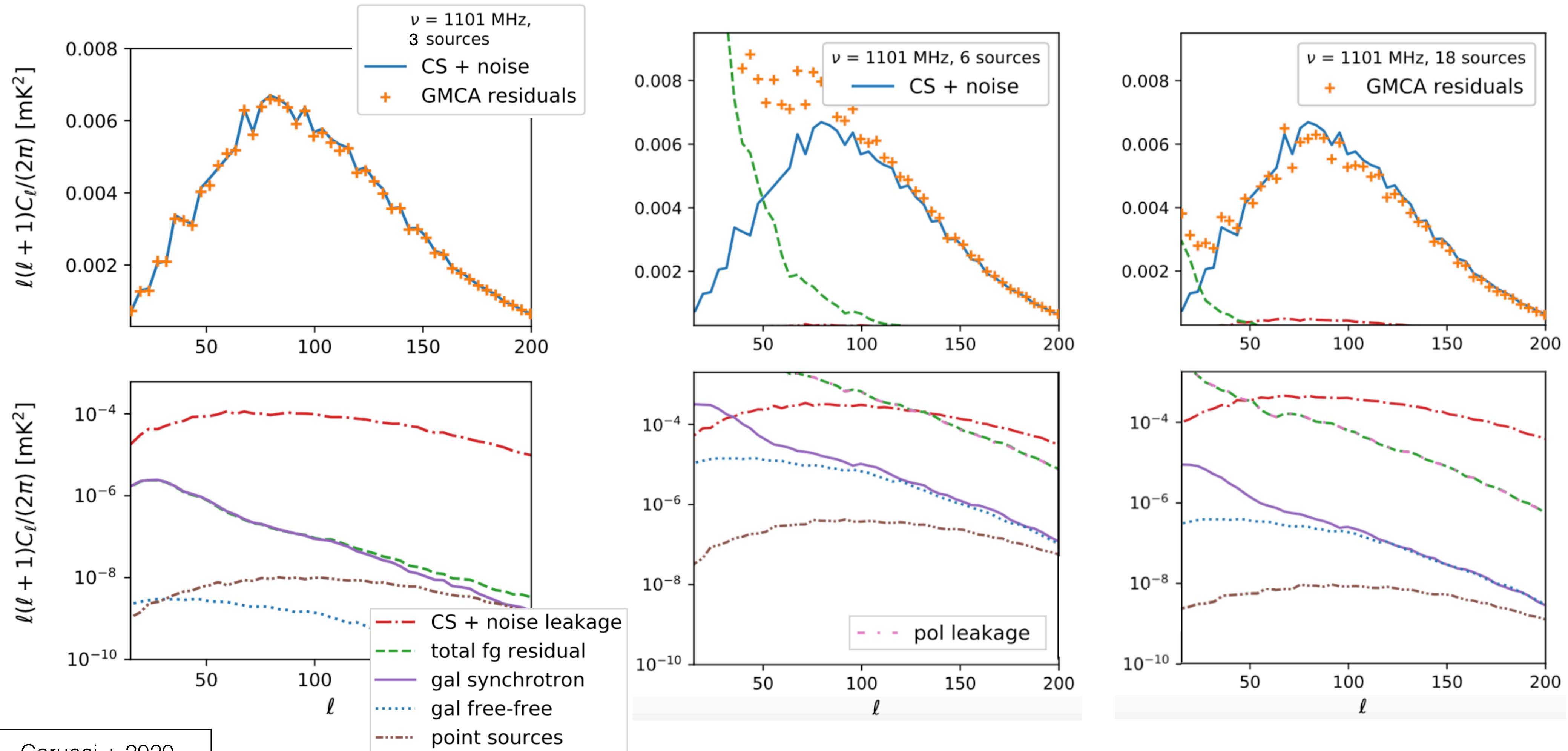


Carucci + 2020



Carucci + 2020

Different scales need different care



Carucci + 2020

MeerKLASS: MeerK_{AT} Large Area Synoptic Survey

ArXiv: 1709:06099

- in South Africa
- an SKA precursor (will be part of SKA-mid)
- 64 dishes of 13.5 m (soon ~100)
- Frequencies: 580 to 3500 MHz ($0 < z < 1.5$)

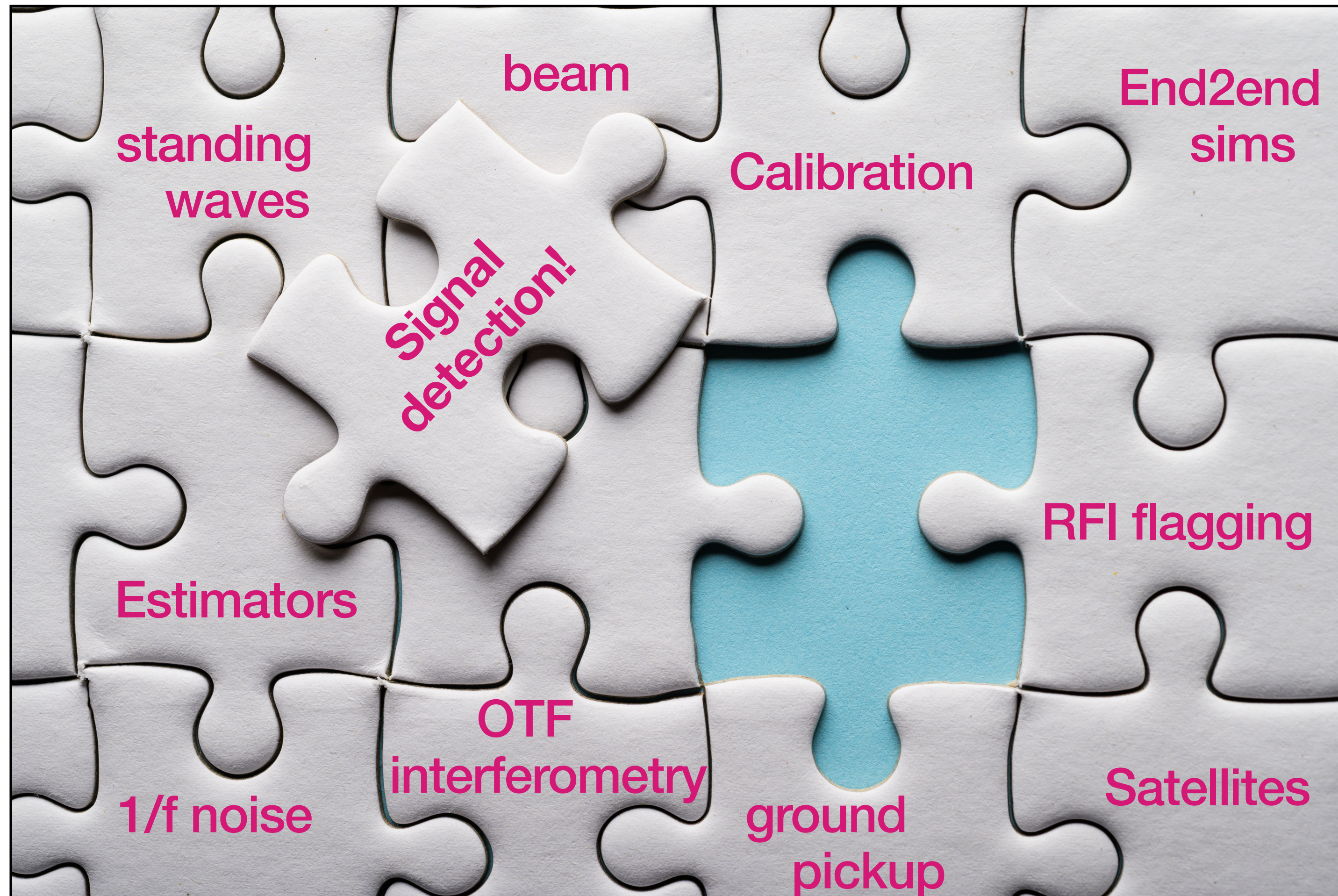
MeerKLASS: MeerK_{AT} Large Area Synoptic Survey

ArXiv: 1709:06099

- in South Africa
- an SKA precursor (will be part of SKA-mid)
- 64 dishes of 13.5 m (soon ~100)
- Frequencies: 580 to 3500 MHz ($0 < z < 1.5$)

- Wang + 2021 — **Calibration** paper
- Li + 2021 — **1/f noise** characterisation
- Irfan + 2022 — **Synchrotron Spectral Index Measurement**
- Cunnington, Li+ 2023 — **Detection in cross-correlation with WiggleZ galaxies**
- MeerKLASS coll. 2024 — Detection in cross with GAMA and **stacking**

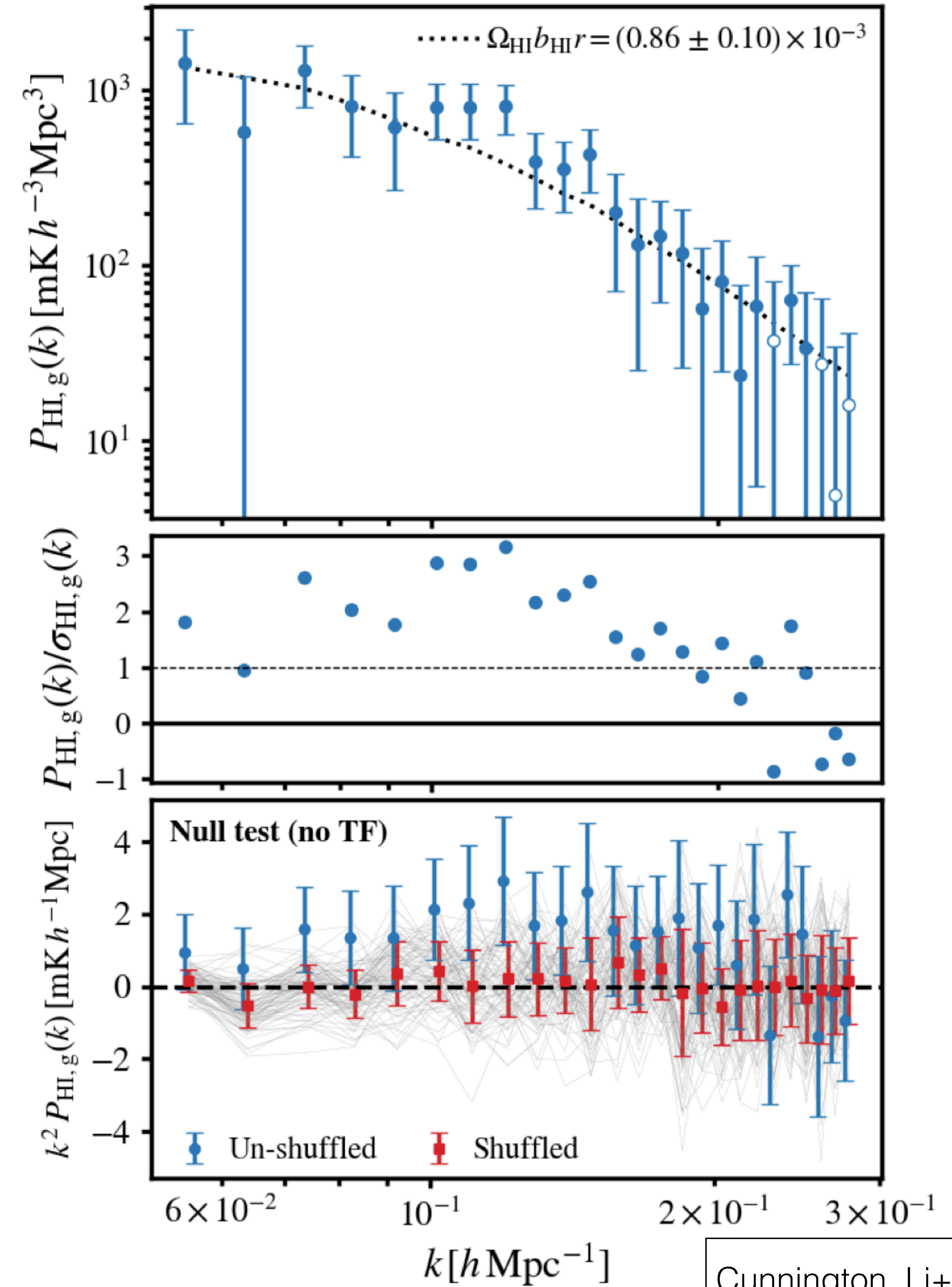
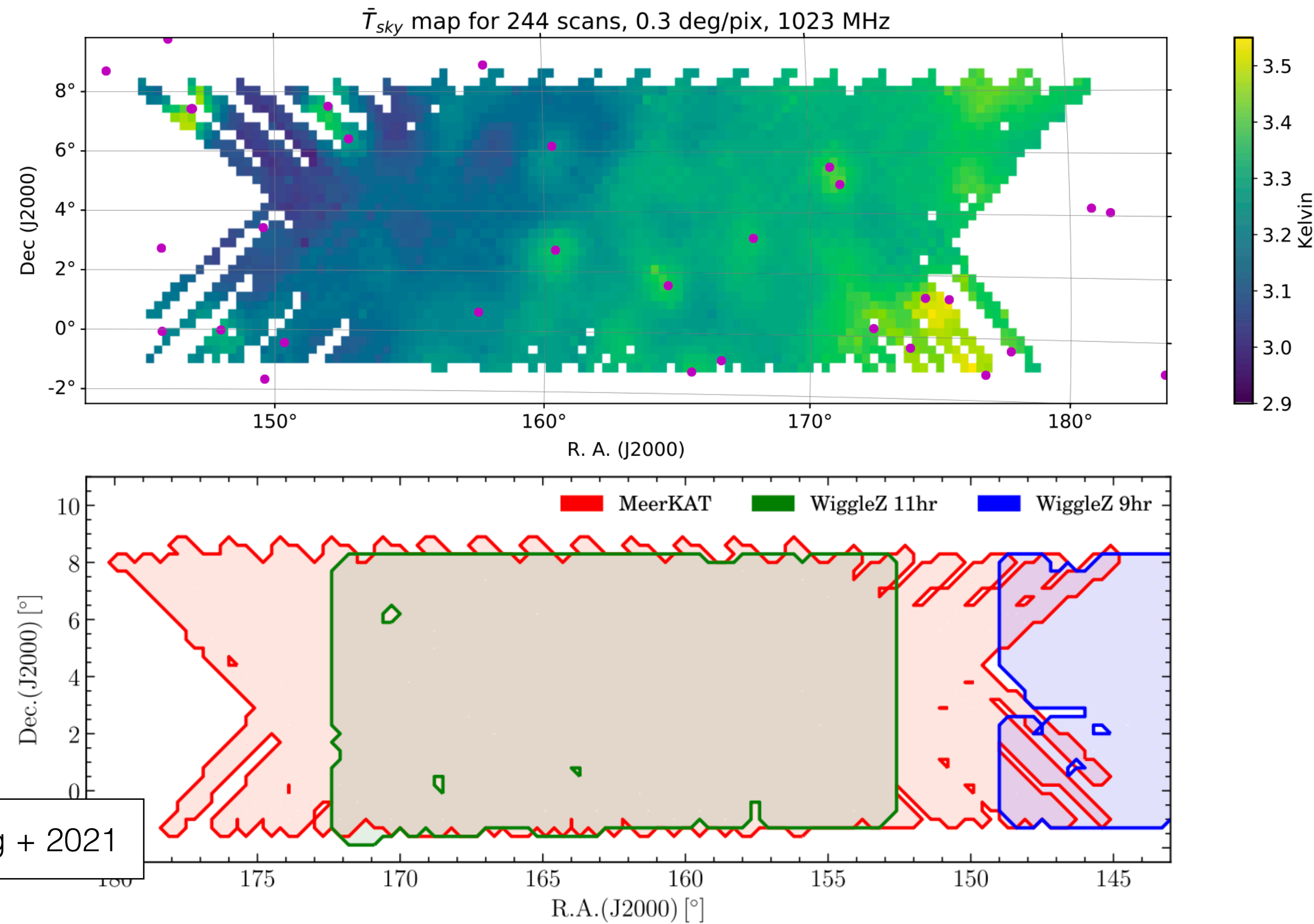
a collective effort



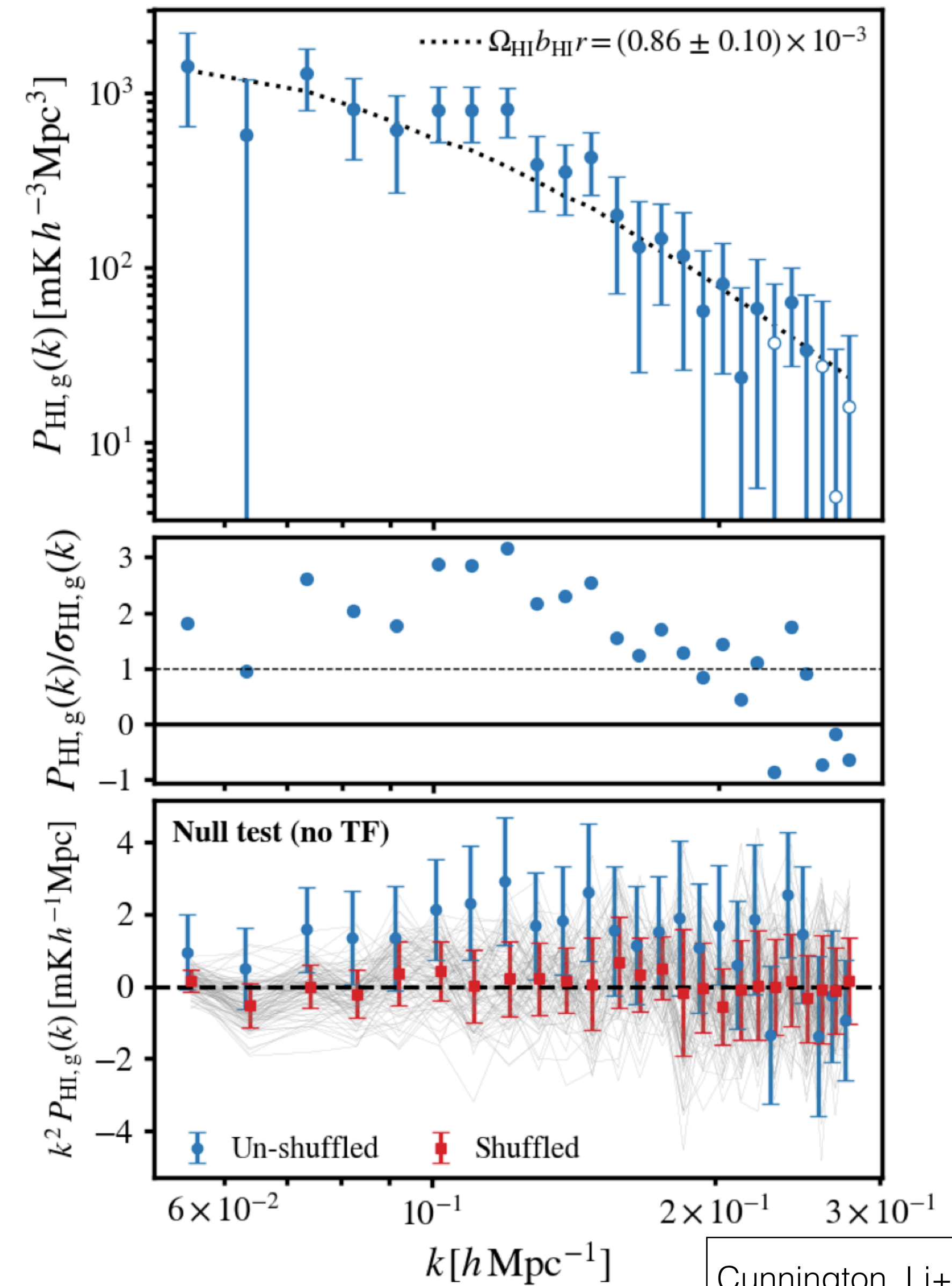
Mário Santos, Aishrila Mazumder, Alkistis Pourtsidou, Brandon Engelbrecht, Isabella P. Carucci, Jingying Wang, José Fonseca, José Luis Bernal, Karin Fornazier, Laura Wolz, Marta Spinelli, Matilde Barberi Squarotti, Piyanat Boom Kittiwisit, Siyambonga Matshawule, Sourabh Paul, Steve Cunnington, Suman Chatterjee, Tamirat Gebeyehu, Wenkai Hu ...

Pilot survey data (2019):

- 10.5 hours of data from six nights of observations
- Overlapping with the WiggleZ 11hr field (~200 deg²)
- We use data in range 973-1015 MHz ($0.40 < z < 0.46$)

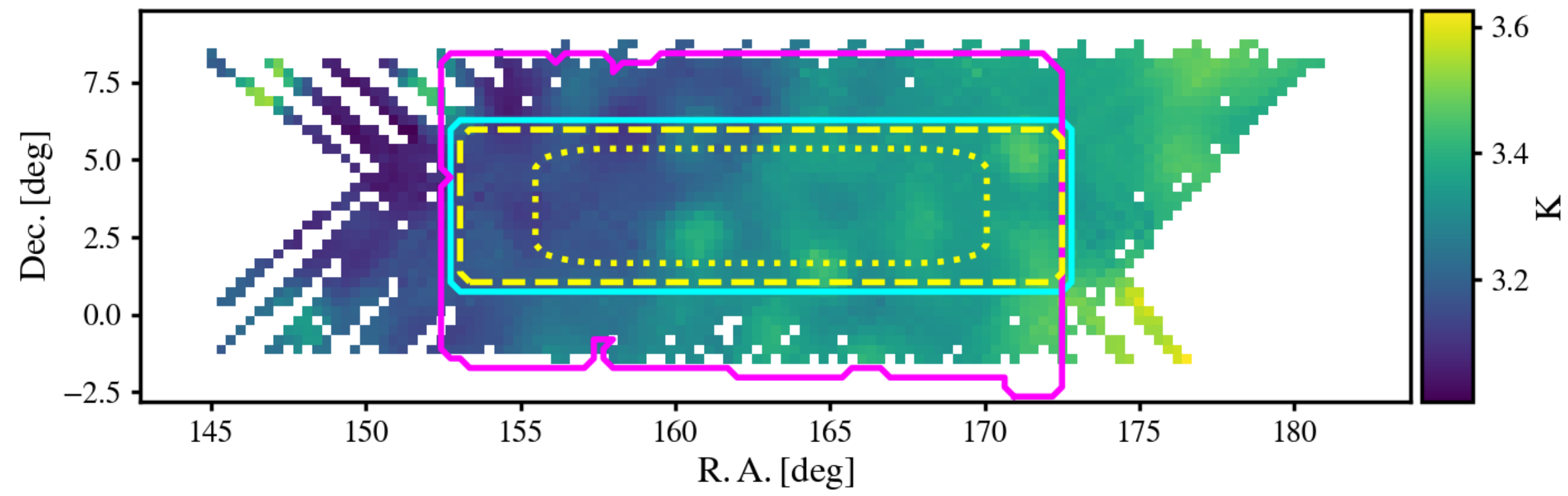


Can we use this cross-corr
detection as a benchmark to
learn something about our
cleaning strategy?



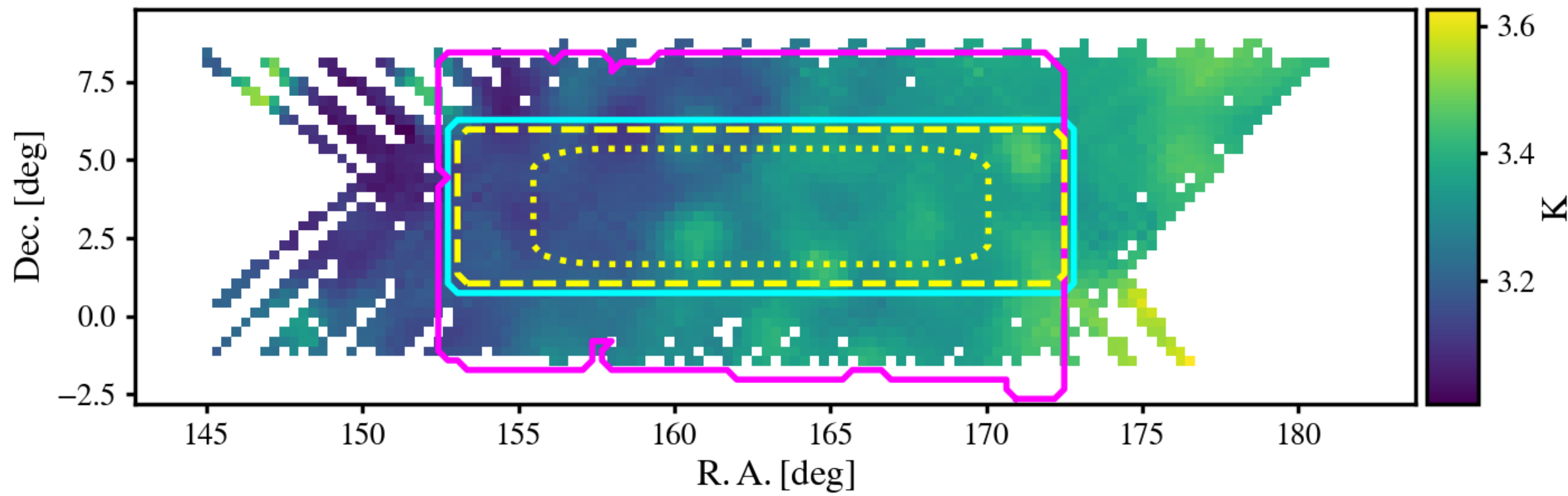
Cunnington, Li+ 2023

original map

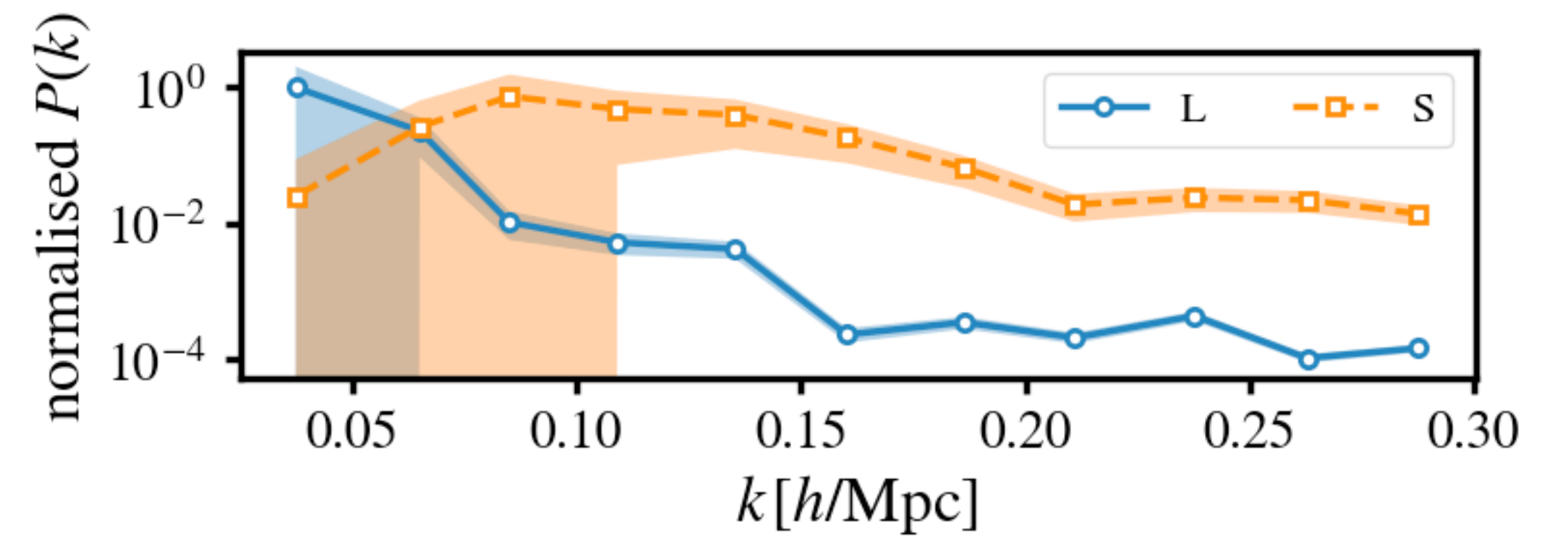
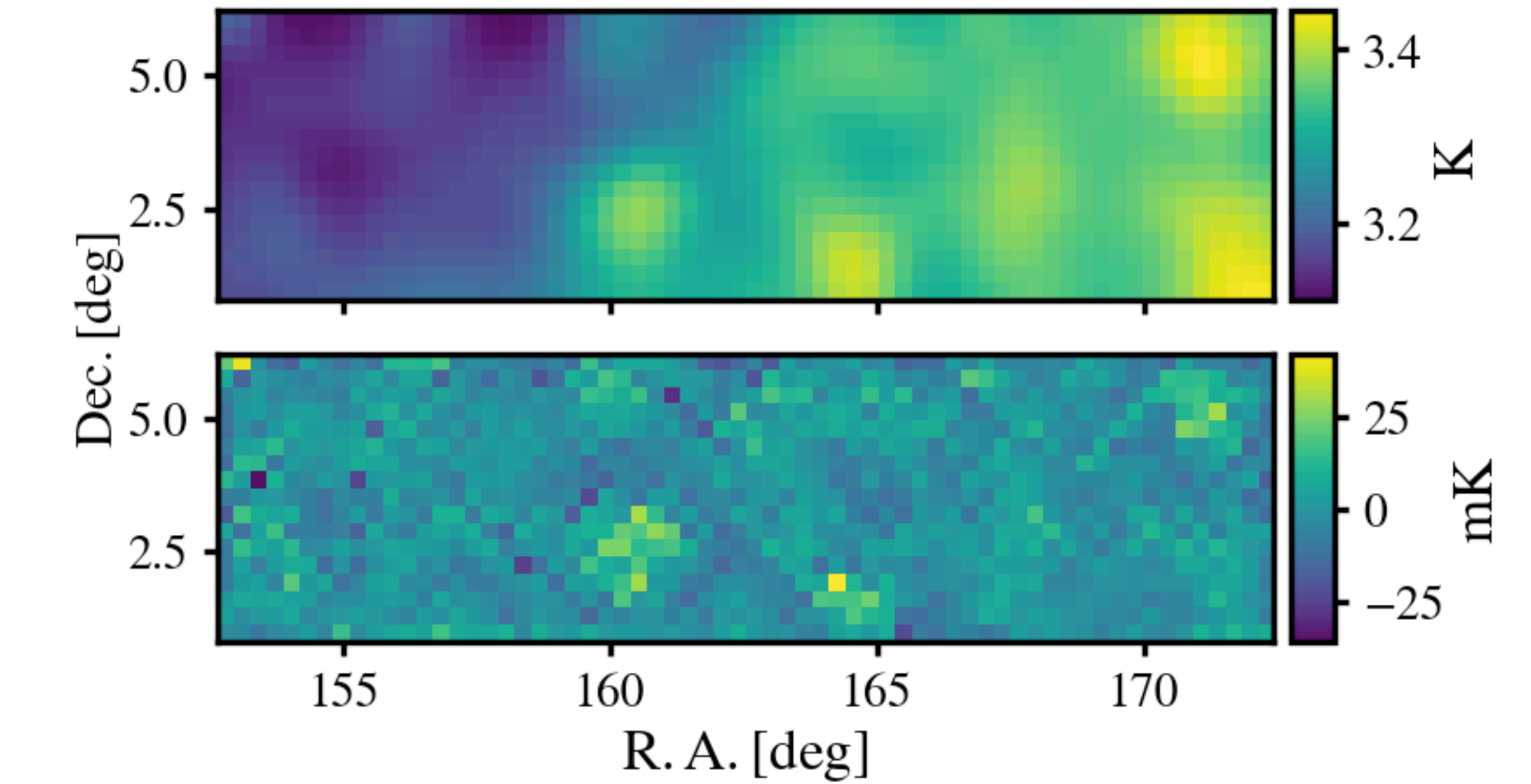


multiscale

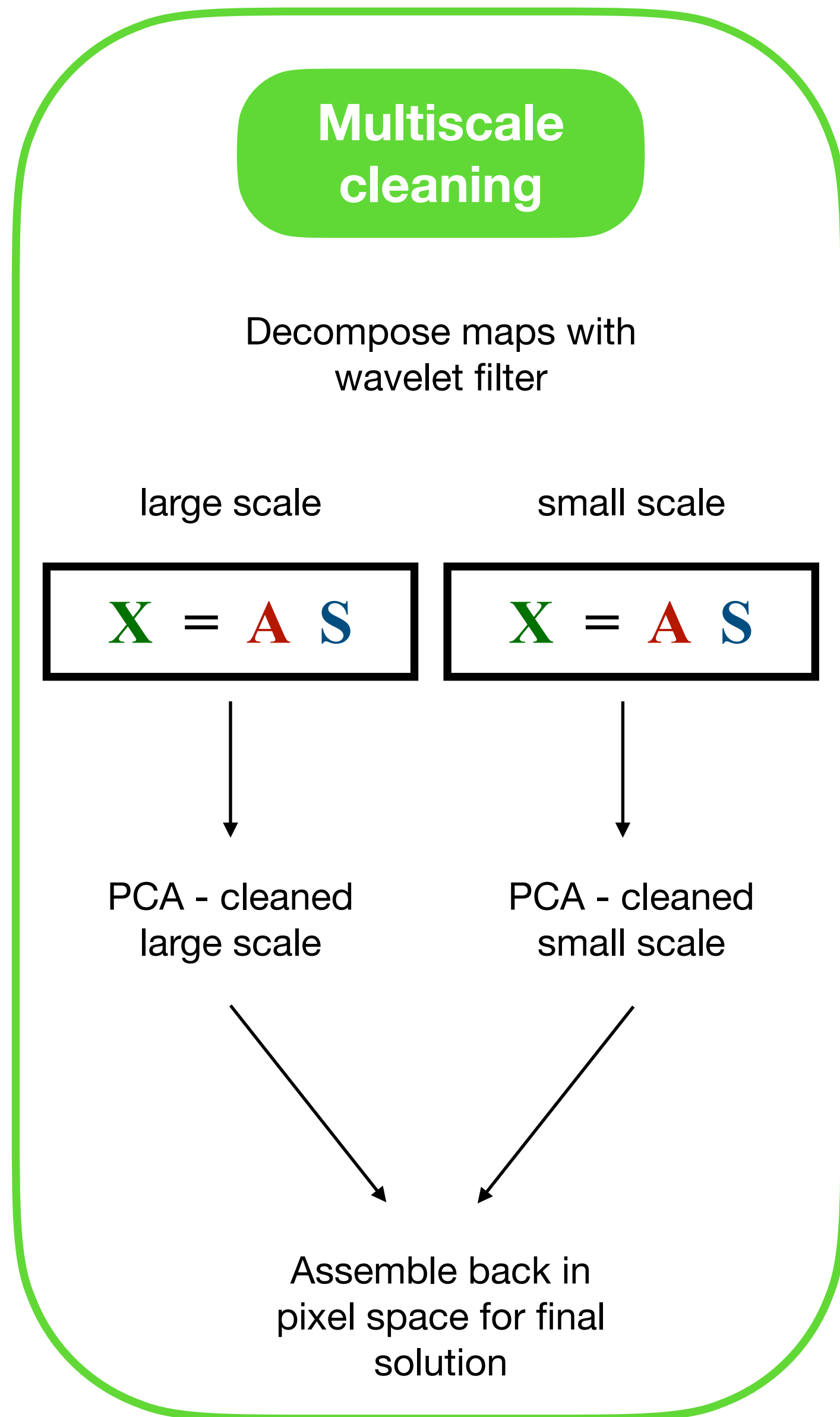
original map



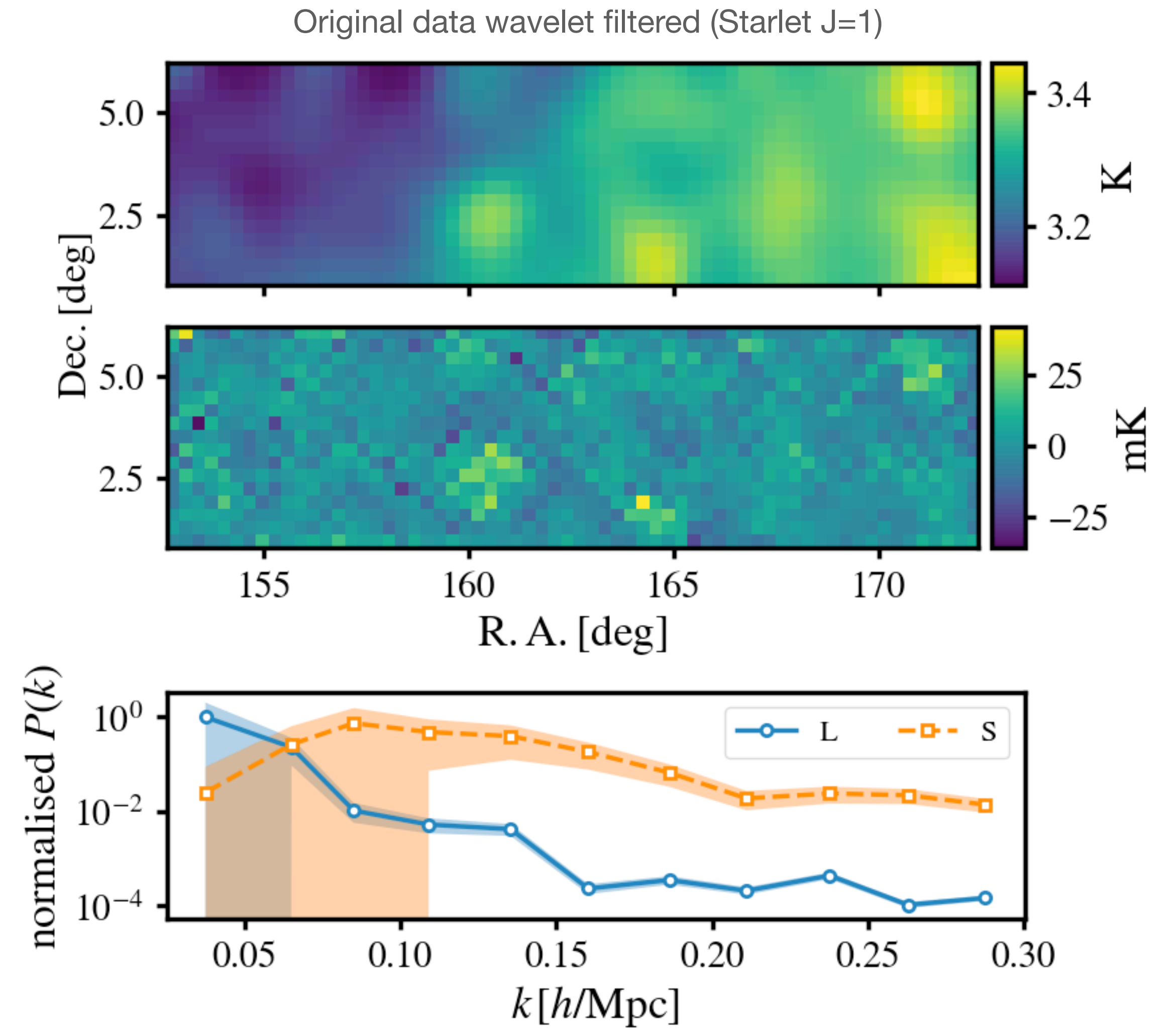
Original data wavelet filtered (Starlet J=1)



Carucci, + in prep.

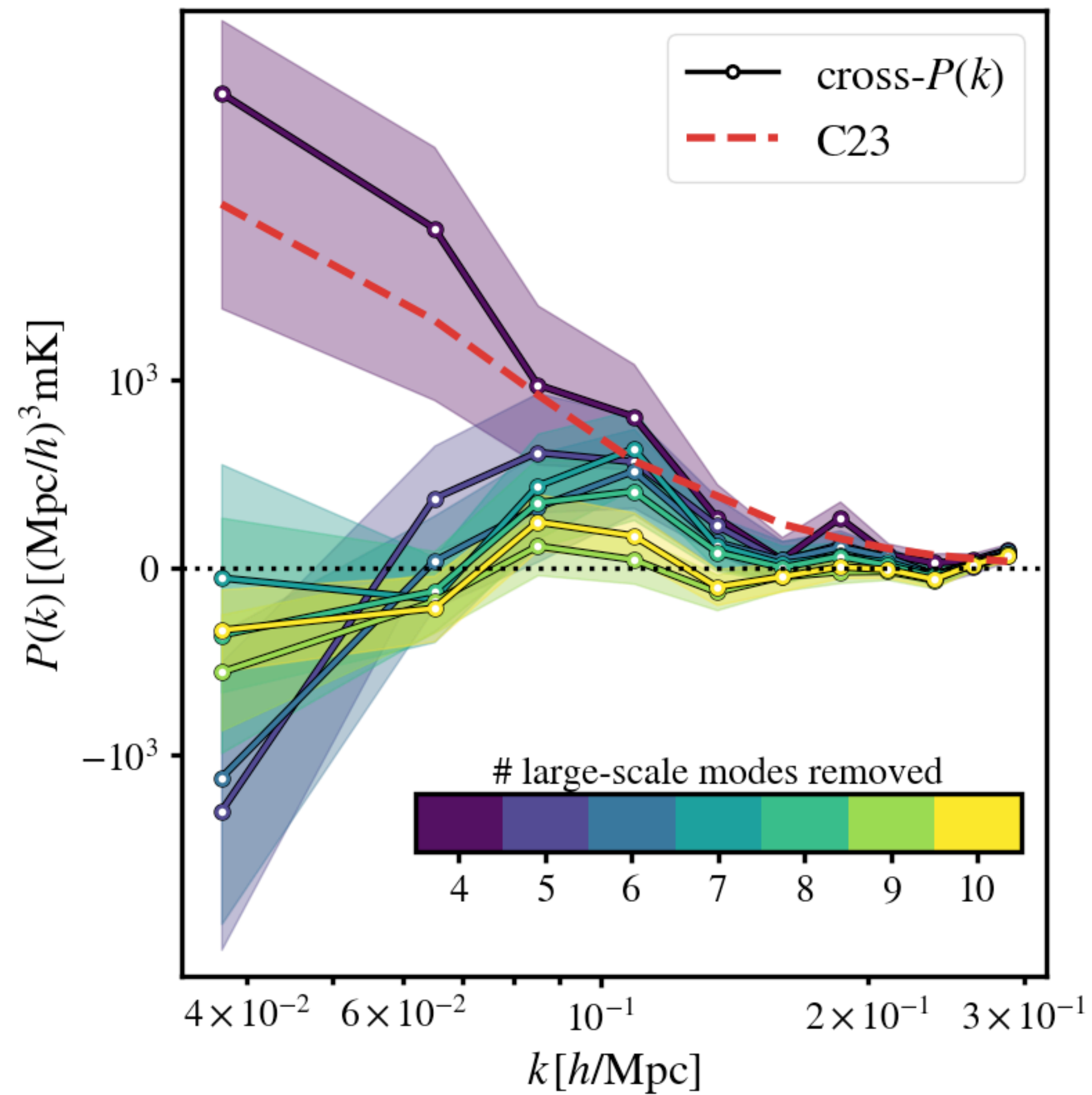


multiscale



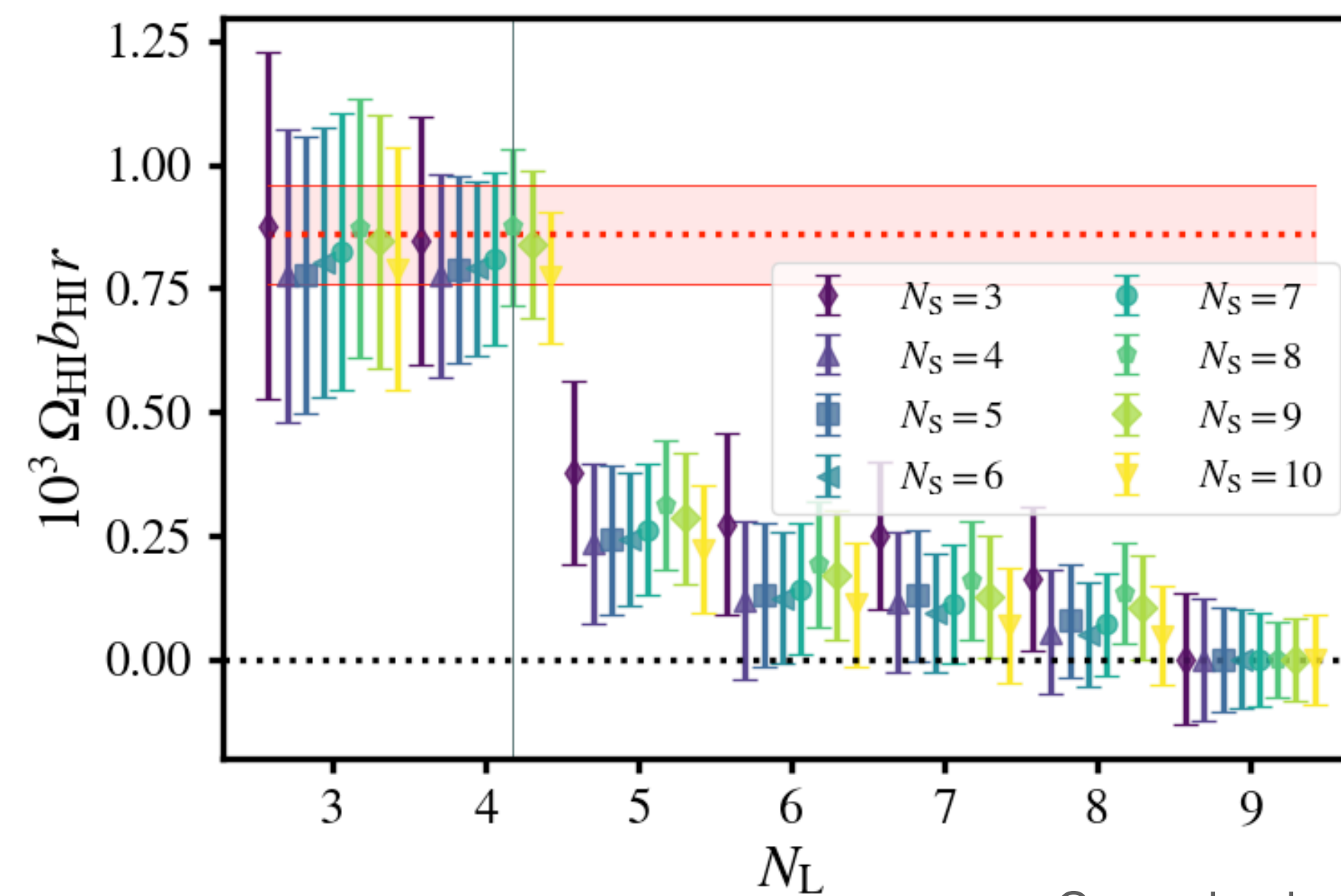
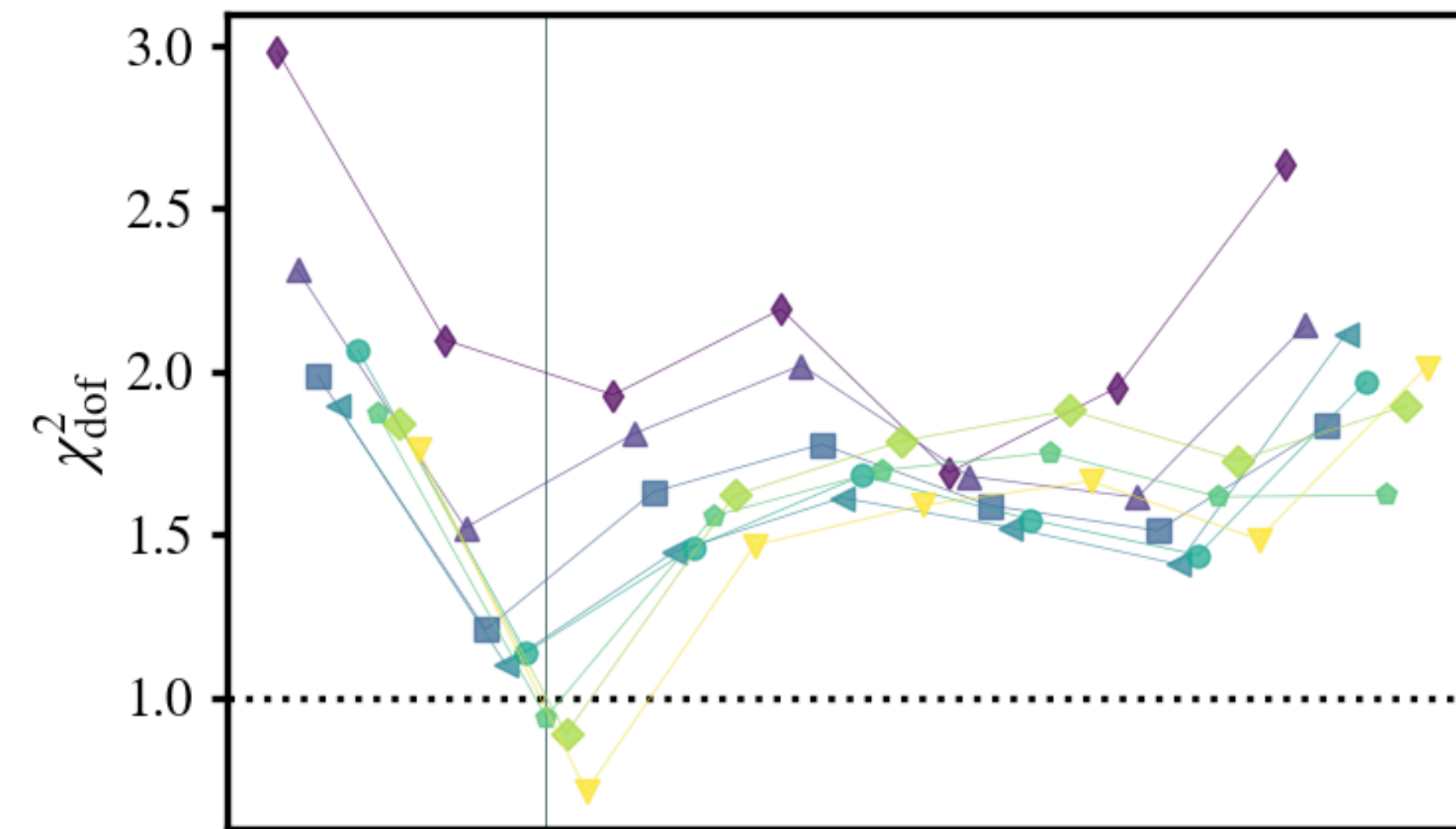
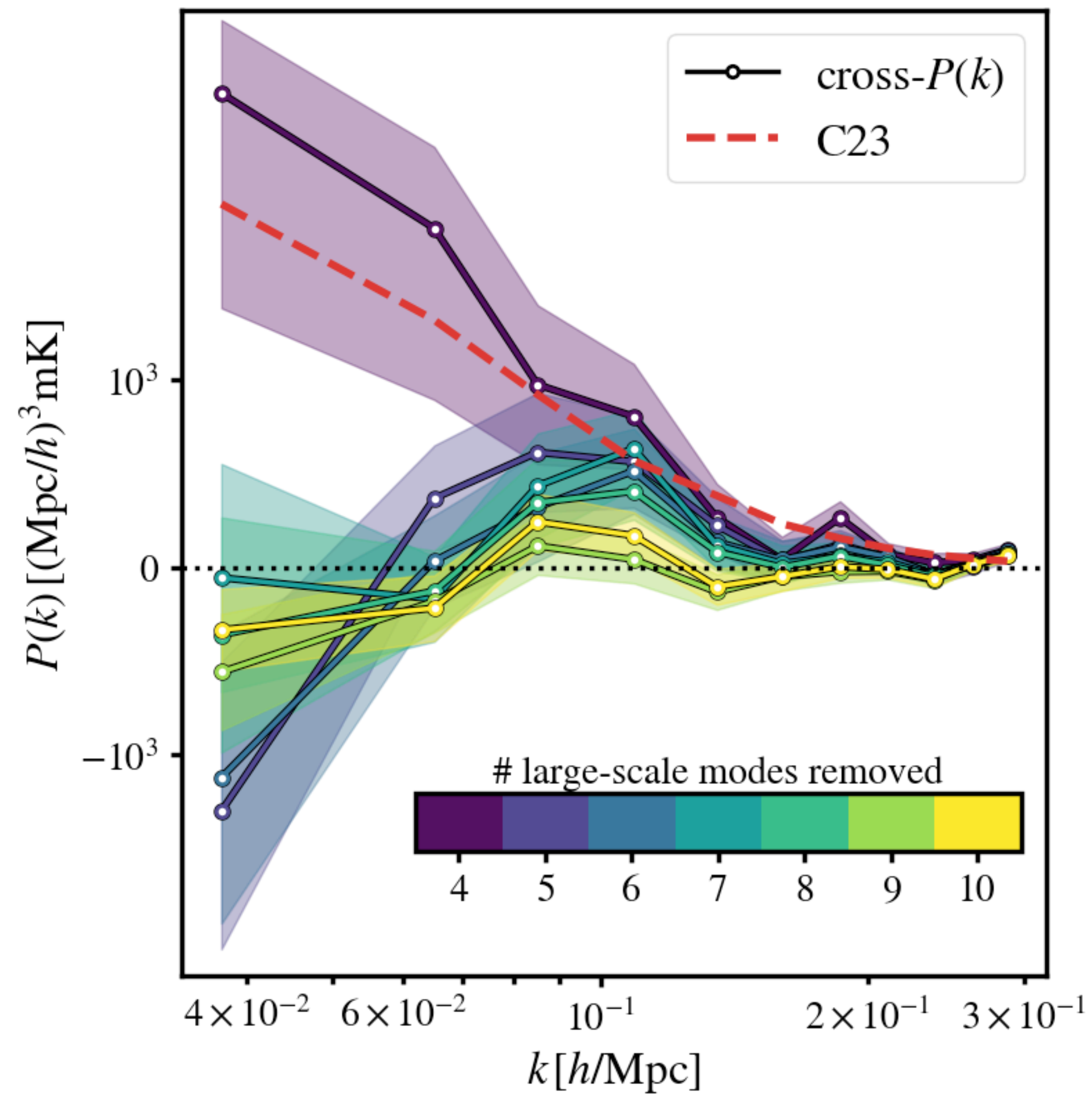
Carucci, + in prep.

30 vs 4 modes removed!



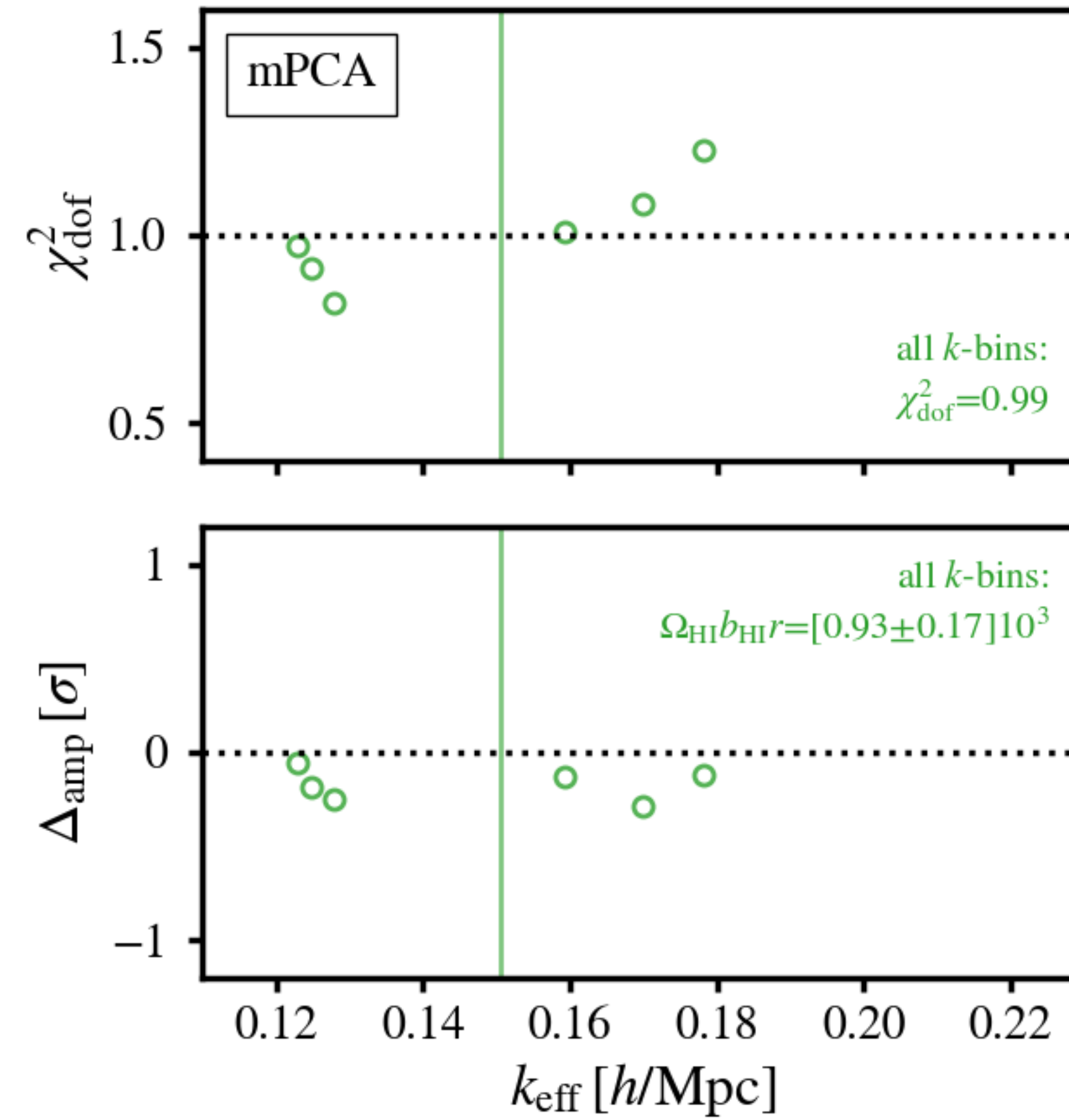
Carucci, + in prep.

30 vs 4 modes removed!



Carucci, + in prep.

A robust result against signal loss

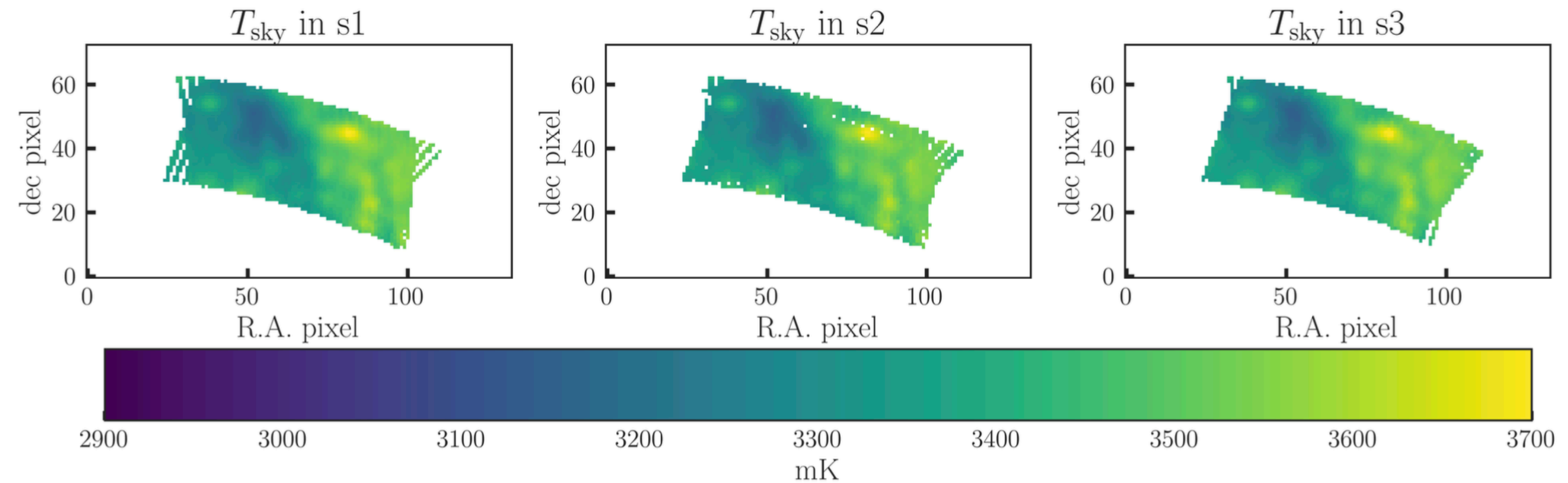


Carucci, + in prep.

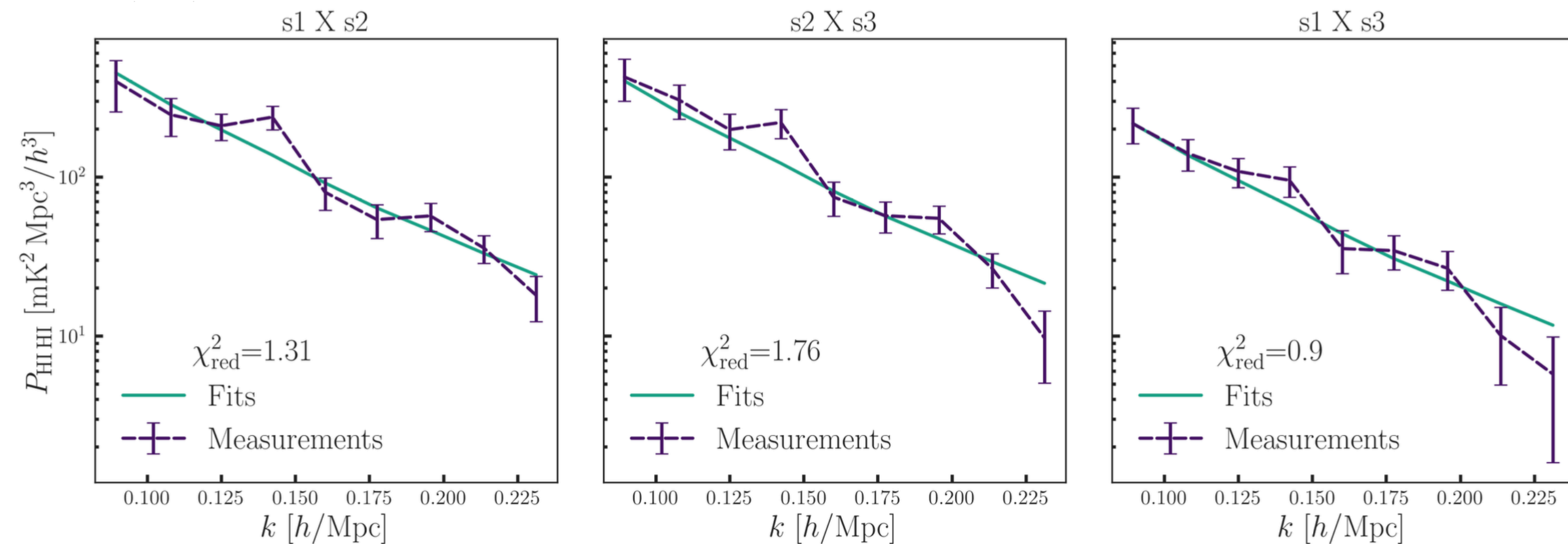
Are we ready to move forward without the galaxies?

Are we ready to move forward without the galaxies?

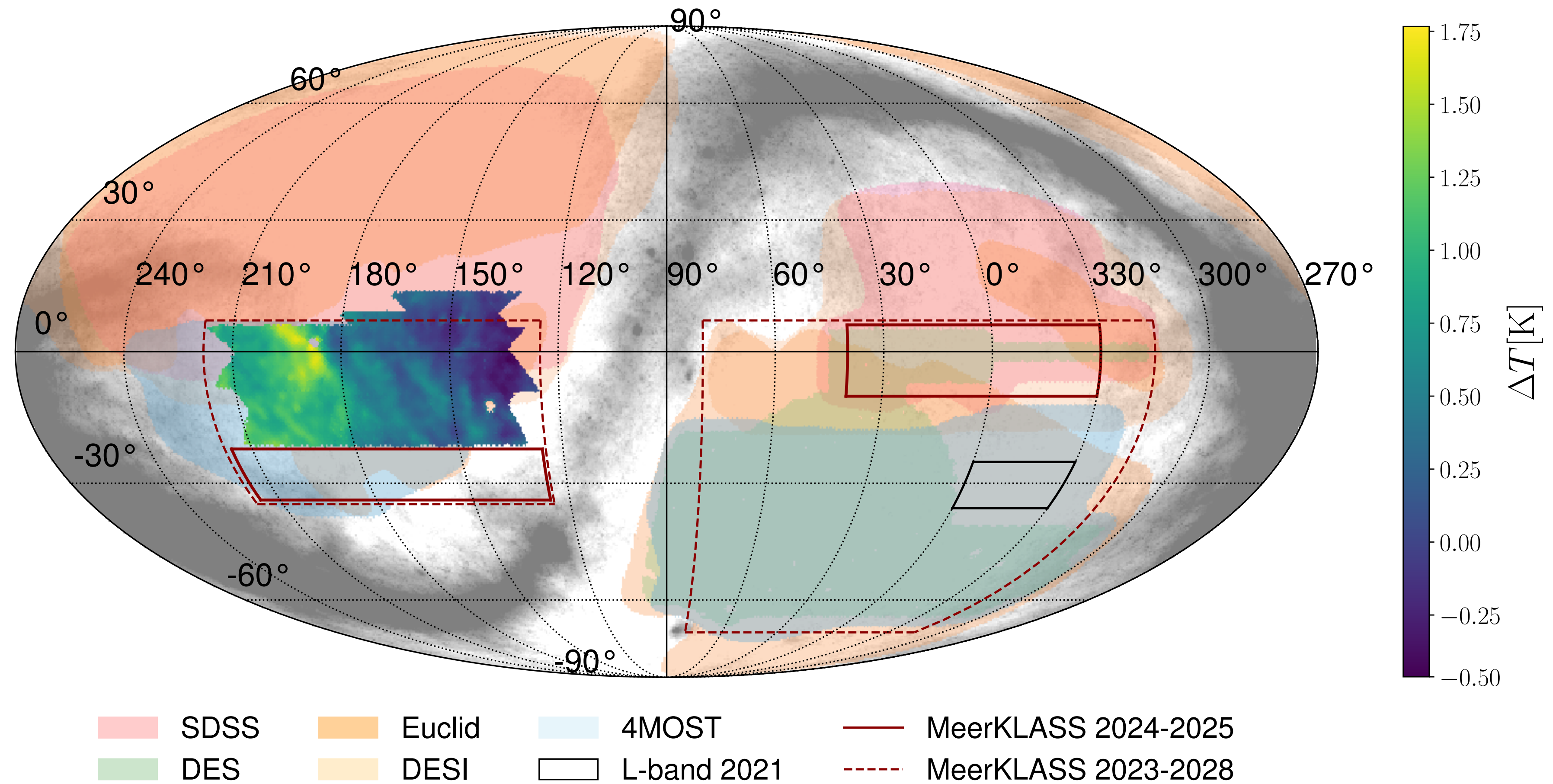
Internal cross-correlations



Work led by Matilde Barberi Squarotti

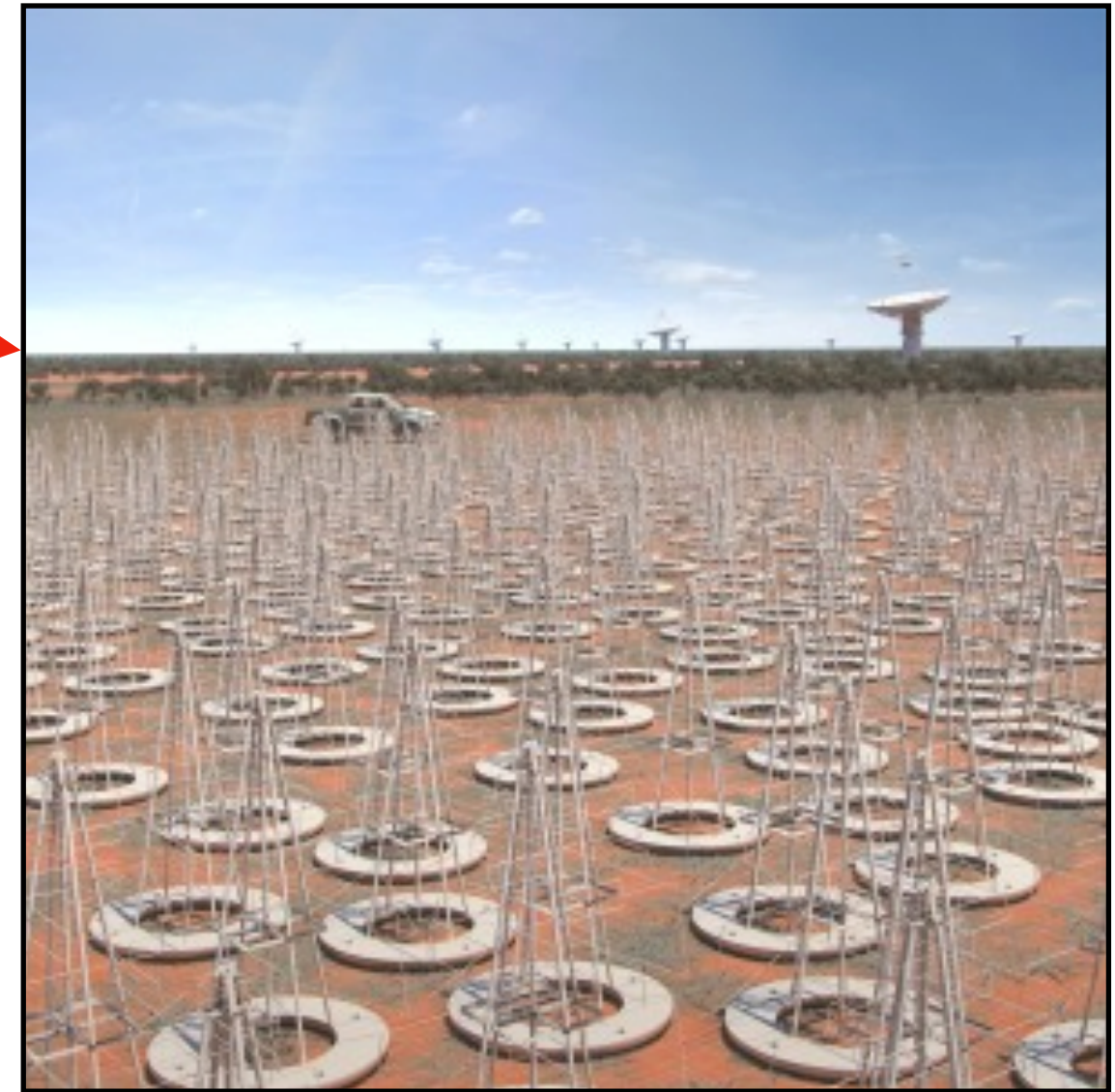
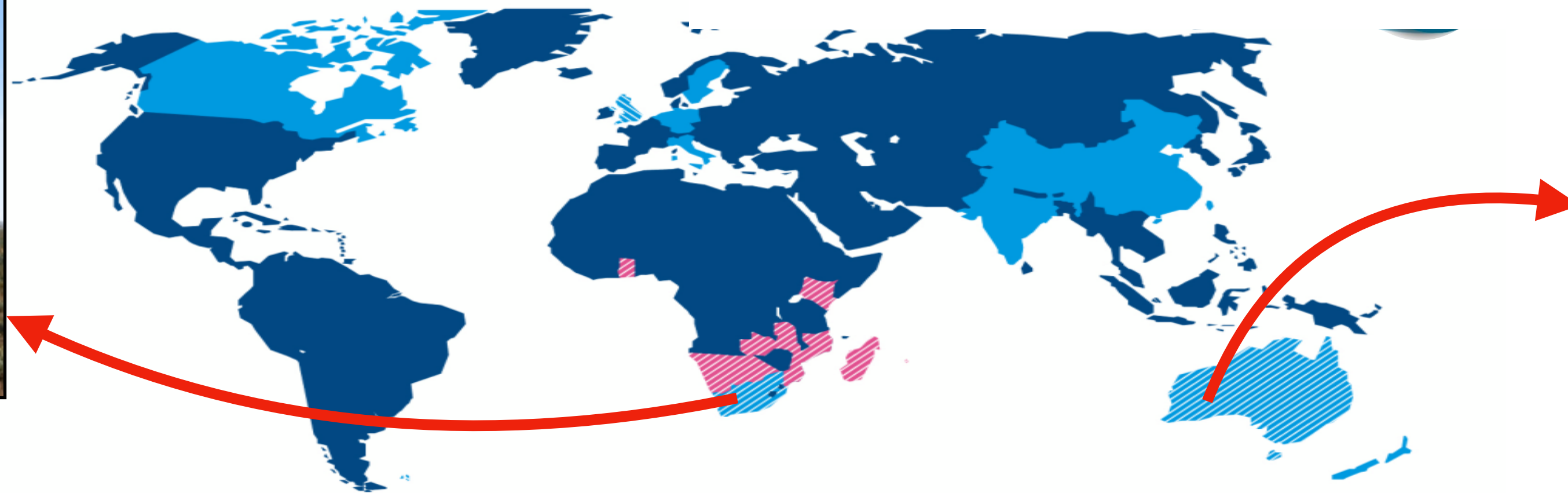


Exciting times ahead !



Credit: MeerKLASS coll.

SKAO



- Full members
- SKA Headquarters host country
- SKA Phase 1 and Phase 2 host countries

- African partner countries (non-member SKA Phase 2 host countries)

This map is intended for reference only and is not meant to represent legal borders

SKA1-mid

the SKA's mid-frequency instrument

$0 < z < 3$



Location:
South Africa



Frequency range:
350 MHz
to
15.3 GHz
with a goal of 24 GHz



197 dishes
(including 64 MeerKAT dishes)



Maximum baseline:
150km

SKA1-low

the SKA's low-frequency instrument

$3 < z < 27$



Location: Australia



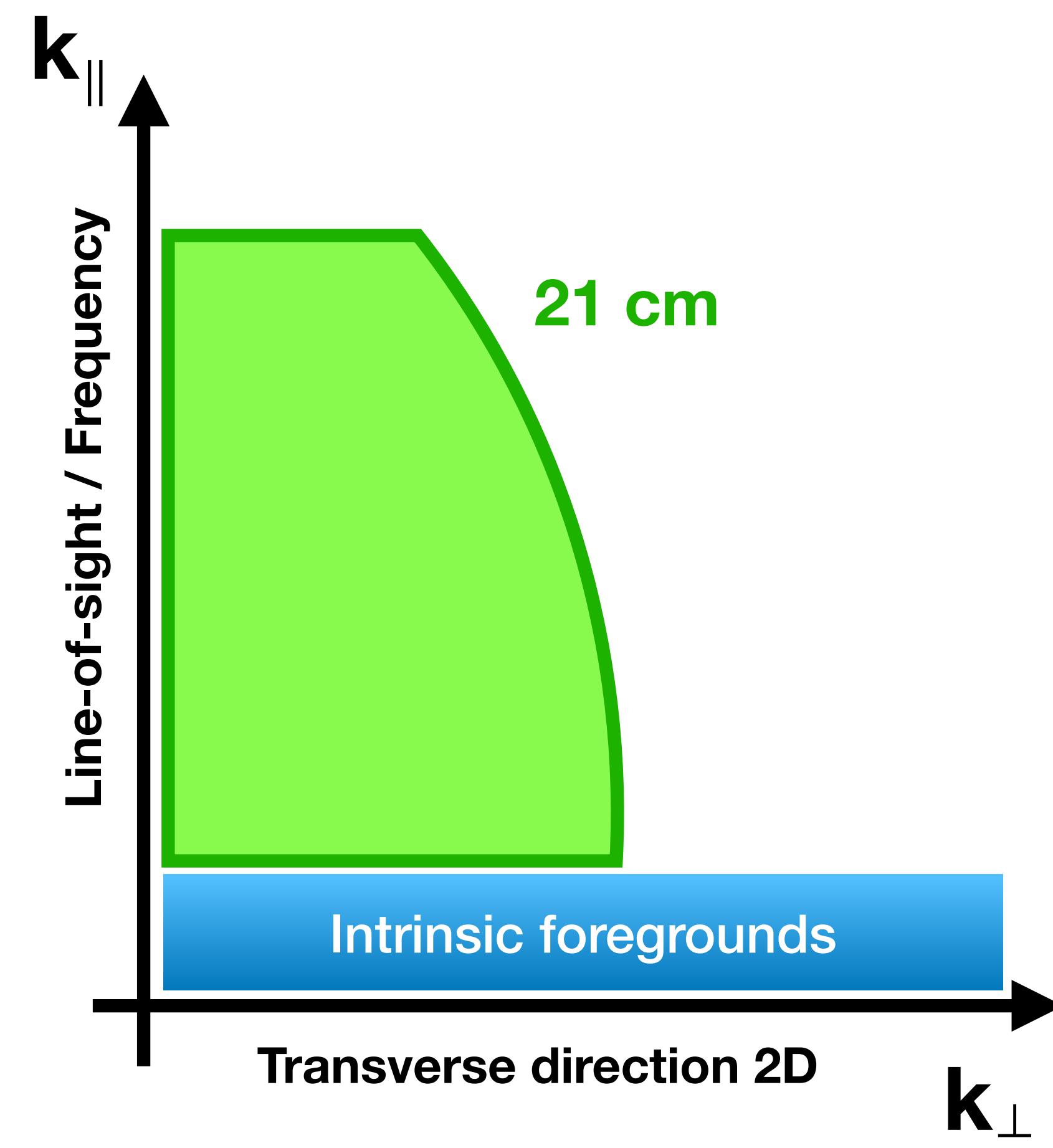
Frequency range:
50 MHz
to
350 MHz



~131,000
antennas spread between
512 stations

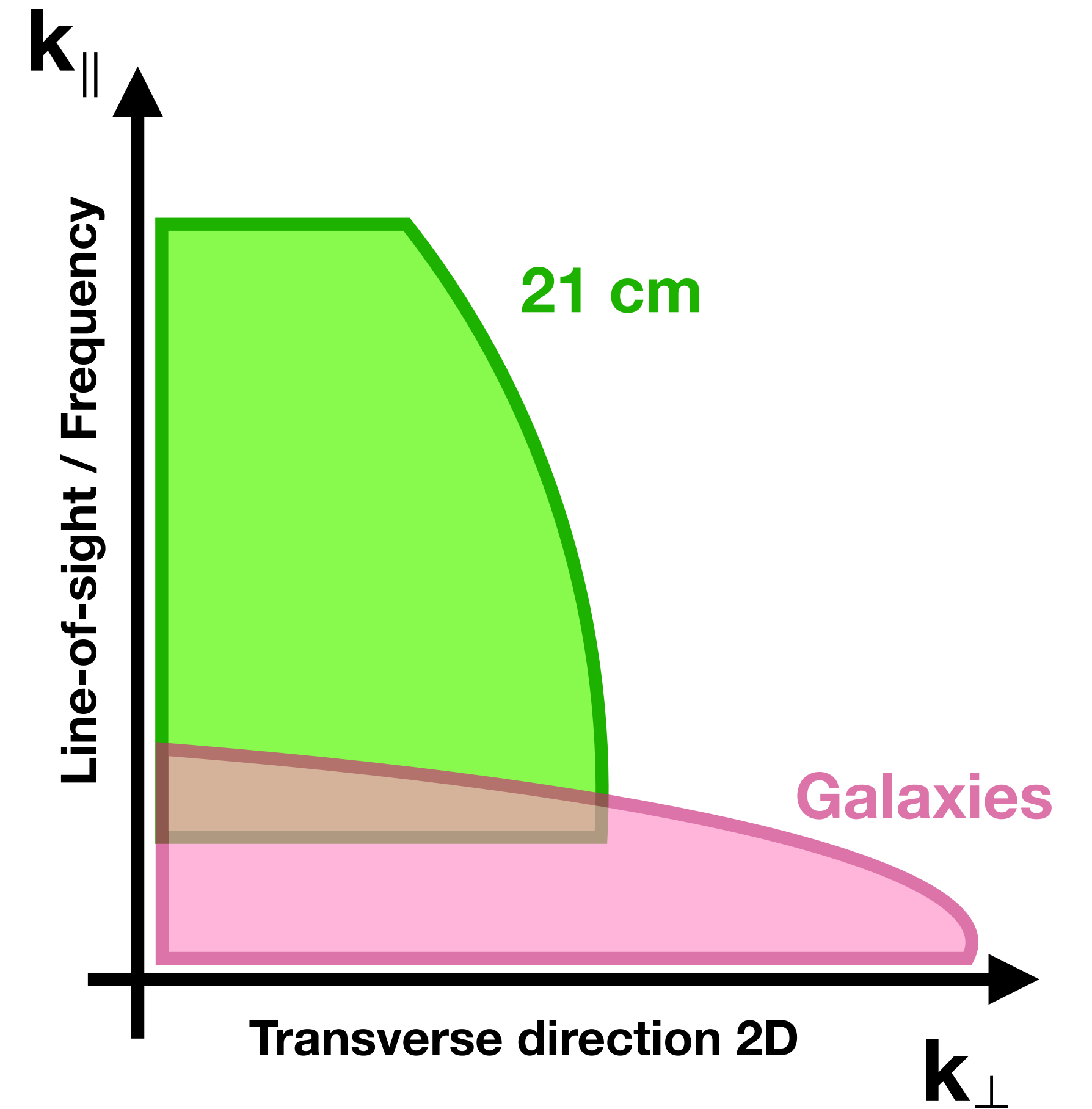


Maximum baseline:
~65km



Beat systematics: cross correlations

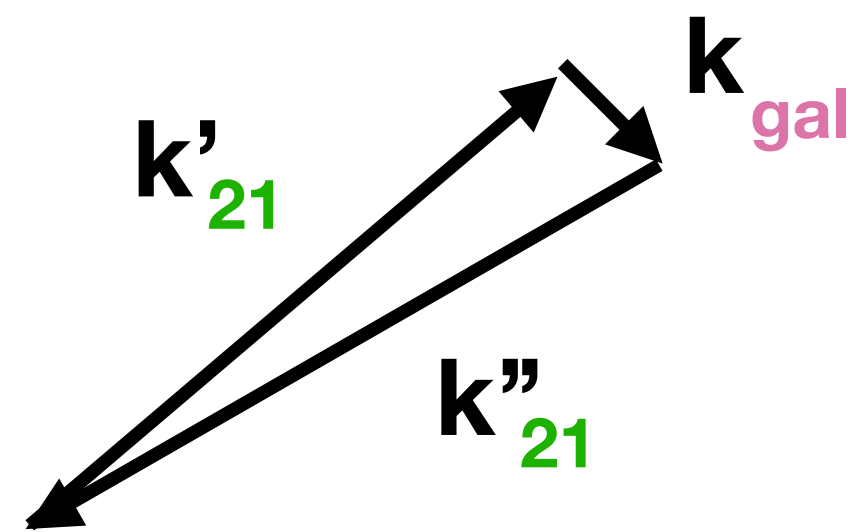
Direct 21 cm x galaxies signal vanishes due to foregrounds in long wavelength line-of-sight modes. Need to use higher order correlations.



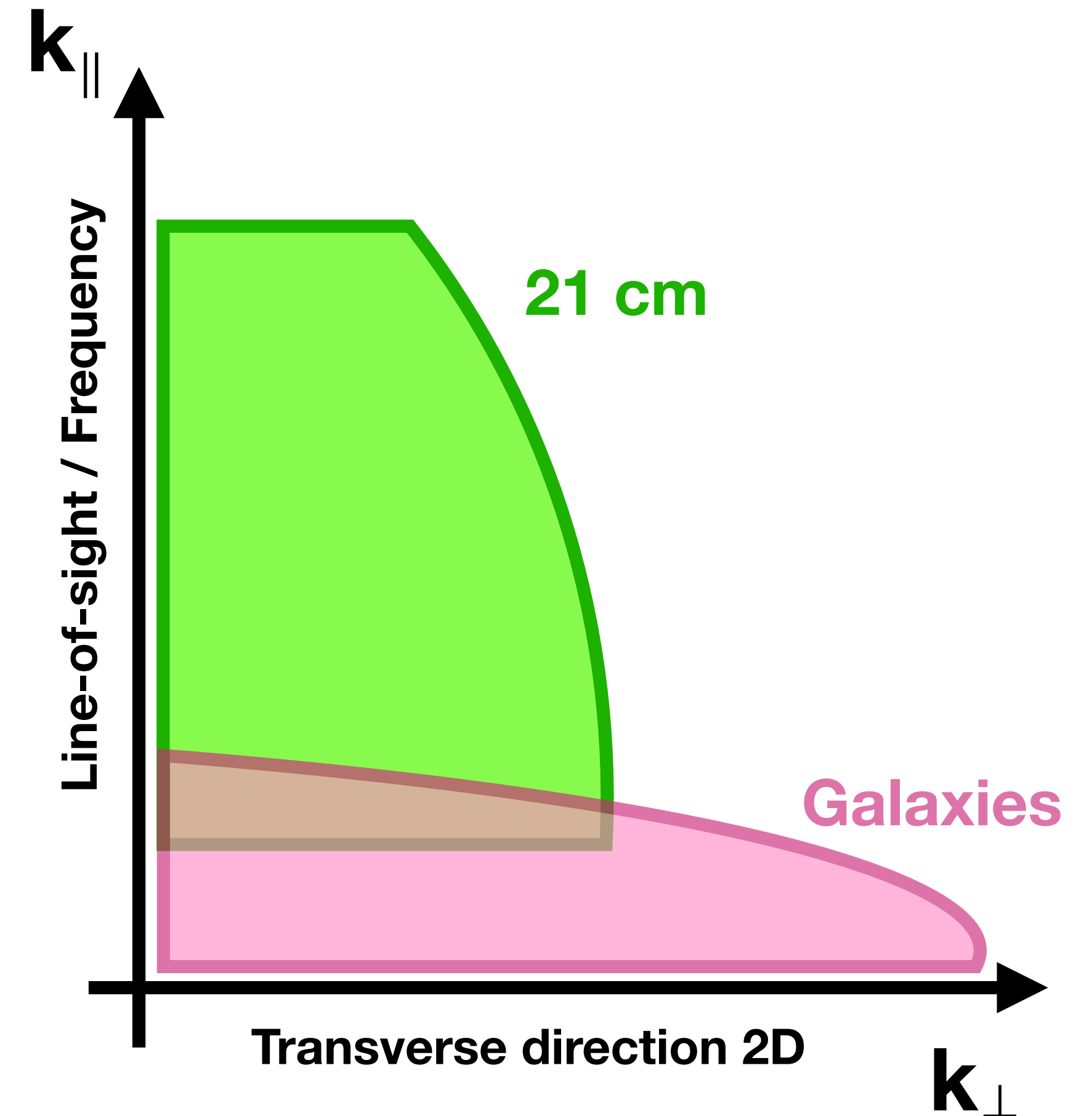
Beat systematics: cross correlations

Direct 21 cm 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.

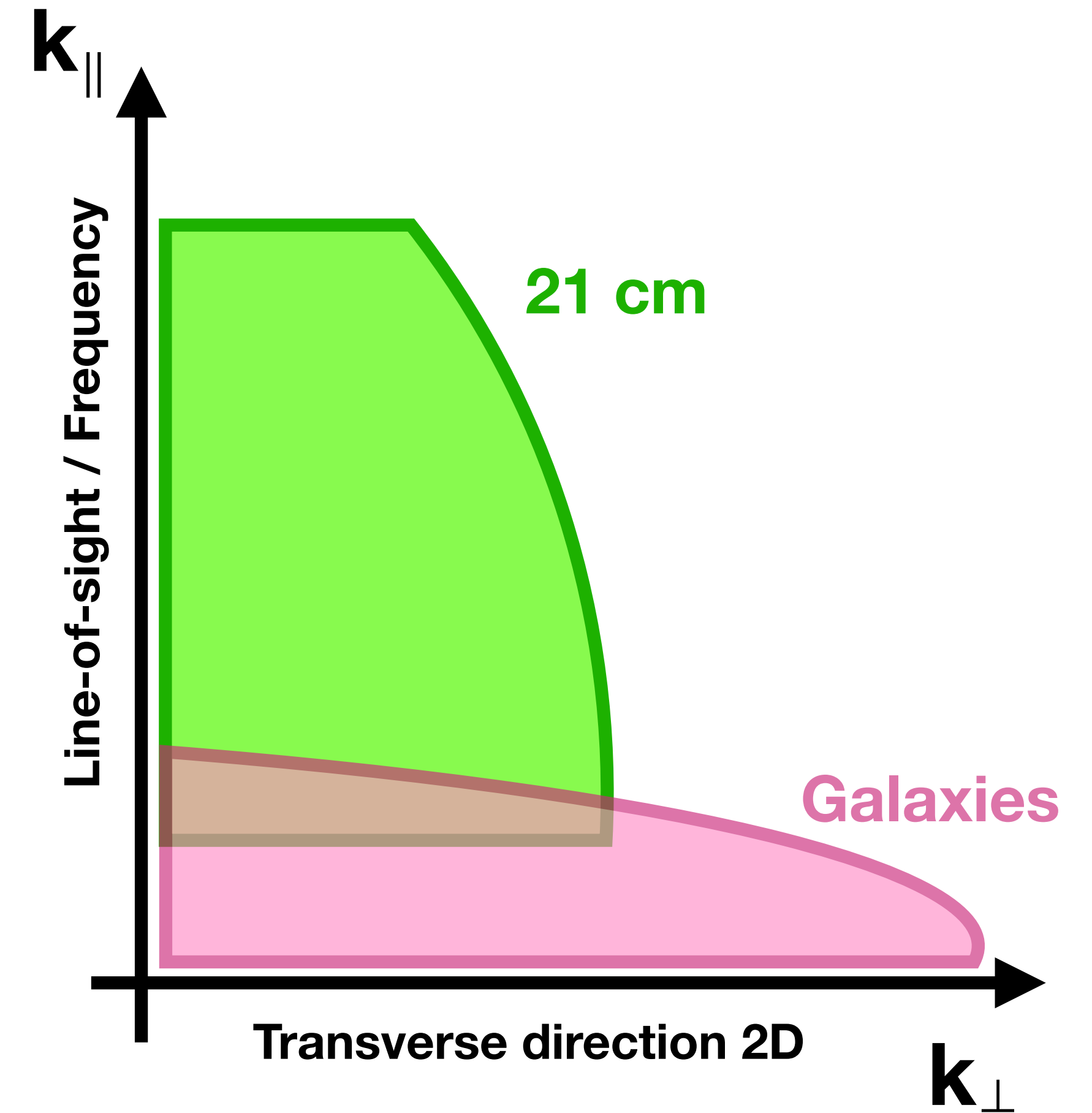
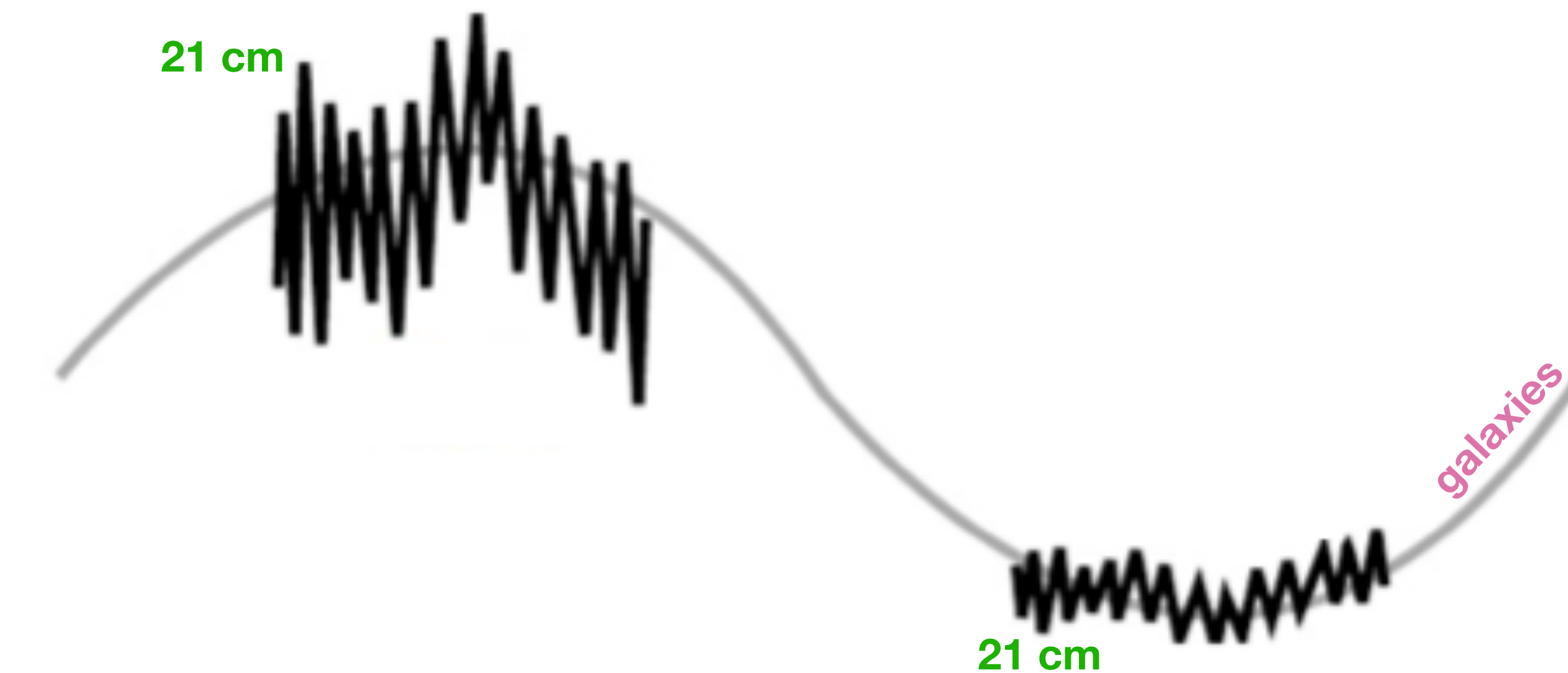


$$\langle \delta(\mathbf{k}) \delta(\mathbf{k}') \delta(\mathbf{k}'') \rangle = \delta_D(\mathbf{k} + \mathbf{k}' + \mathbf{k}'') B(\mathbf{k} + \mathbf{k}')$$



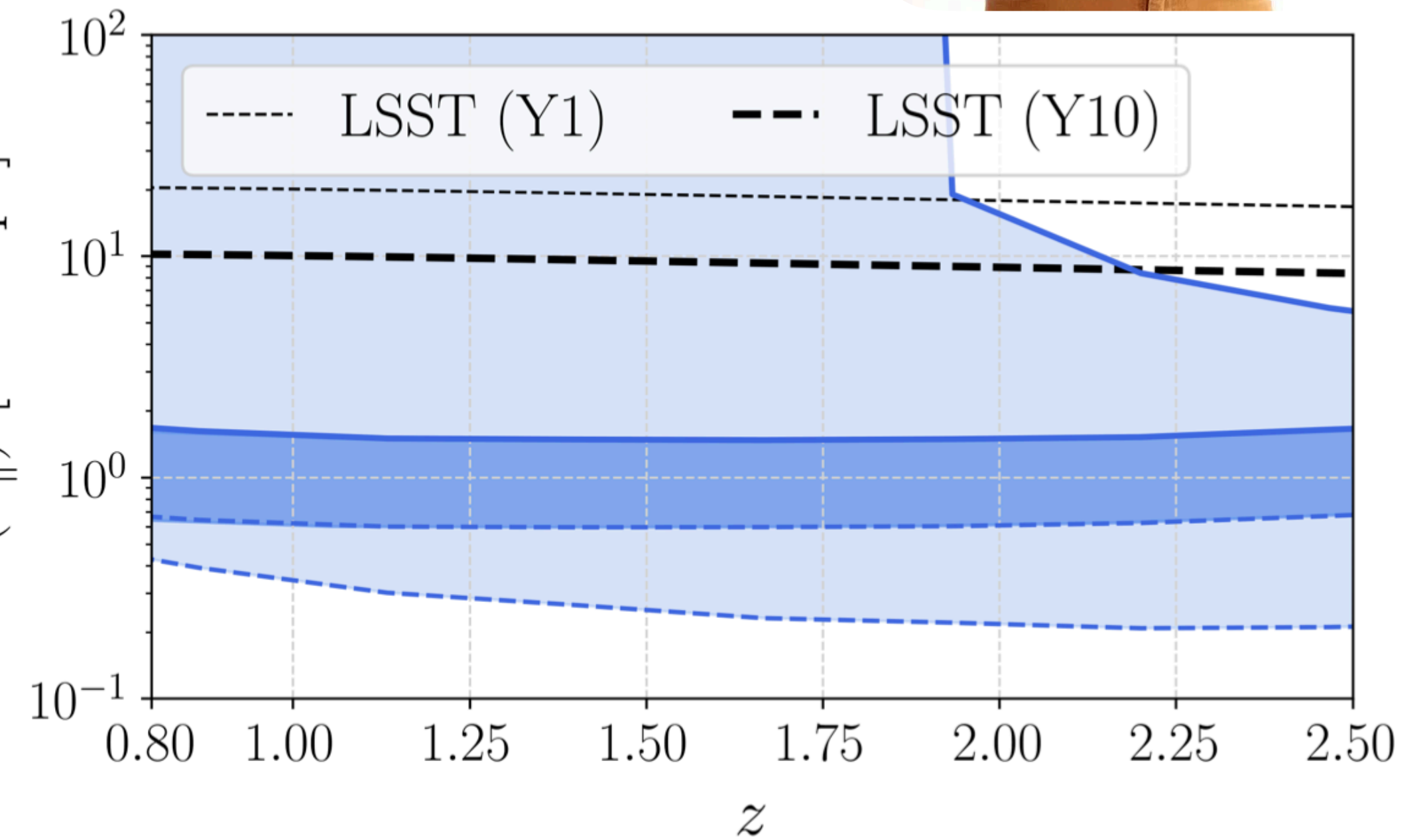
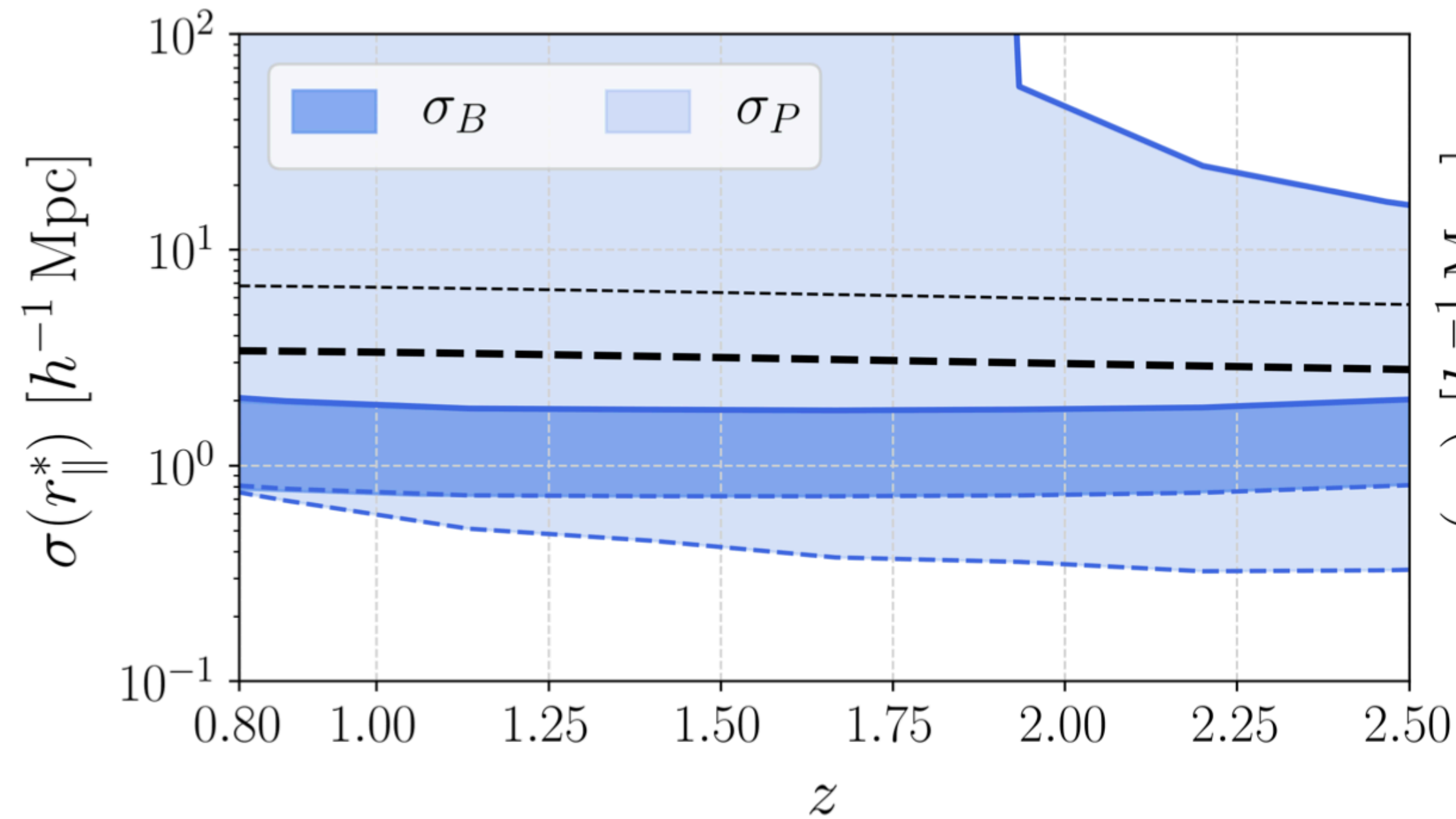
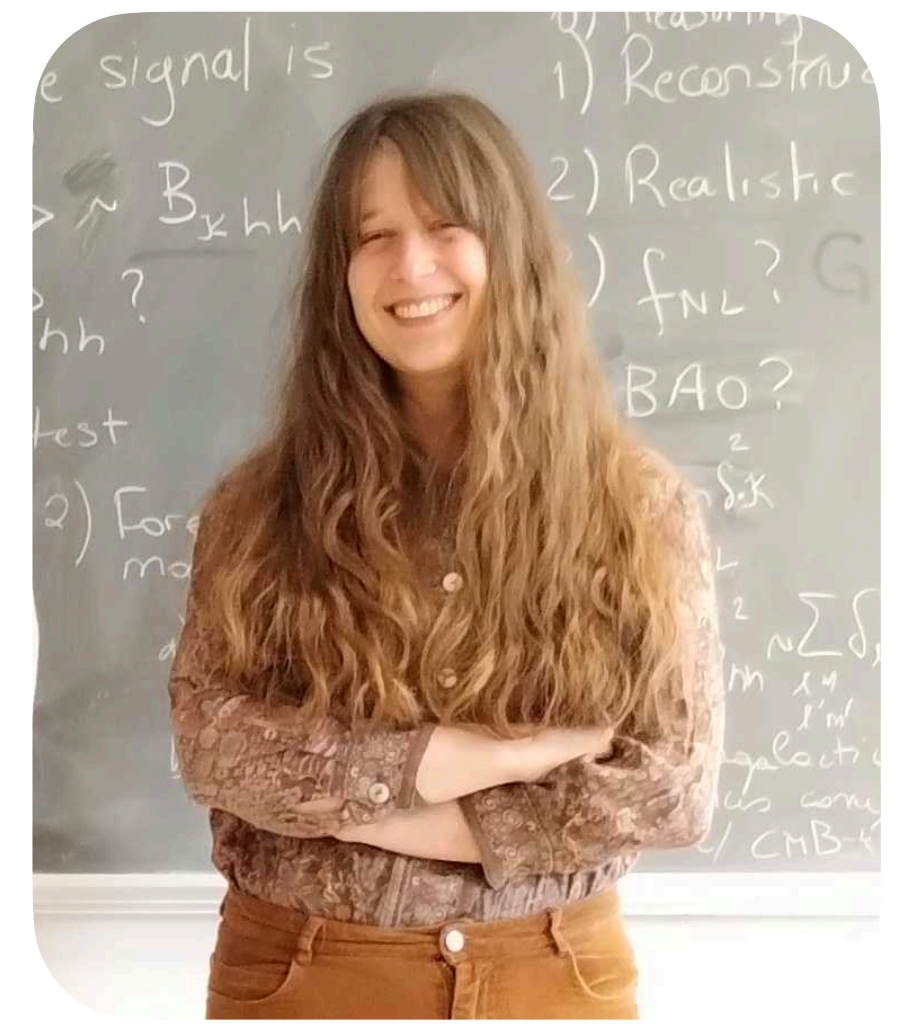
Beat systematics: cross correlations

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



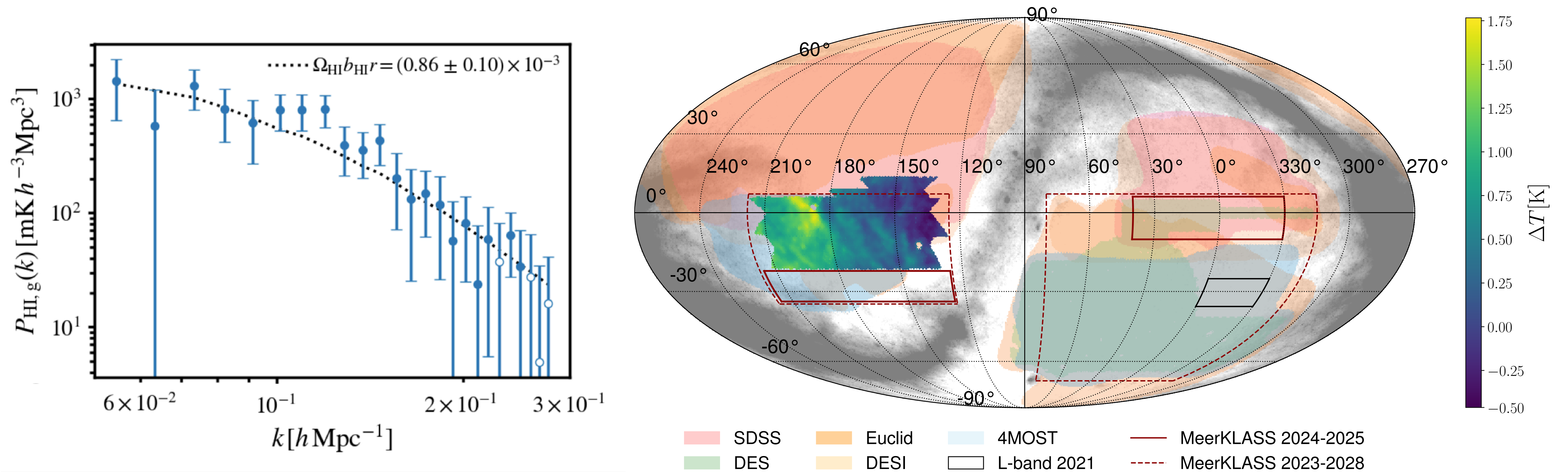
Beat systematics: cross correlations

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



Carol Guandalin+ 2022

- **Hydrogen Intensity Mapping (IM):** complementary, promising tracer
- **IM is hard... yet we are getting there!**



Credit: MeerKLASS coll.