Hydrogen intensity mapping with MeerKAT observations: **optimising the contaminant subtraction**

Isabella Paola Carucci

SKA Cosmology SWG Annual Meeting
16th January 2024

INAF
ISTITUTO NAZIONALE
DI ASTROFISICA

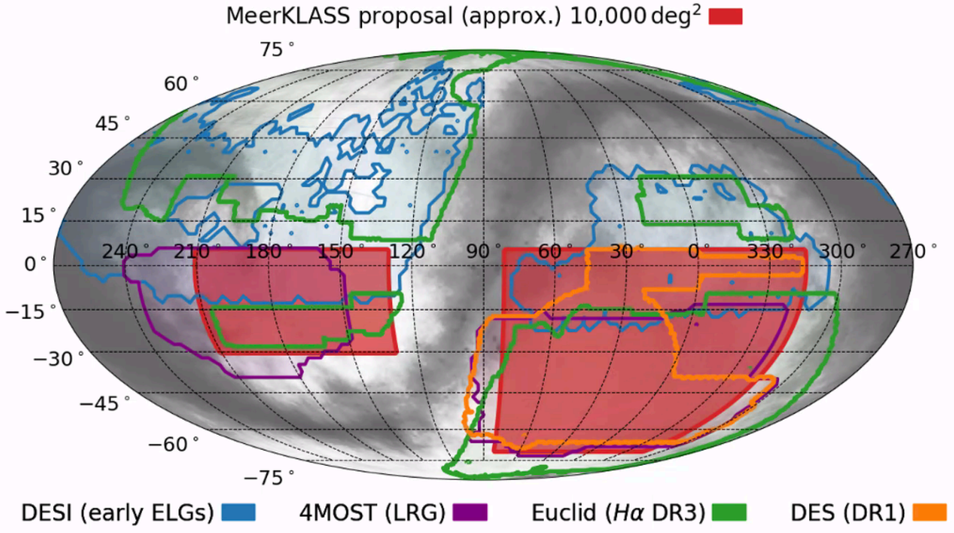
Funded by
the European Union
NextGenerationEU

Slide: Mário Santos from yesterday!

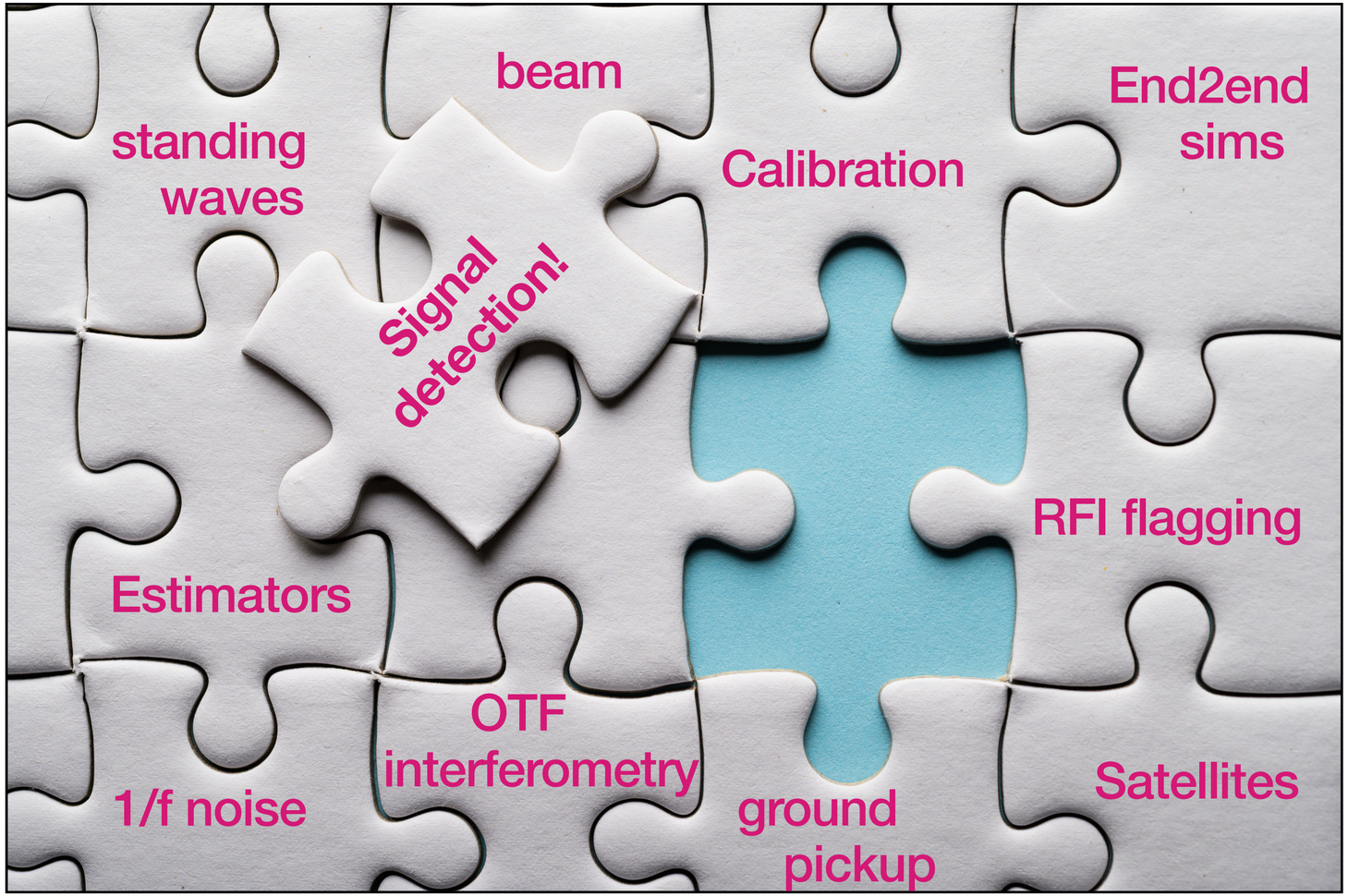
The present: an SKA cosmology survey precursor with MeerKAT

- ***MeerKLASS: MeerKAT Large Area Synoptic Survey*** (> 50 members - Santos et al., arXiv:1709.06099)
- Aim: Cosmology (HI intensity mapping) but commensal with lots of other science
- Use single dish data for cosmology and interferometer data for a continuum galaxy survey

- **L-band:**
  - 900-1670 MHz (z<0.58)
  - ~ 100 hours observed
  - MeerKLASS+ proposal submitted: 2,000 h over 5,000 deg2 (continuum: 9 uJy rms, 5'')
- **UHF band:**
  - 580 MHz-1015 MHz (0.40 < z < 1.45)
  - ~ 120 hours observed
  - Project "approved": 2,500 hours over 10,000 deg^2 (continuum: 25 uJy rms, 13'')



a collective effort



I am going to focus on the cleaning and report what has been (partially) going on within the Foreground&PowerSpectrum working Group of MeerKLASS:

Alkistis Pourtsidou, Jingying Wang, José Luis Bernal, Keith Grainge, Laura Wolz, Mario Santos, Marta Spinelli, Matilde Barberi Squarotti, Mel Irfan, Steve Cunnington, Zé Fonseca, …

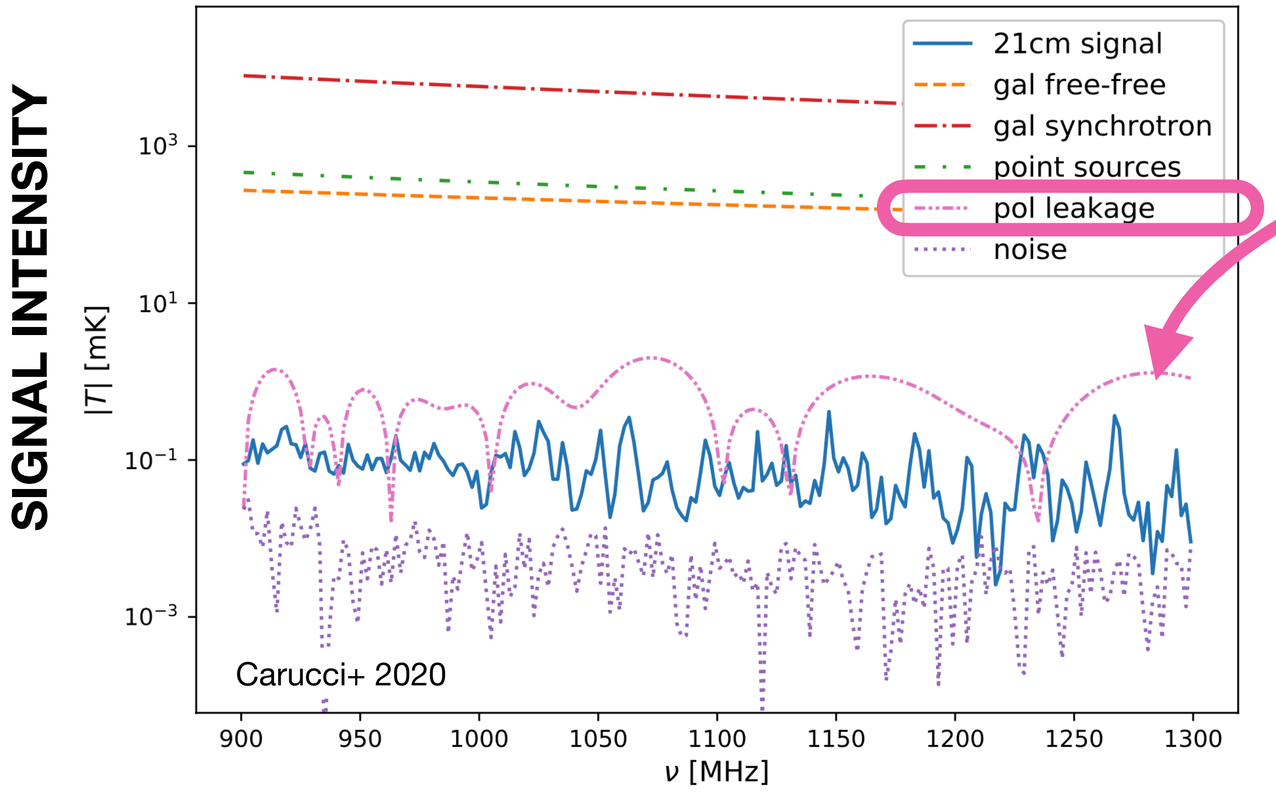
Outline

- It's not only "foregrounds": very short overview of the contaminant subtraction problem in HI IM
- Blind Source Separation methods as cleaning strategy
- Testing the methods with data to reproduce our detection in X-corr with galaxies

ongoing work

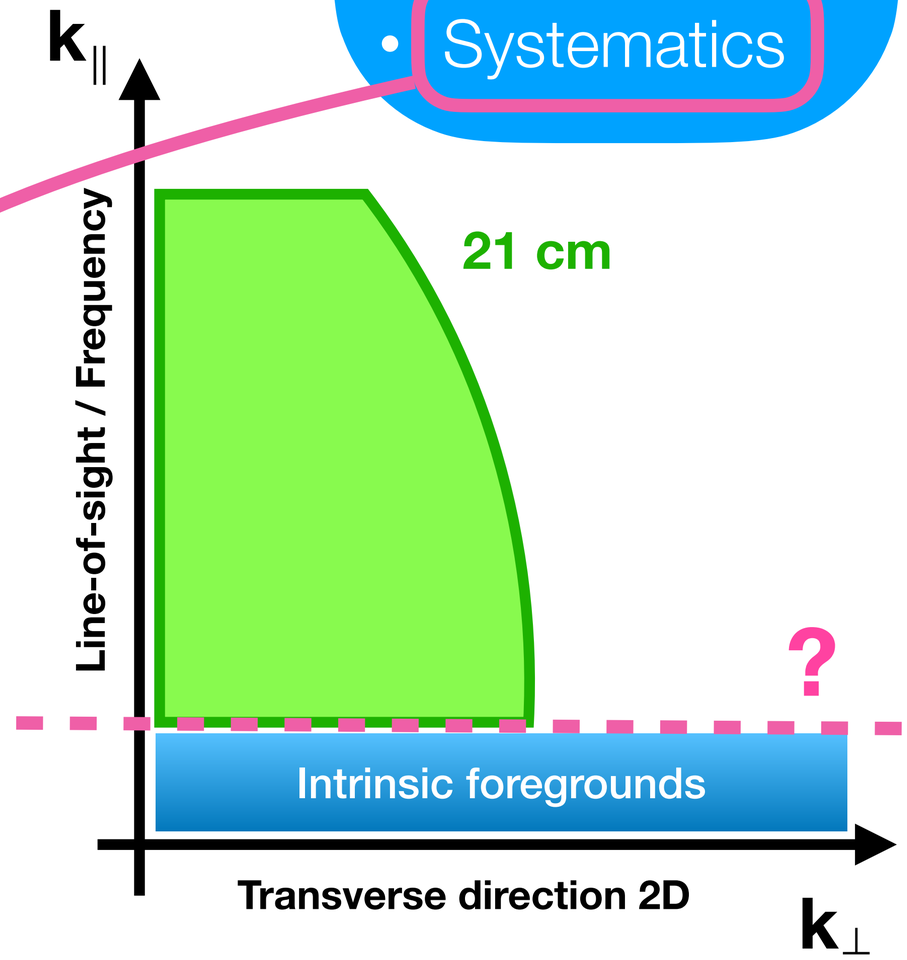
HI intensity mapping

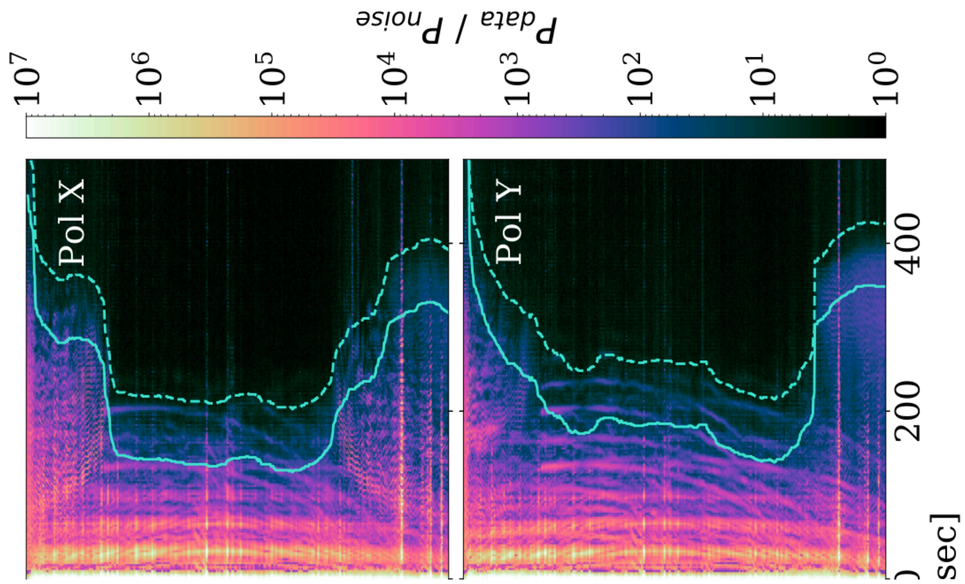
buried under the contaminants



CHALLENGES:

- Foregrounds
- Systematics

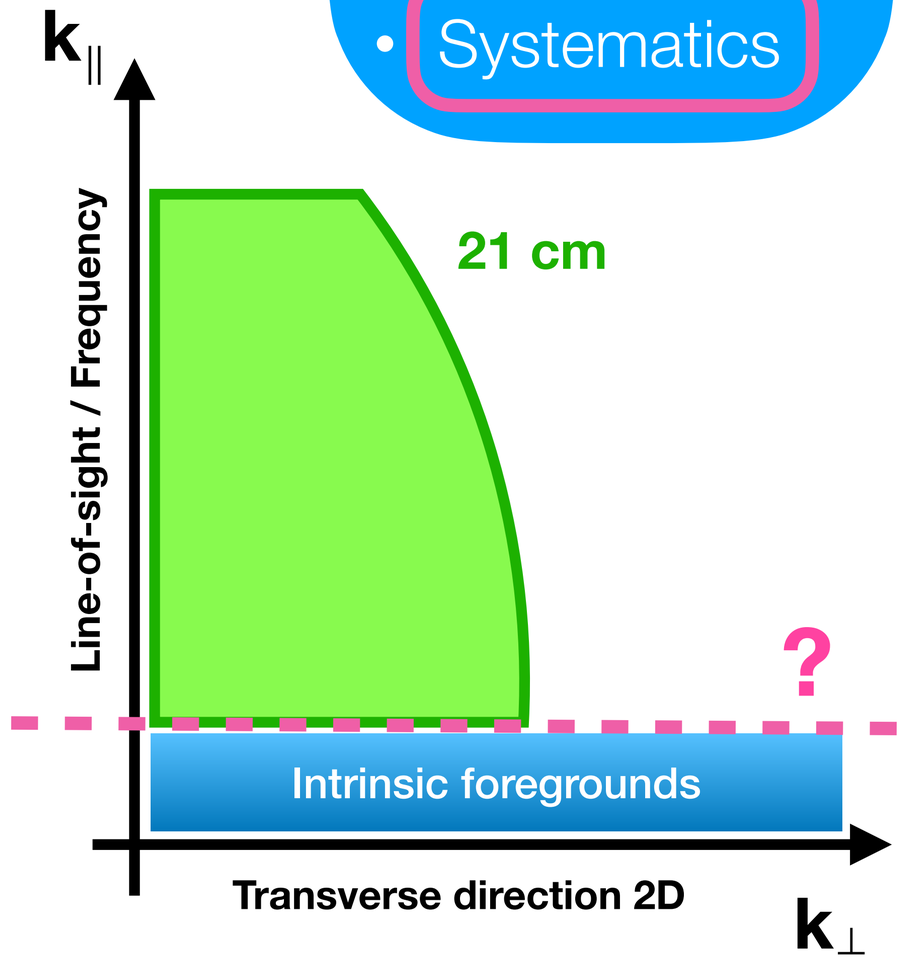




CHIME collab. 2022

CHALLENGES:

- Foregrounds
- Systematics



Filtering or ‘avoiding’ the contaminants is not an option

(Unless you are Aishrila, Sourabh, Mario, Laura, Zhaoting and work with Mightee data)

Filtering or ‘avoiding’ the contaminants is not an option for cosmo science

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!

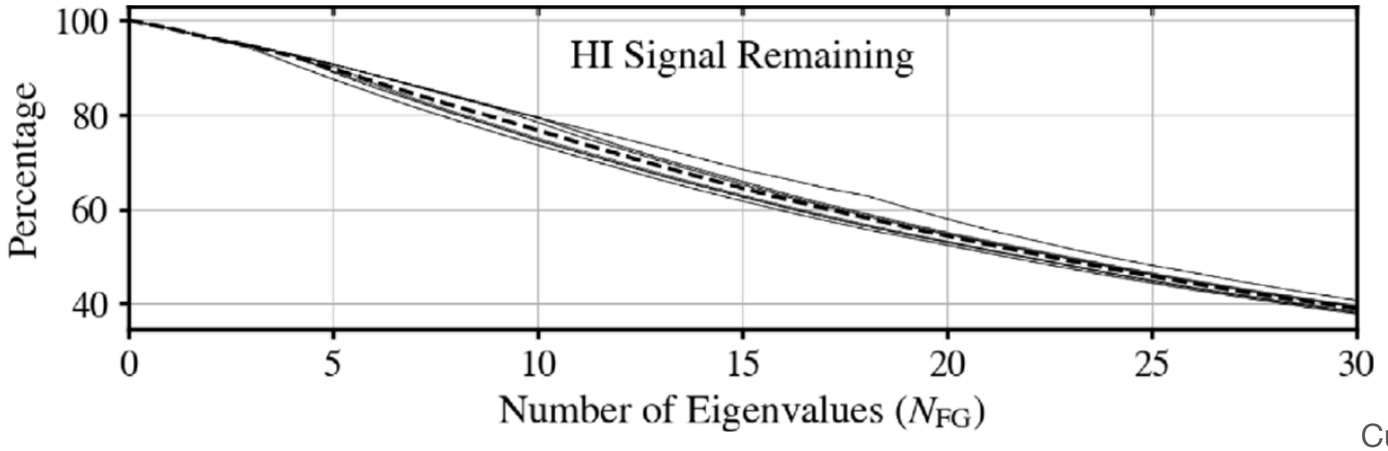
mixing matrix (f,n)

HI signal!

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

signal (f,p)

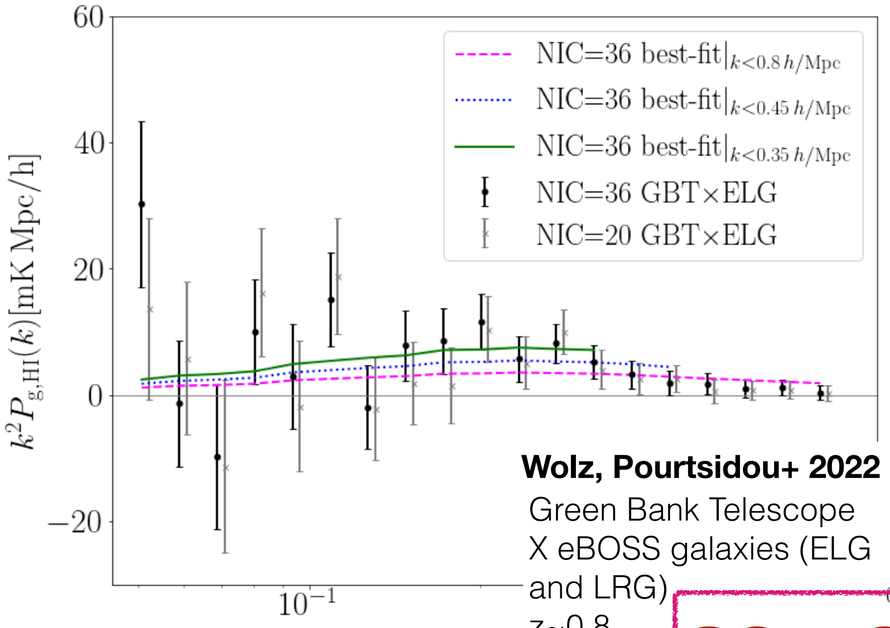
sources (n,p)



Cunnington+ 2021

Blind Source Separation algorithms

State-of-the-art

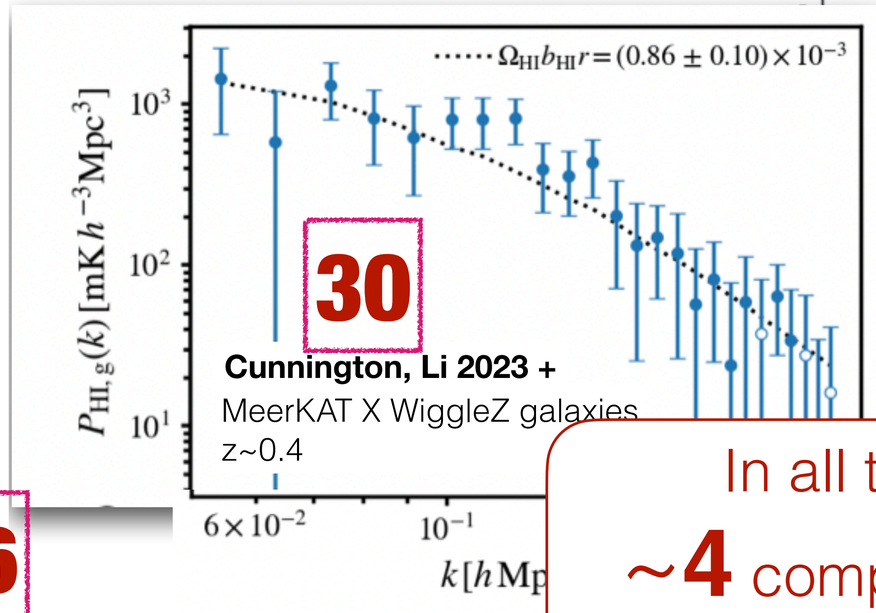


Wolz, Pourtsidou+ 2022
Green Bank Telescope
X eBOSS galaxies (ELG and LRG)
z~0.8

20 - 36

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

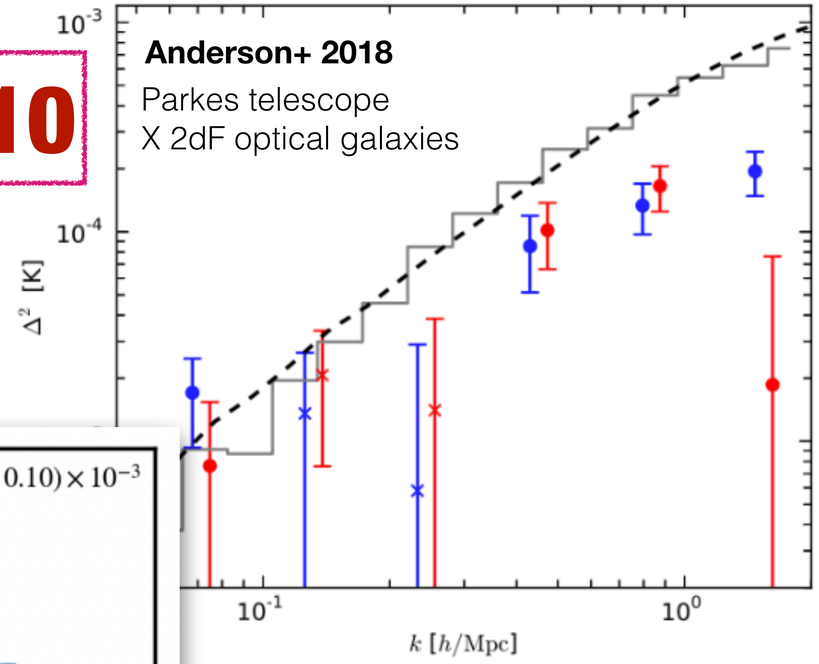
10 - 20



30

Cunnington, Li 2023 +
MeerKAT X WiggleZ galaxies
z~0.4

10



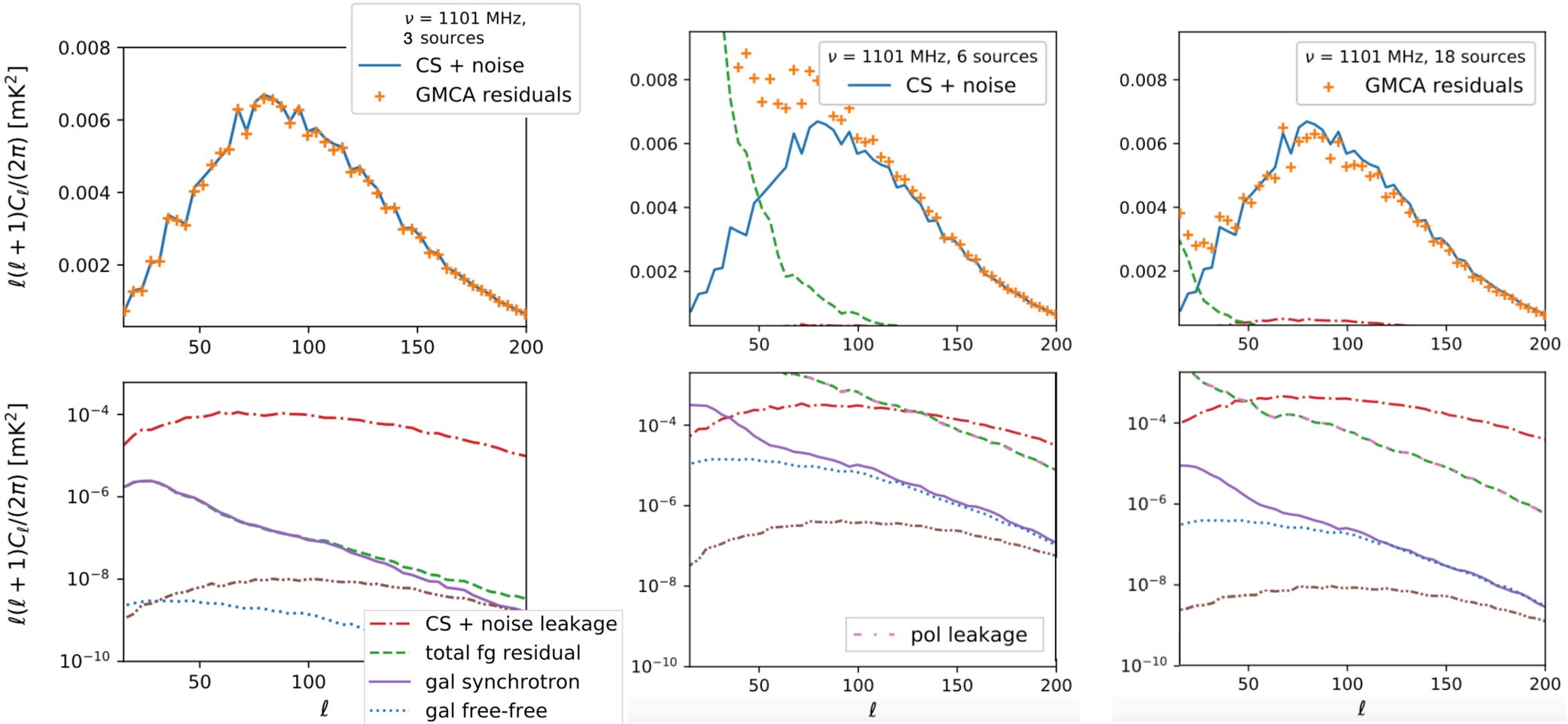
Anderson+ 2018
Parkes telescope
X 2dF optical galaxies

In all theoretical works:

~4 components removed are enough

(e.g., Wolz+ 2014, Alonso+ 2015, Cunnington+ 2019, …)

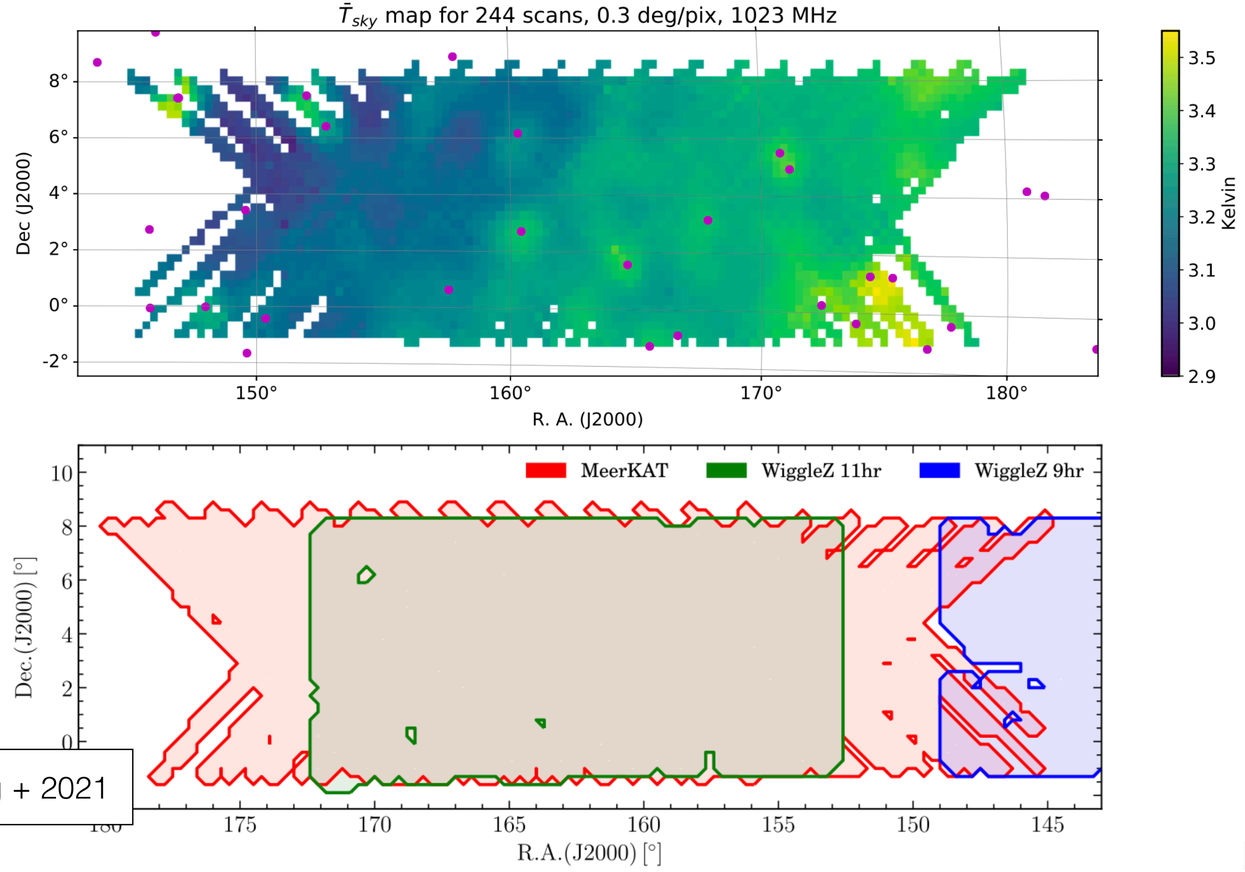
Different scales need different care



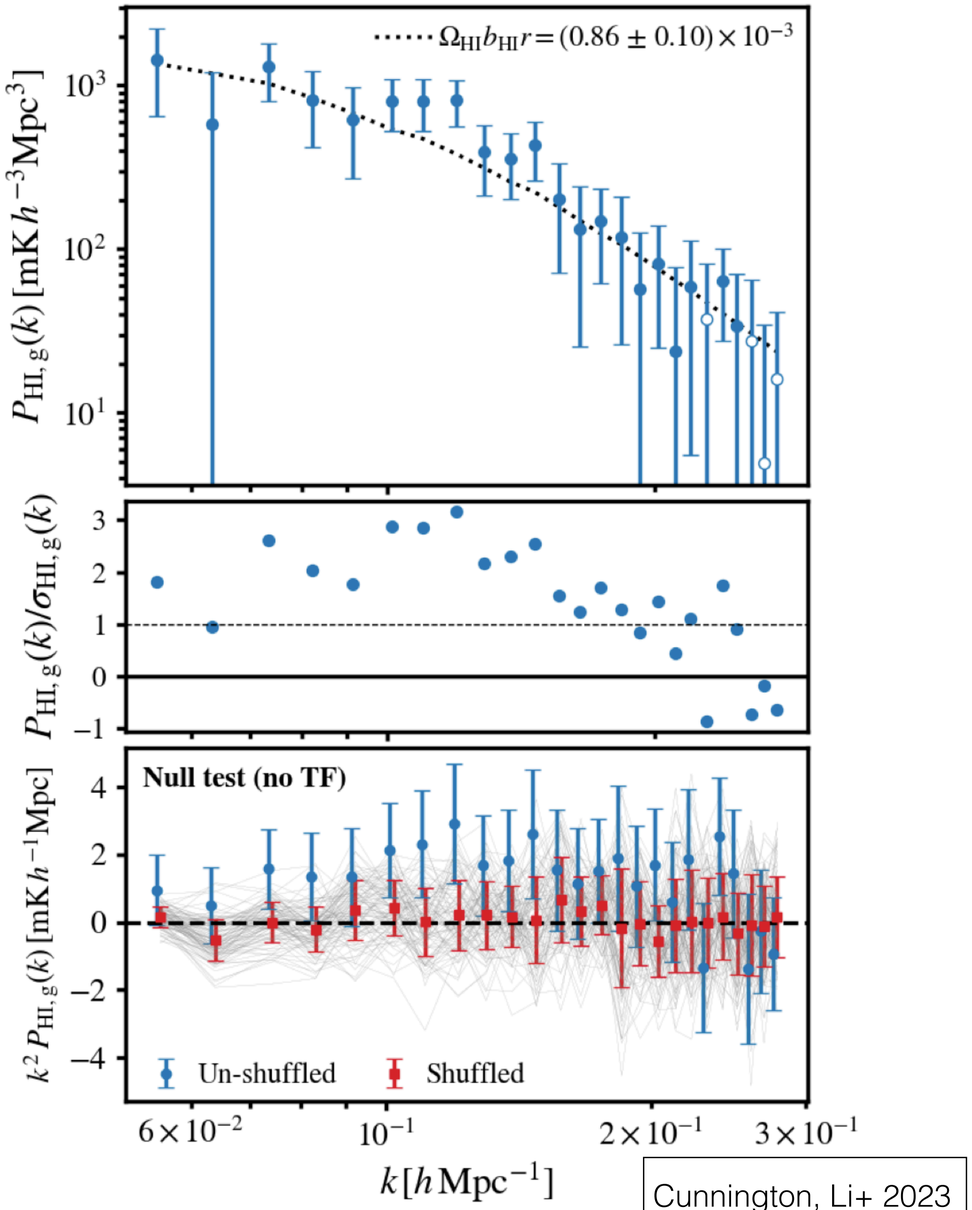
Carucci + 2020

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)

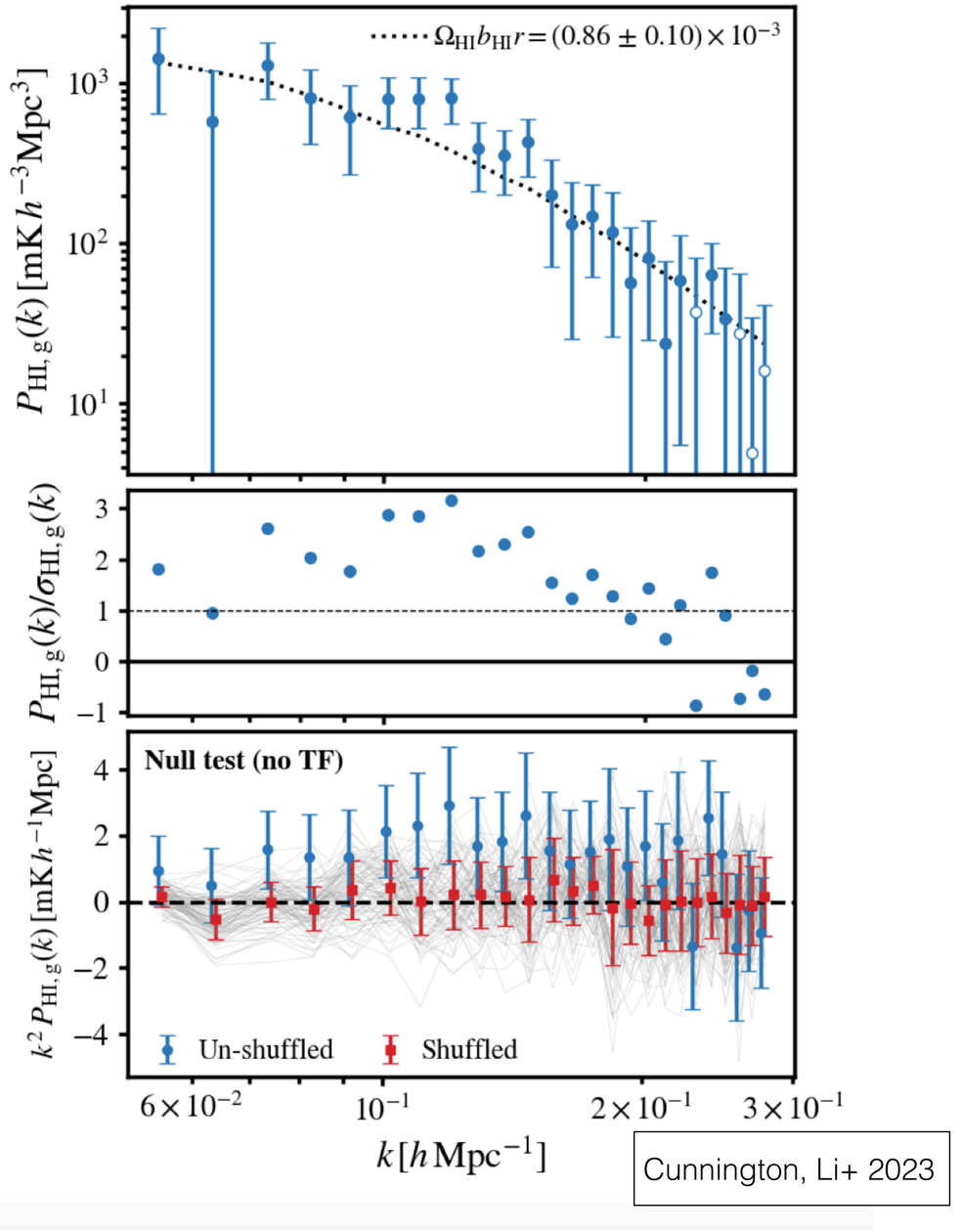


JY Wang + 2021



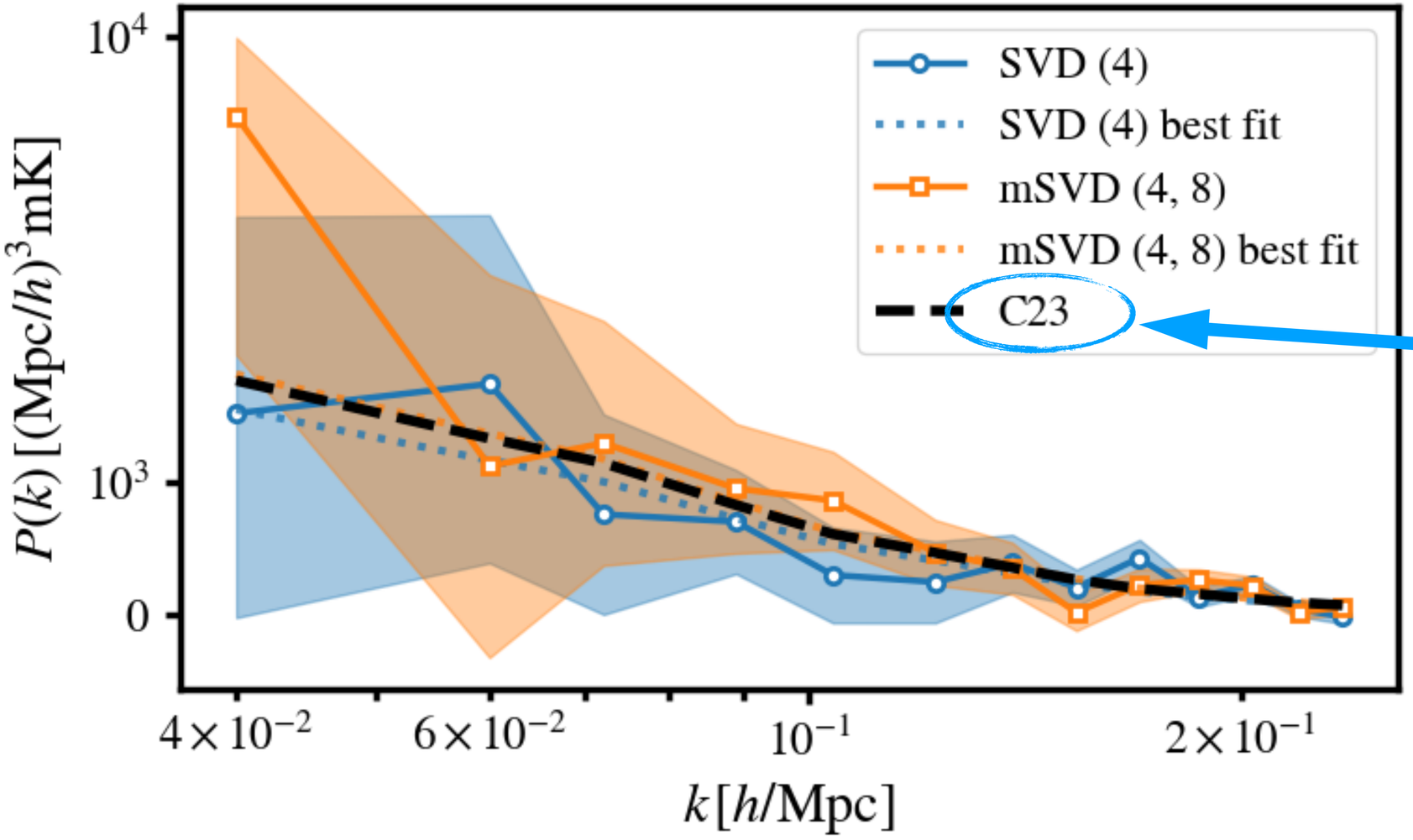
Cunnington, Li+ 2023

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



Cunnington, Li+ 2023

Preliminary



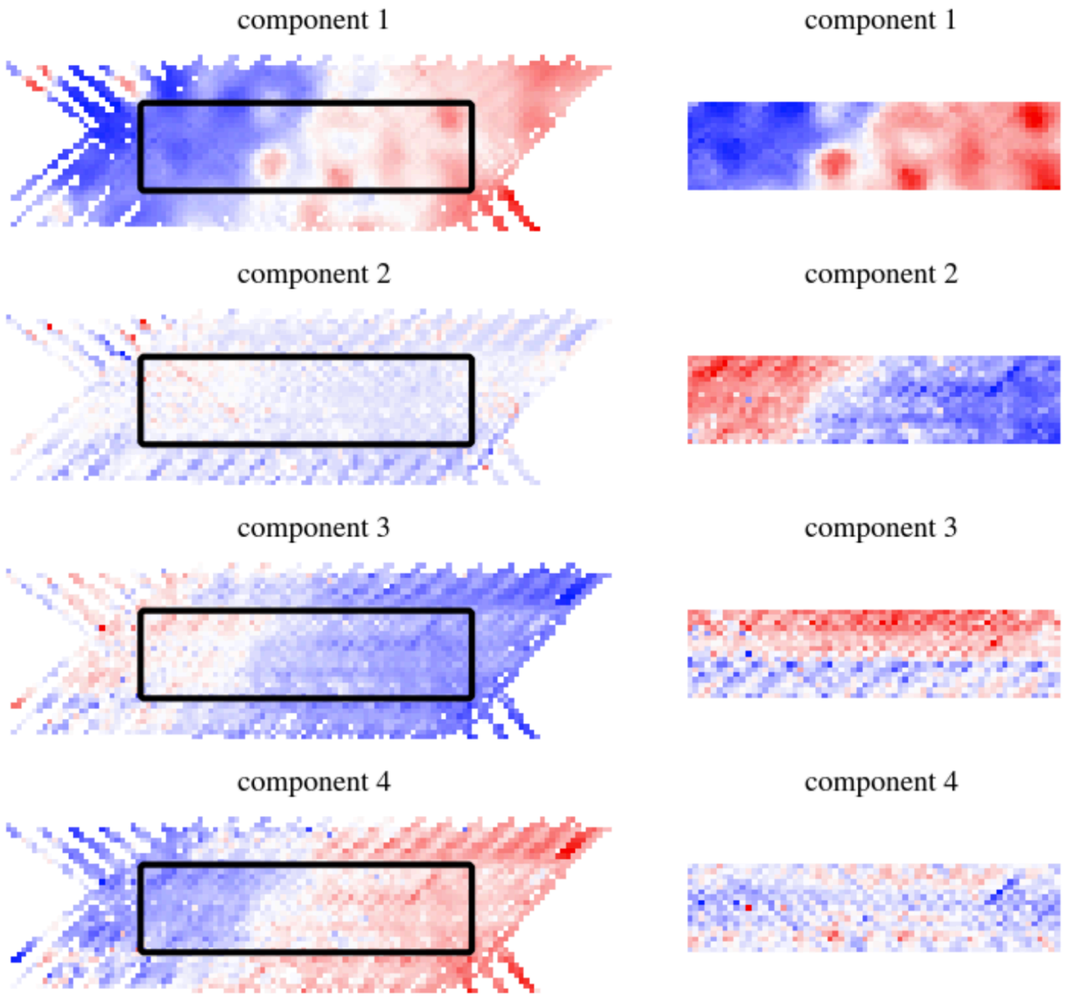
30 modes removed + Transfer Function correction

Re-analyis of 2019 data

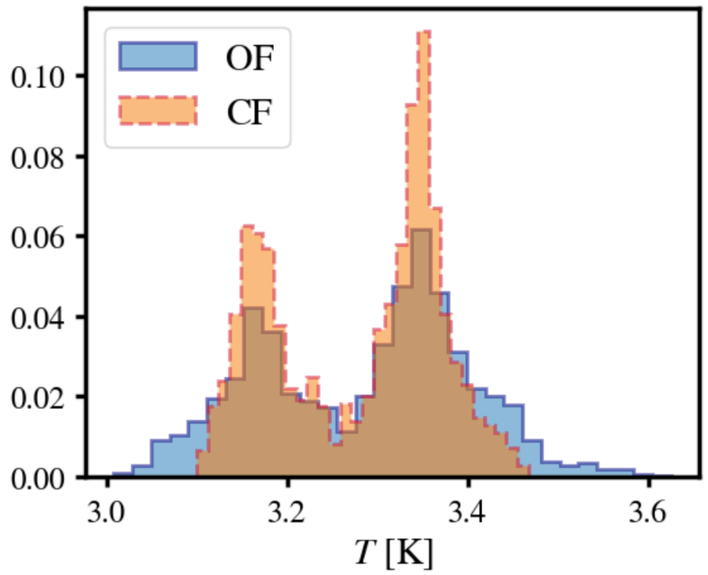
$$X = A\,S$$

Preliminary

1. PCA-informed pixel flagging



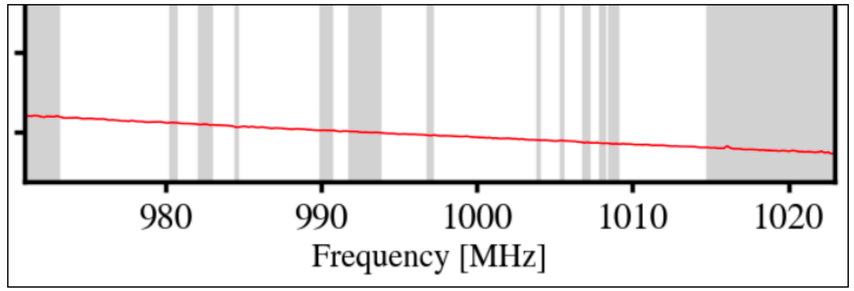
original data cube



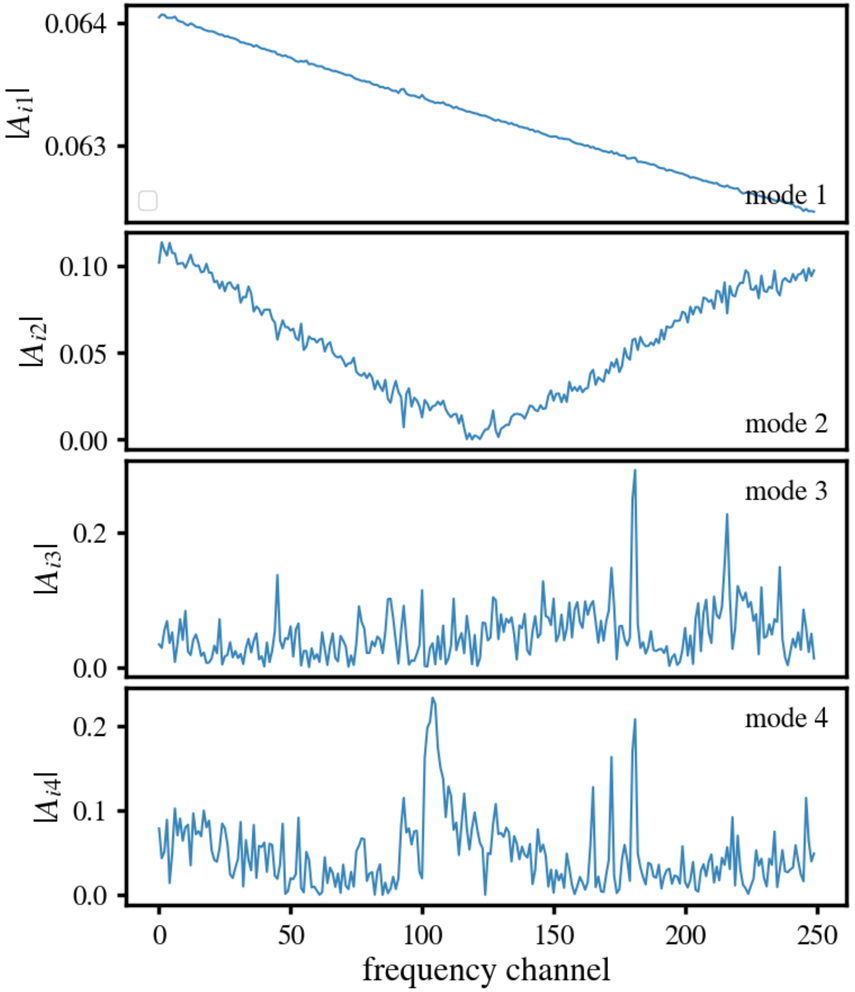
Re-analyis of 2019 data

1. PCA-informed pixel flagging
2. Keep *bad* channels

see also discussion in Carucci+ 2020



$$\mathbf{X} = \mathbf{A}\ \mathbf{S}$$



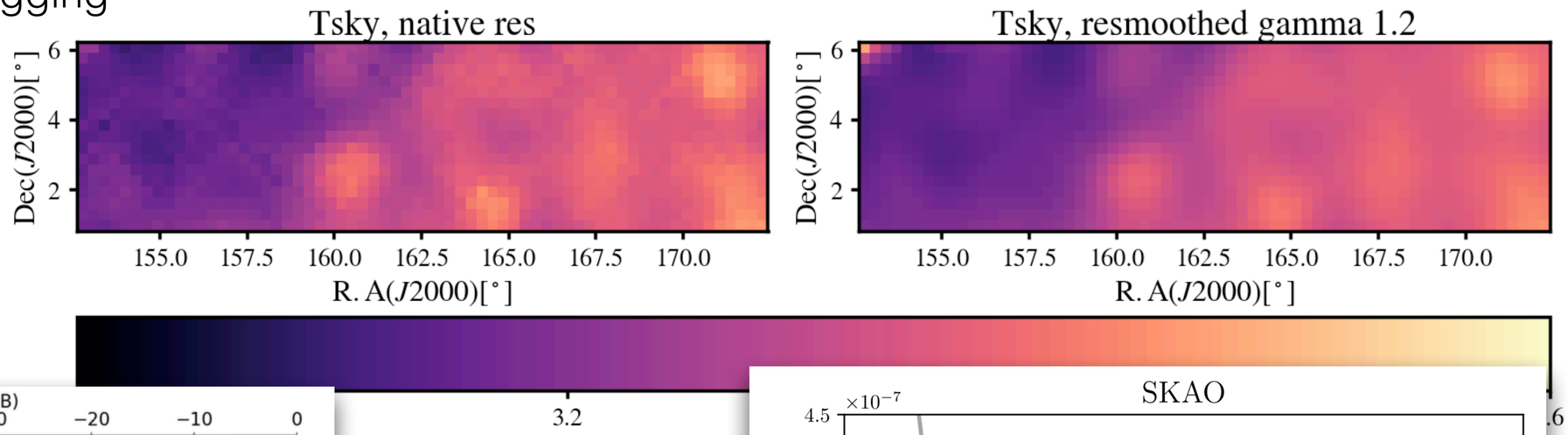
Preliminary

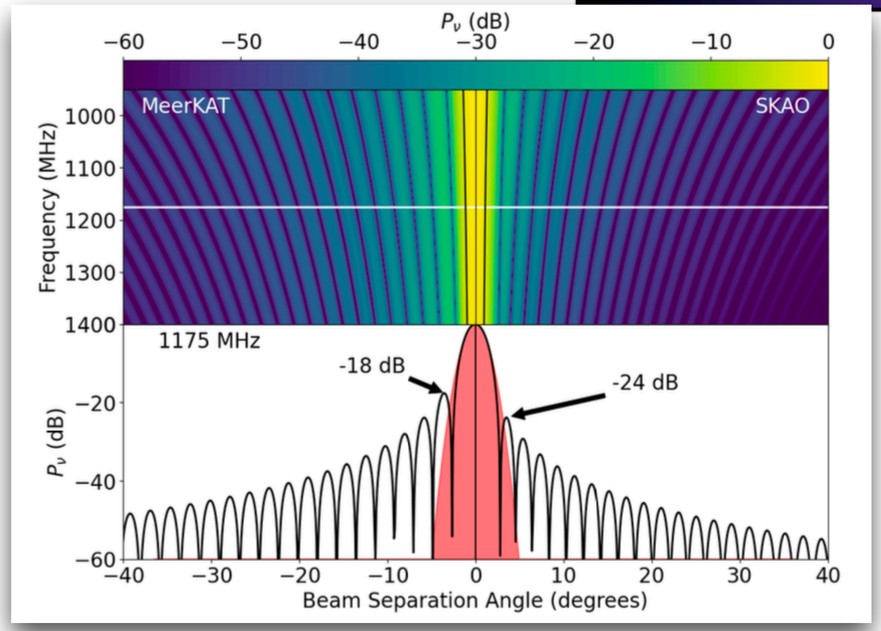
See **Irfan+ 2022** for discussion on how to relate that first mode to the galactic synchrotron

Re-analyis of 2019 data

Preliminary

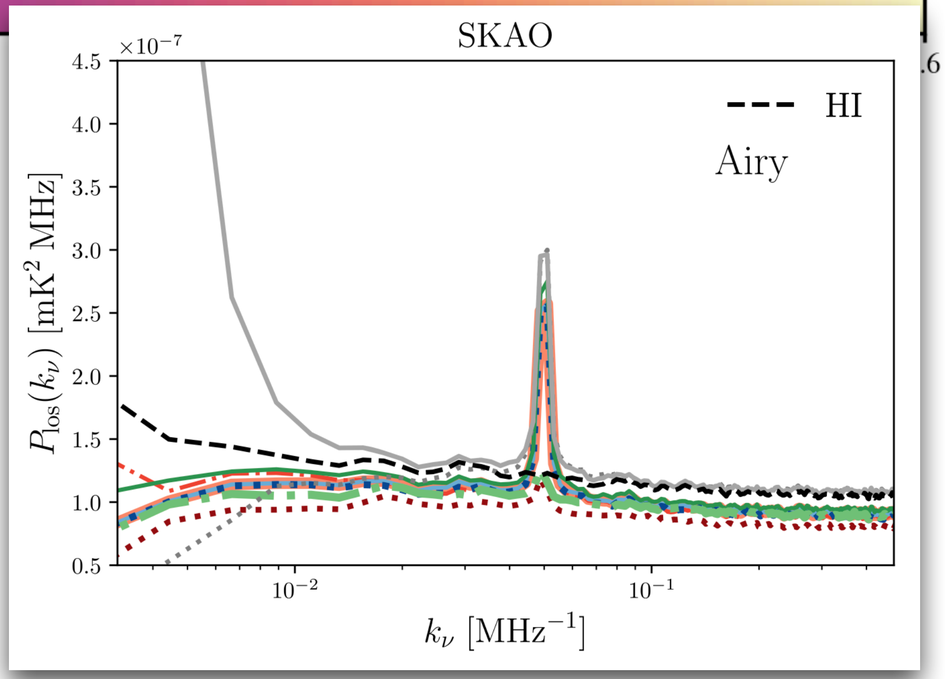
1. PCA-informed pixel flagging
2. Keep *bad* channels
3. No re-smoothing





Blind foreground subtraction challenge

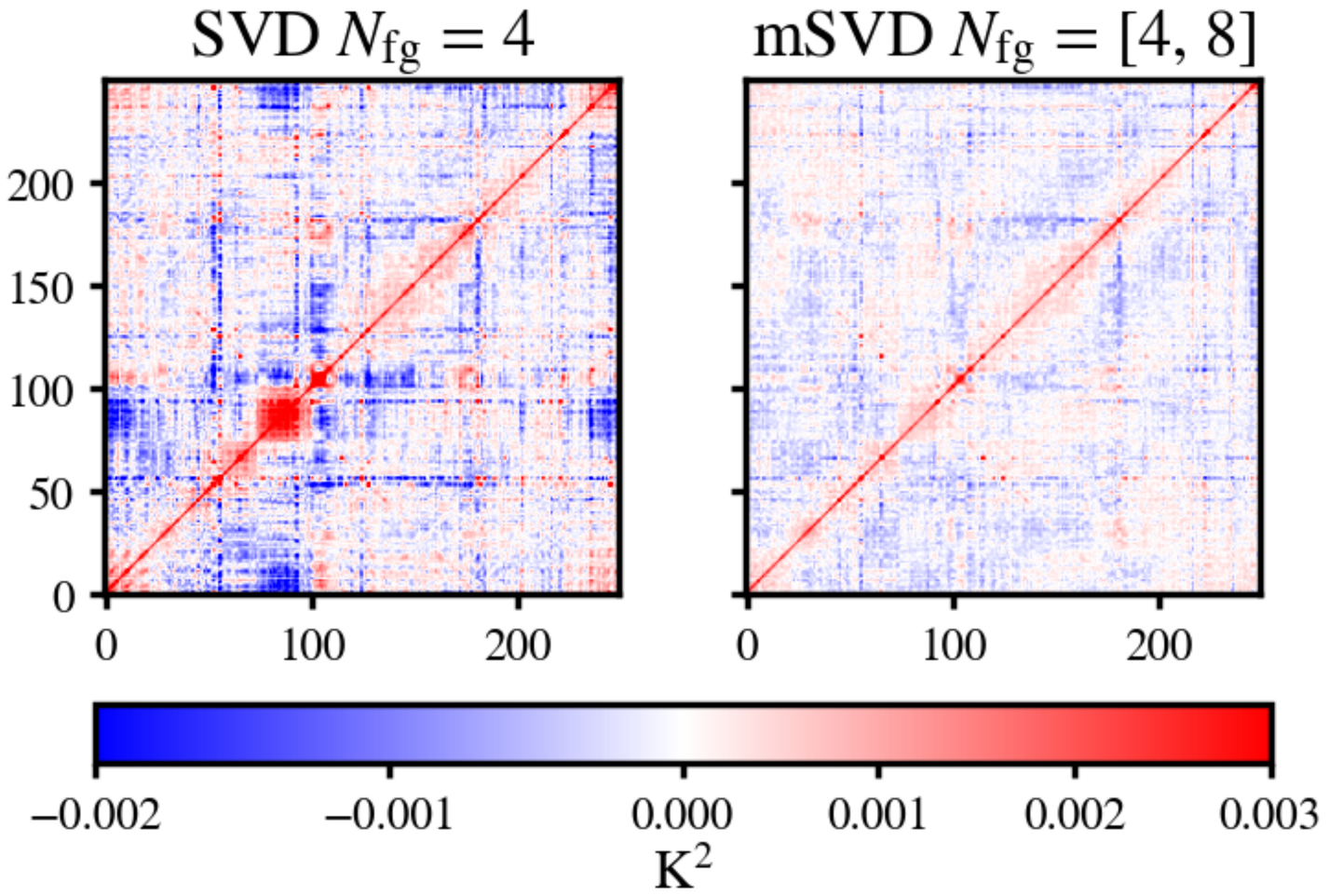
Spinelli+ 2022



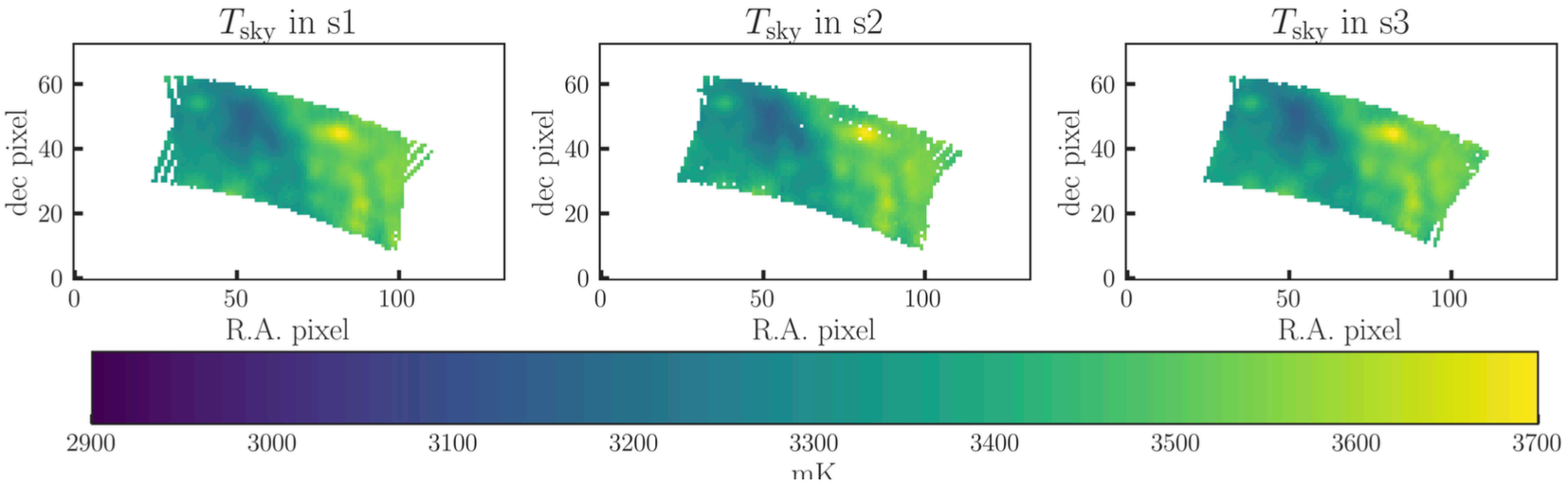
Re-analyis of 2019 data

Preliminary

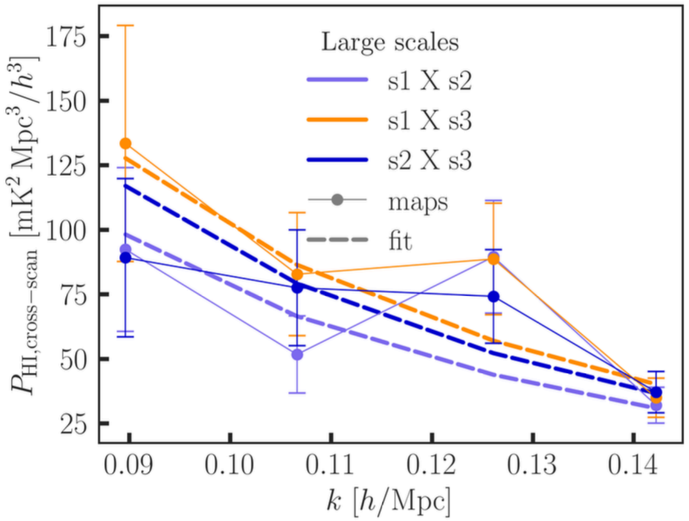
1. PCA-informed pixel flagging
2. Keep *bad* channels
3. No re-smoothing
4. Going multiscale



Very Preliminary



Work led by Matilde Barberi Squarotti



Summary

- We are detecting (again!) the cross signal with WZ galaxies to test different pre-processing steps and cleaning algorithms
- We did learn things!
- PCA /SVD is still our best friend

- Separating scales for the cleaning is more efficient at reducing the cube variance (multiscale cleaning)
- Even a PCA/SVD run should not be applied as a black-box

Getting ready for the SKAO HI IM science