21cm Intensity Mapping: opportunities and challenges on the road to the SKA Observatory

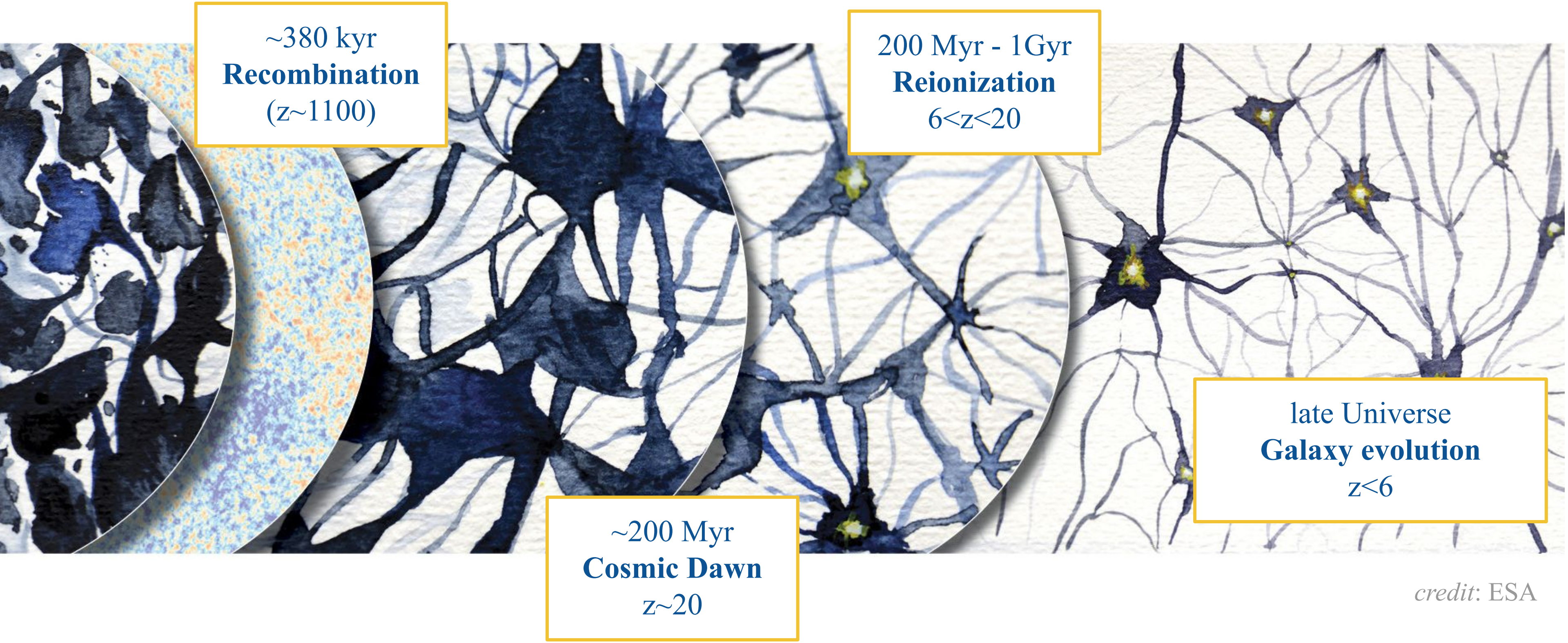
Marta Spinelli

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

RESPICE PROSPICE
UNIVERSITY of the WESTERN CAPE

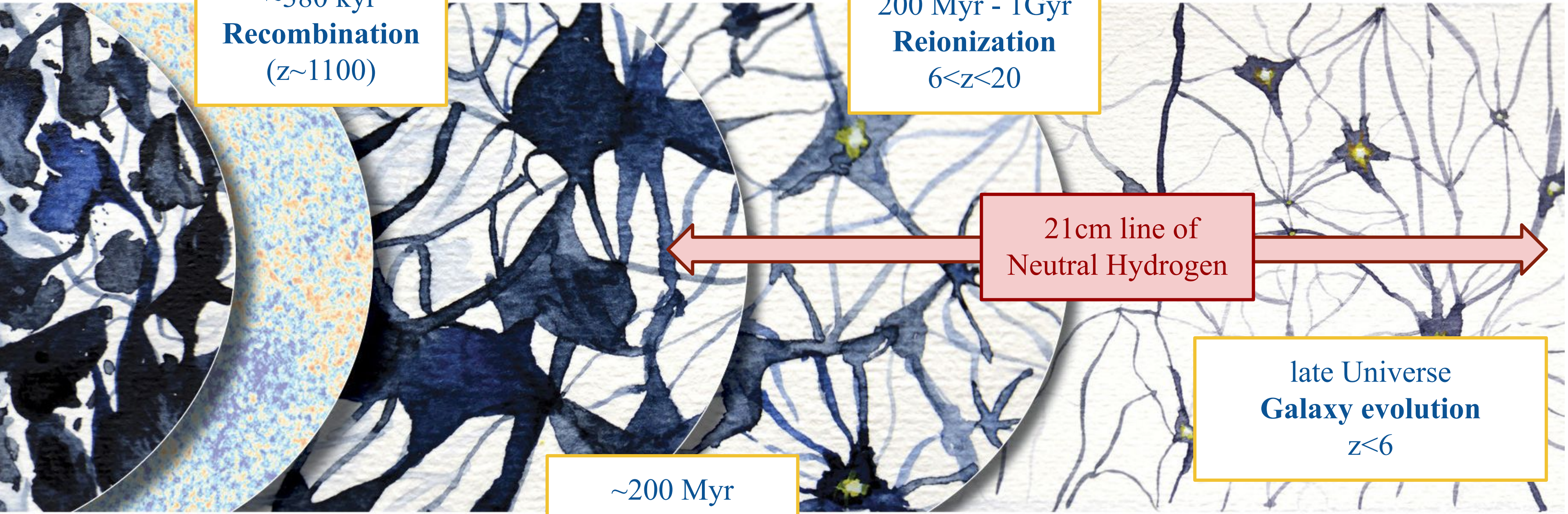
ISTITUTO NAZIONALE DI ASTROFISICA
INAF

Hydrogen through cosmic time



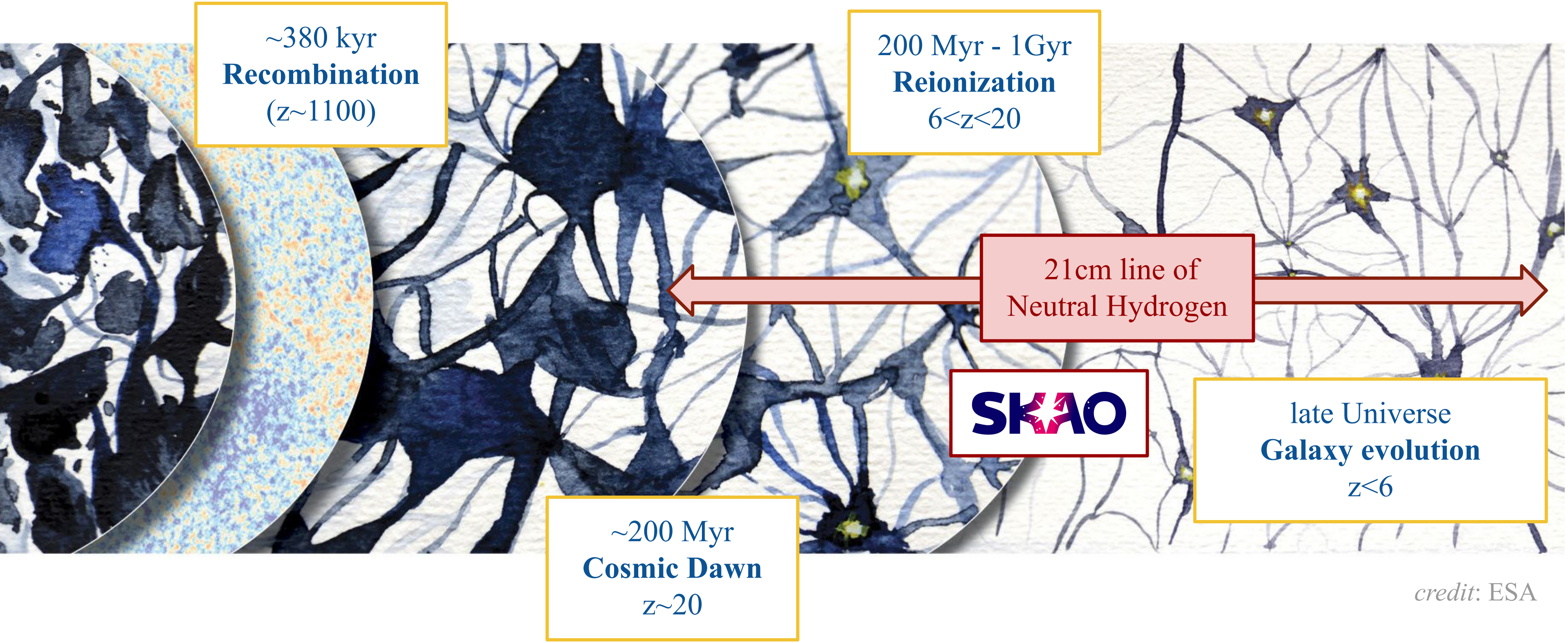
credit: ESA

Hydrogen through cosmic time



credit: ESA

Hydrogen through cosmic time

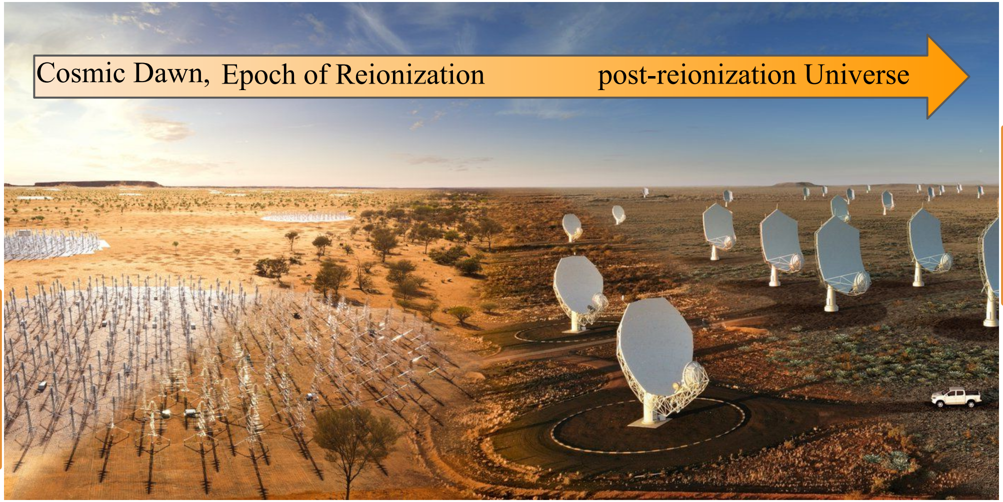


credit: ESA

21cm Cosmology

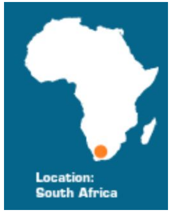
- signal *redshifted* due to the expansion of the Universe to **Radio Frequencies**
- **SKA Observatory**: cover **all the relevant frequencies** with unprecedented sensitivity

SKA-Low
50 MHz - 350 MHz
30>z>3



credit: skatelescope.org

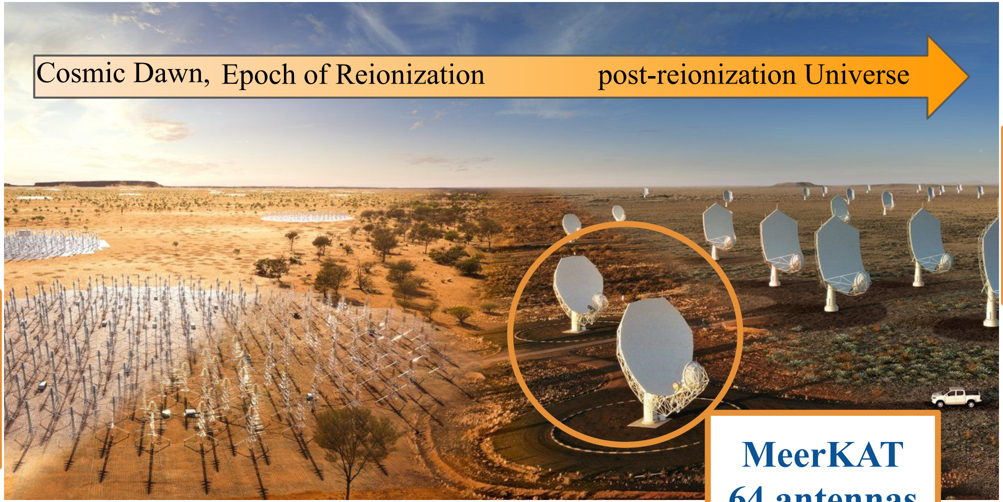
SKA-Mid
350 MHz - 13.5 GHz
3>z>0



21cm Cosmology

- signal *redshifted* due to the expansion of the Universe to **Radio Frequencies**
- **SKA Observatory**: cover **all the relevant frequencies** with unprecedented sensitivity

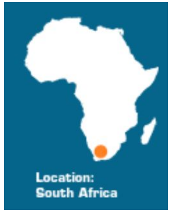
SKA-Low
50 MHz - 350 MHz
30>z>3



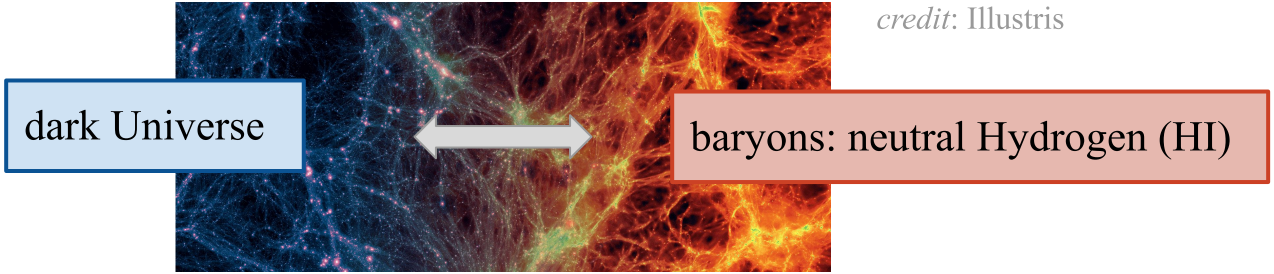
credit: skatelescope.org

SKA-Mid
350 MHz - 13.5 GHz
3>z>0

MeerKAT
64 antennas
1.5>z>0



Mapping neutral hydrogen



credit: Illustris

What is the nature of dark matter and dark energy?

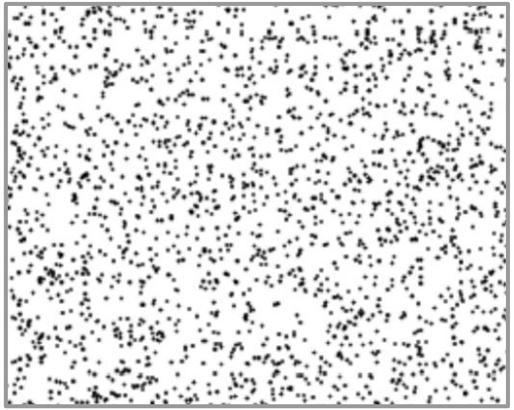
- ❑ how is dark matter distributed on large scales?
- ❑ how does its distribution evolve with cosmic time?
- ❑ what is the role of dark energy?

How do baryons trace dark matter?

- ❑ what is the link between galaxies and dark matter halos?
- ❑ how are HI galaxies distributed in the cosmic web?
- ❑ how does the total cosmic HI evolve with redshift?

Intensity Mapping

credit: A. Pourtsidou

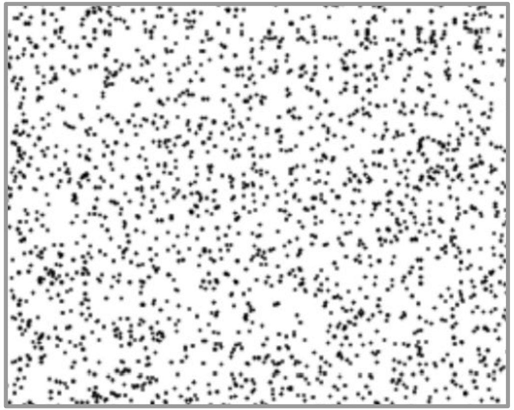


The distribution of **neutral Hydrogen** is a biased tracer of the **matter clustering** *similar to galaxy surveys*

In cosmology, **large scales** are fundamental

Intensity Mapping

credit: A. Pourtsidou



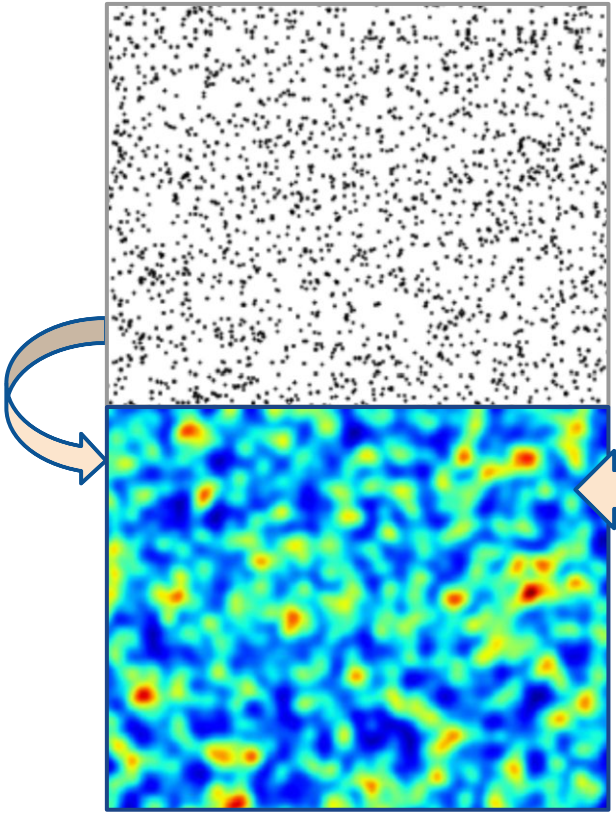
The distribution of **neutral Hydrogen**
is a biased tracer of the **matter clustering**
similar to galaxy surveys

In cosmology, **large scales** are fundamental

How can we efficiently observe cosmological volumes?

Intensity Mapping

credit: A. Pourtsidou



The distribution of **neutral Hydrogen**
is a biased tracer of the **matter clustering**
similar to galaxy surveys

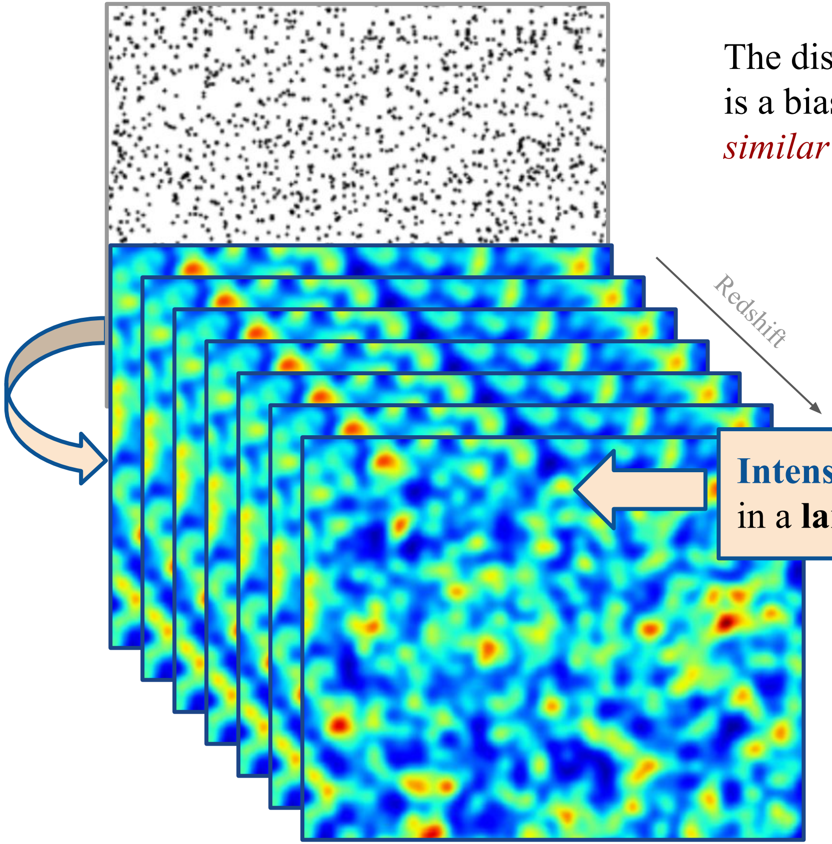
In cosmology, **large scales** are fundamental

How can we efficiently observe cosmological volumes?

Intensity Mapping: total intensity of the 21cm emission line in a **large pixel** (low spatial resolution)

Intensity Mapping

credit: A. Pourtsidou



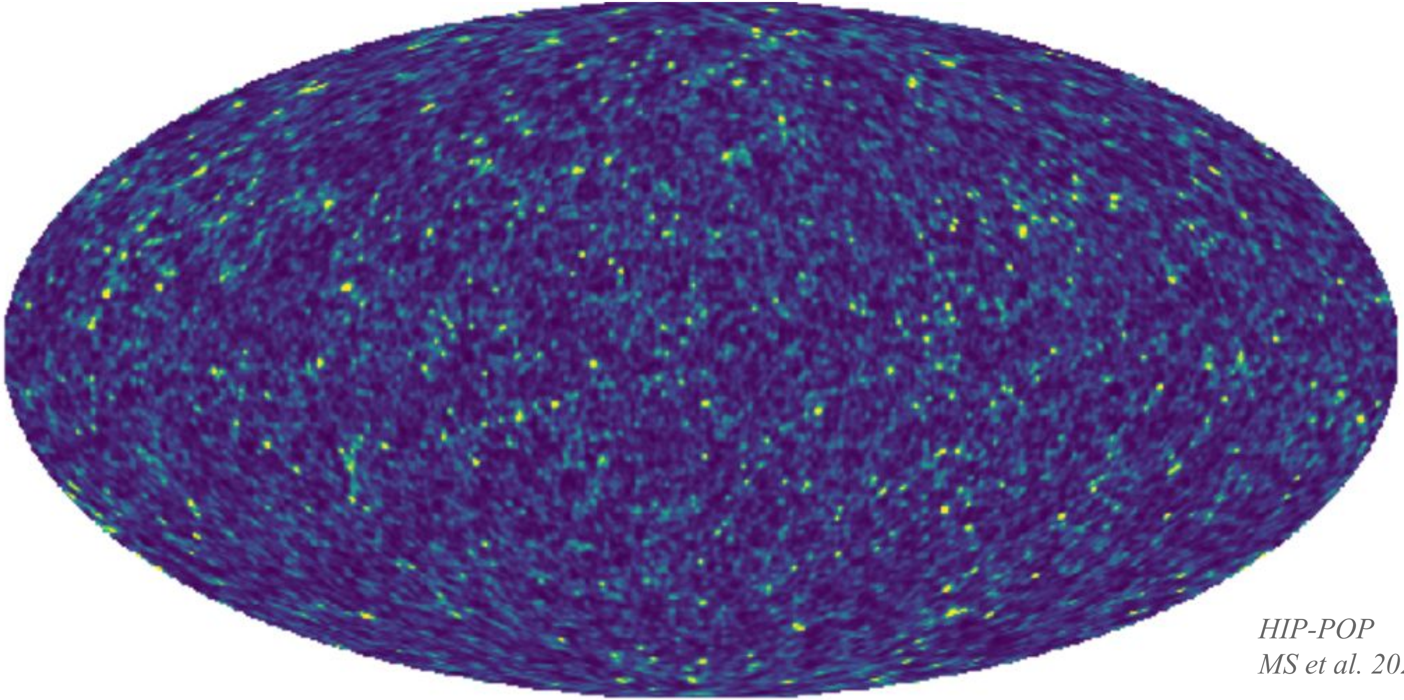
The distribution of **neutral Hydrogen**
is a biased tracer of the **matter clustering**
similar to galaxy surveys

Intensity Mapping: total intensity of the 21cm emission line in a **large pixel** (low spatial resolution)

different frequencies,
different z
high spectral resolution (tomography)

Key cosmological probe

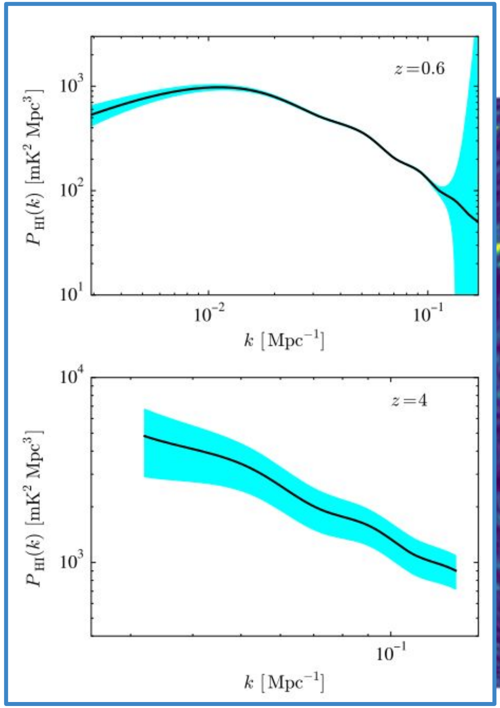
Key cosmological probe

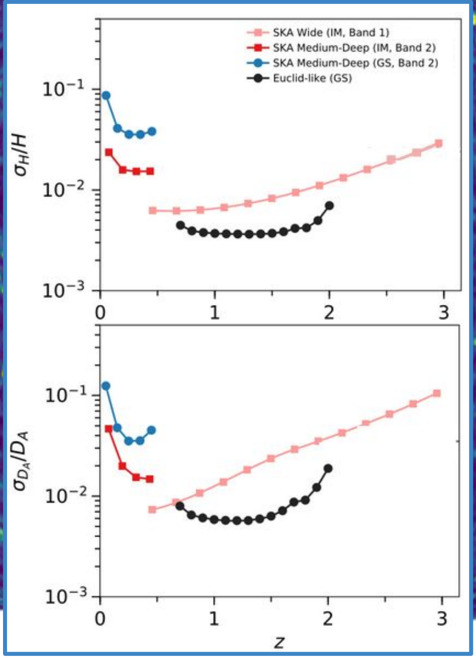


HIP-POP
MS et al. 2022

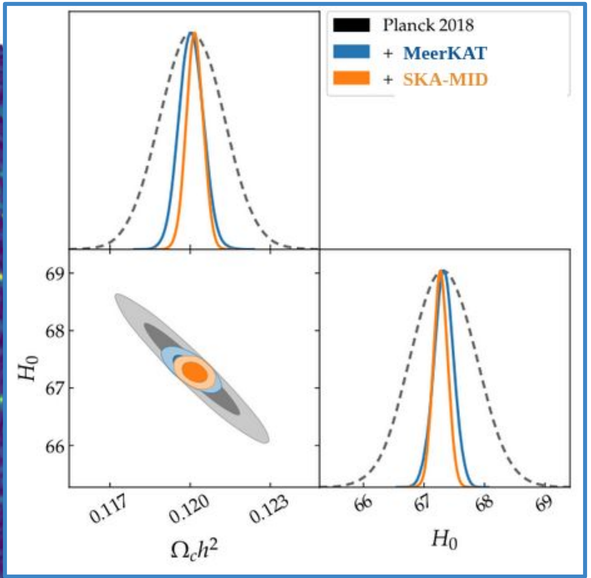
Key cosmological probe

SKA Red Book (2020)





SKA Red Book (2020)



Berti, MS et al. 2022

SKAO forecasts

$$P_{21}(z,k,\mu) = \bar{T}_b^2(z)\left[\, b_{\rm HI}(z) + f(z)\,\mu^2 \,\right]^2 P_{\rm m}(z,k)$$

SKAO forecasts

$$P_{21}(z,k,\mu) = \bar{T}_b^2(z)\left[\,b_{\rm HI}(z) + f(z)\,\mu^2\,\right]^2 P_{\rm m}(z,k)$$

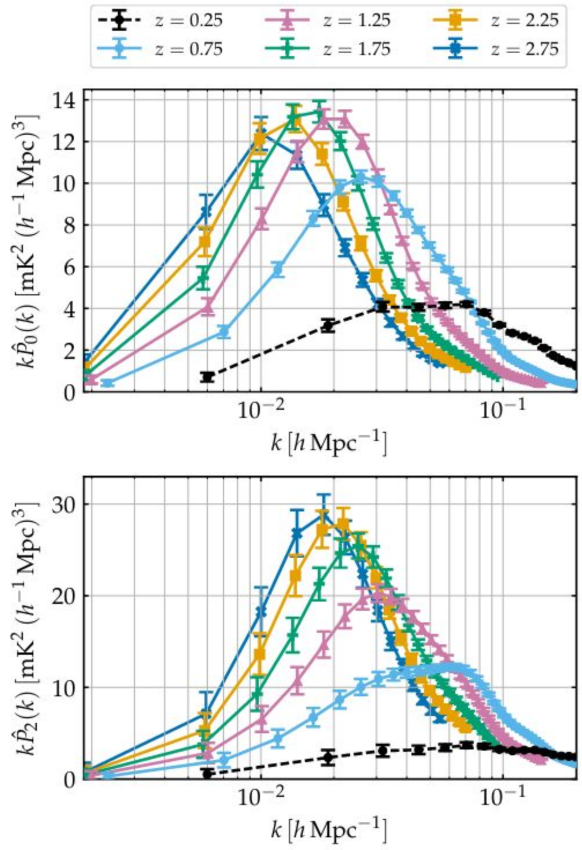
$$P_\ell(z,k) = \frac{(2\ell+1)}{2}\,\bar{T}_b^2(z)\,P_{\rm m}(z,k)\int_{-1}^{1} {\rm d}\mu\,\mathscr{L}_\ell(\mu)\,\left[b_{\rm HI}(z) + f(z)\,\mu^2\right]^2$$

SKAO forecasts

$$P_{21}(z,k,\mu) = \bar{T}_b^2(z)\left[\, b_{HI}(z) + f(z)\,\mu^2 \,\right]^2 P_m(z,k)$$

$$P_\ell(z,k) = \frac{(2\ell+1)}{2}\,\bar{T}_b^2(z)\,P_m(z,k)\int_{-1}^{1} d\mu\,\mathscr{L}_\ell(\mu)\,\left[b_{HI}(z)+f(z)\,\mu^2\right]^2$$

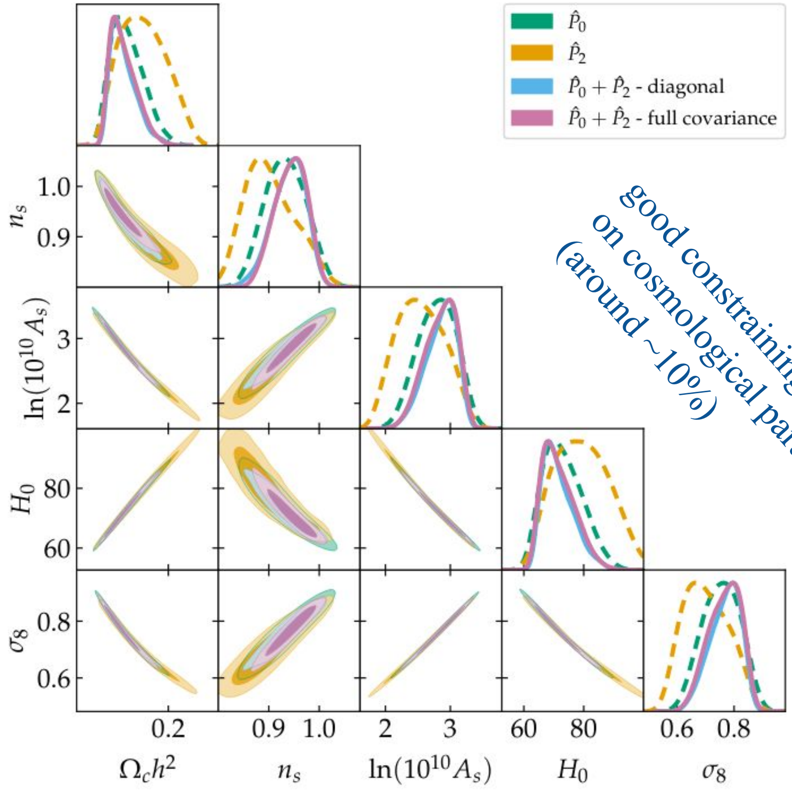
- ❑ We consider only monopole and quadrupole l=0,2
- ❑ SKA-Mid like observations
  - **tomographic** (6 redshift between 0 and 3)
  - **Single-dish**: beam effect
  - expected noise and sky area



Berti, MS et al. 2023

SKAO forecasts

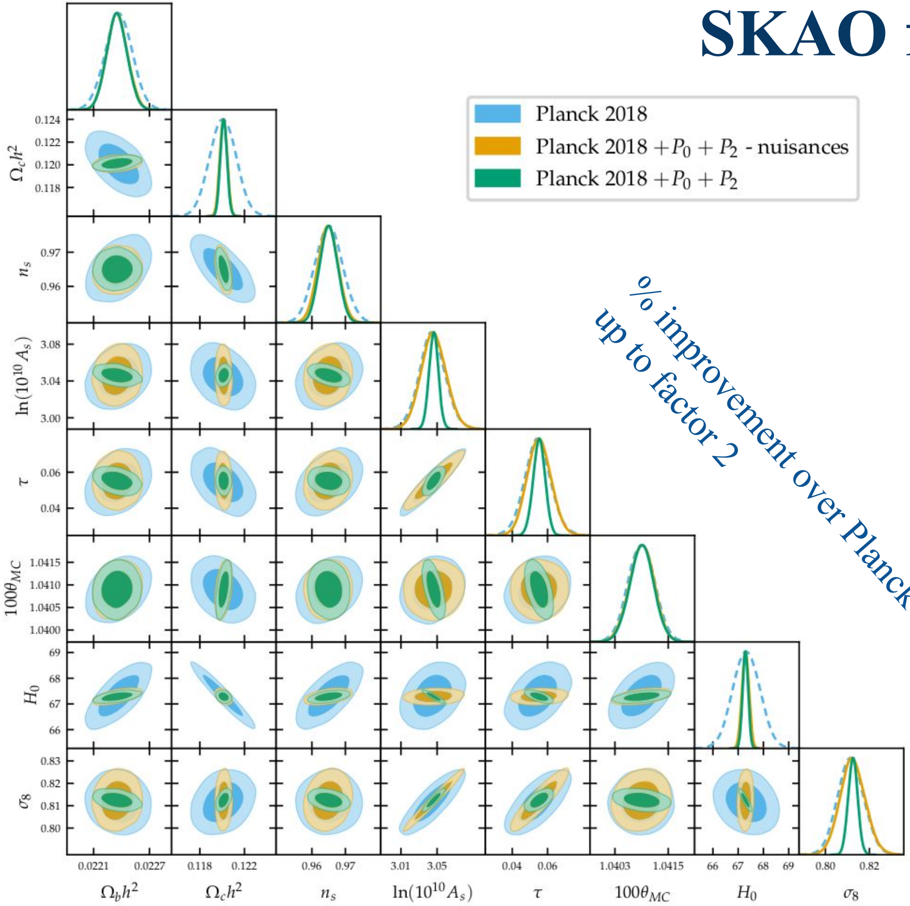
Berti, MS et al. 2023



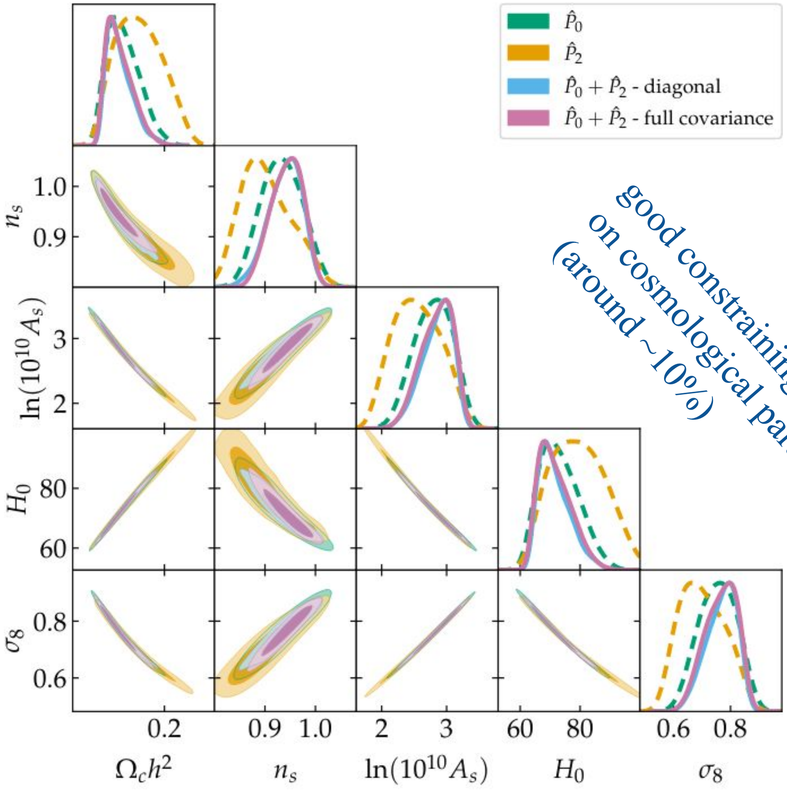
good constraining power on cosmological parameters (around ~10%)

SKAO forecasts

Berti, MS et al. 2023

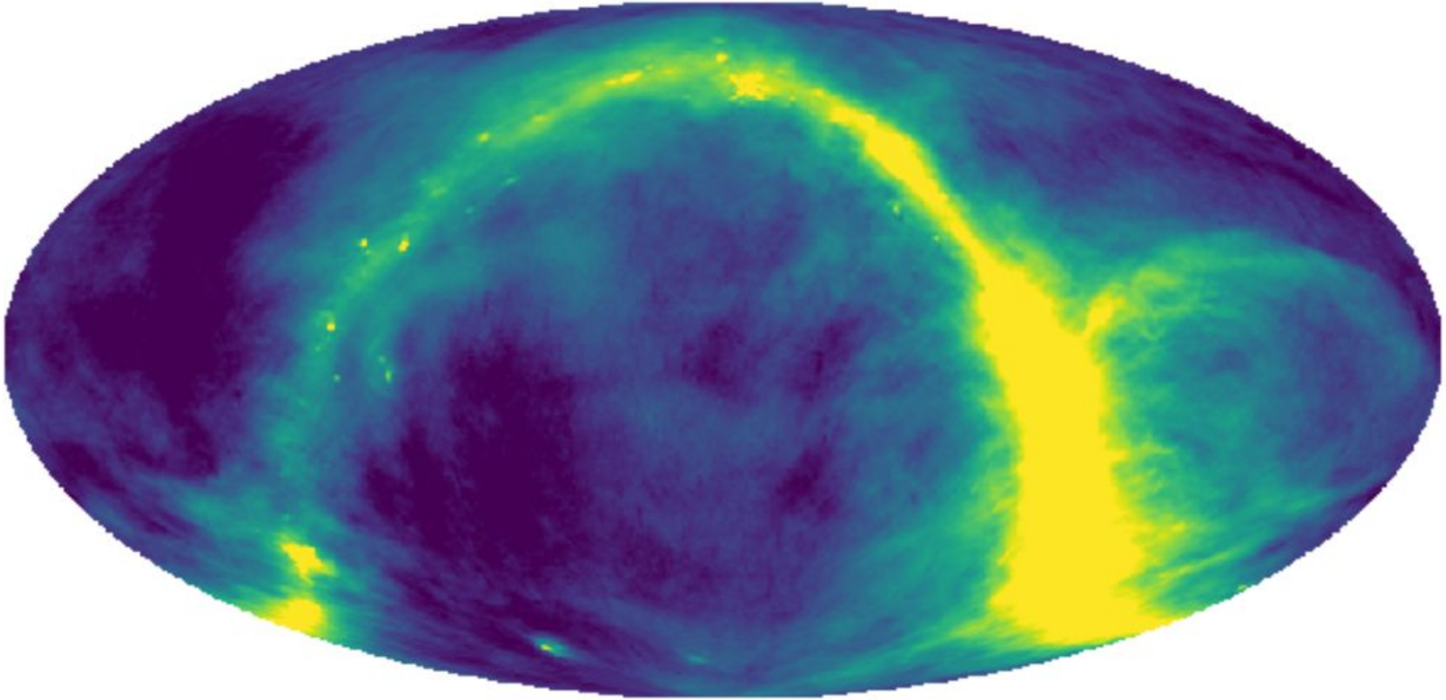


% improvement over Planck
up to factor 2



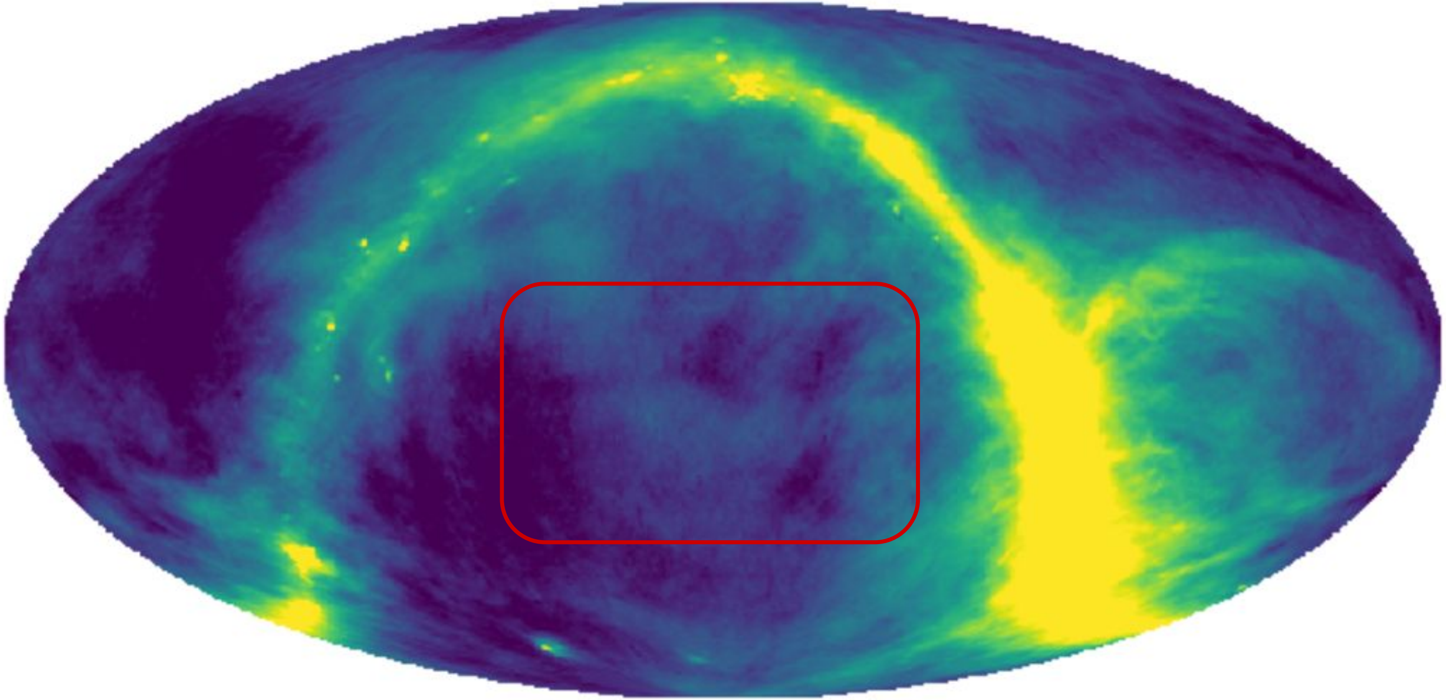
good constraining power
on cosmological parameters
(around ~10%)

The challenge of foregrounds



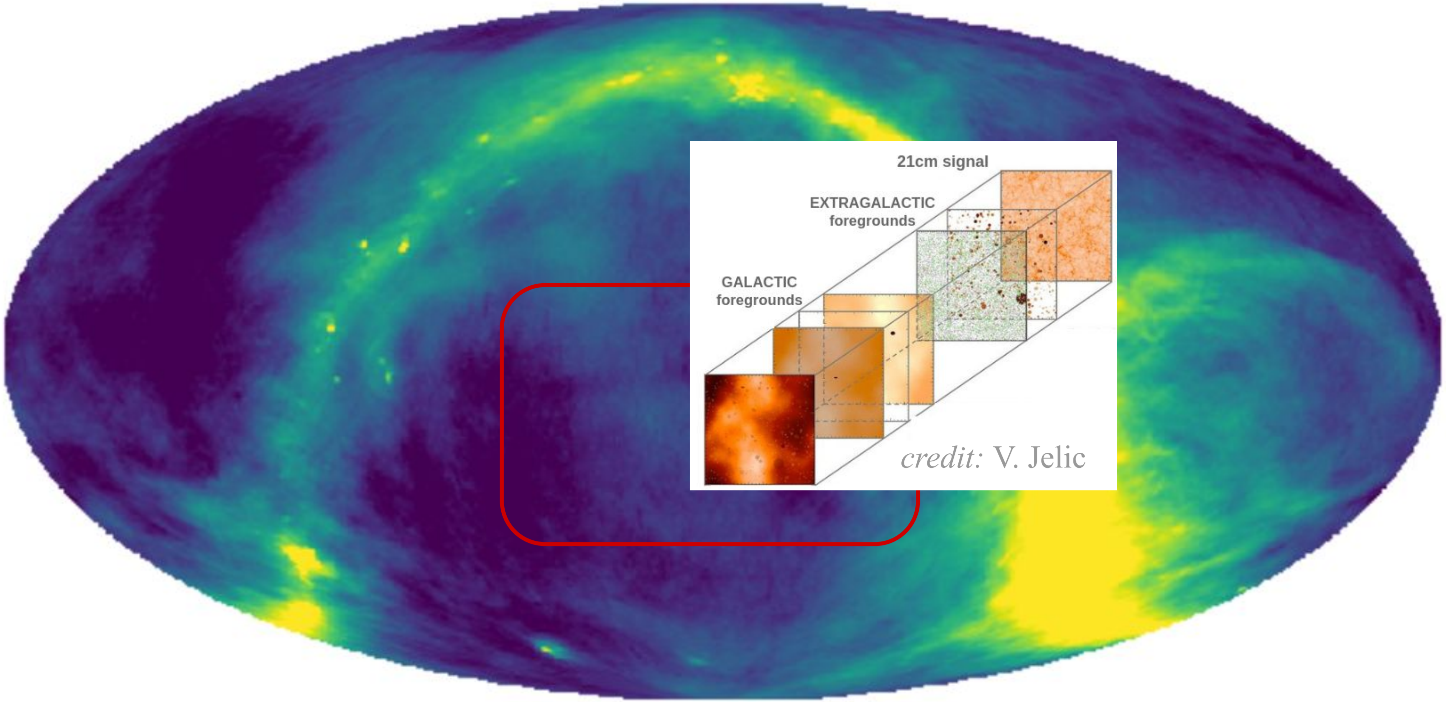
adapted from Haslam et al. (1982)

The challenge of foregrounds



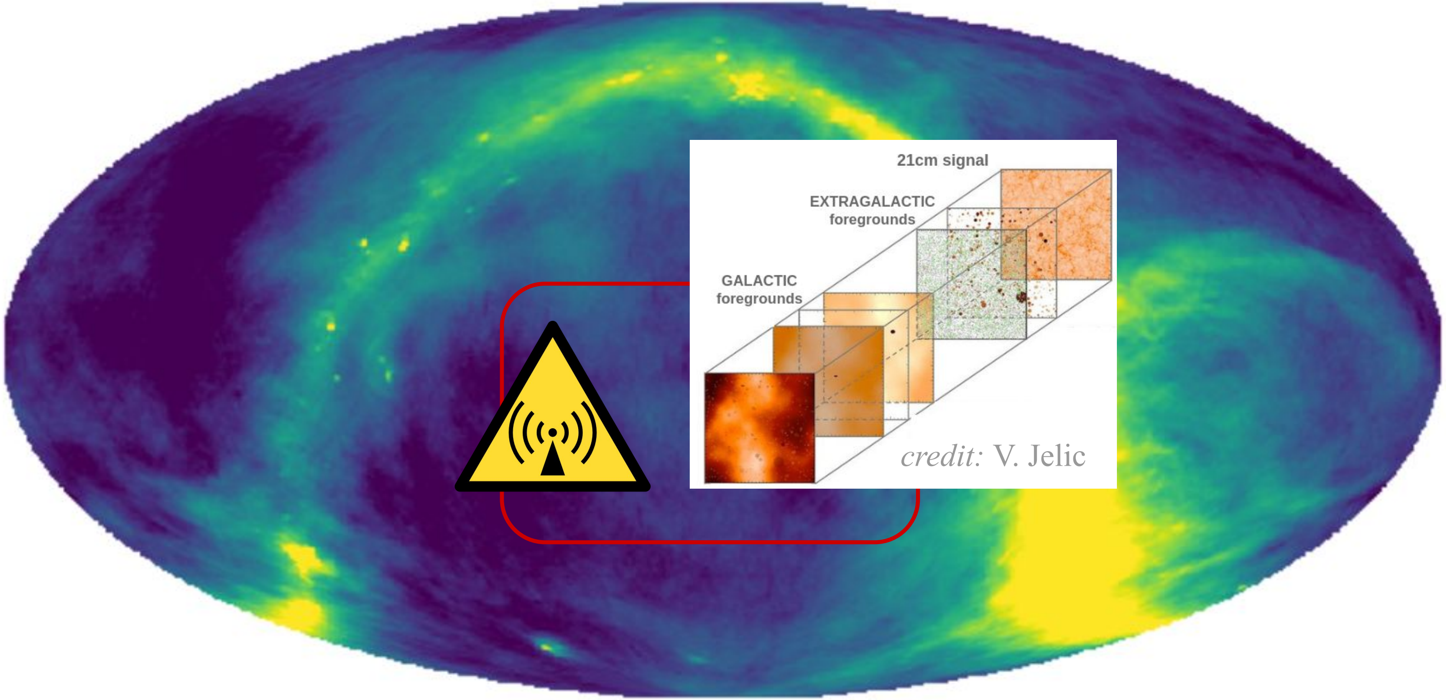
adapted from Haslam et al. (1982)

The challenge of foregrounds



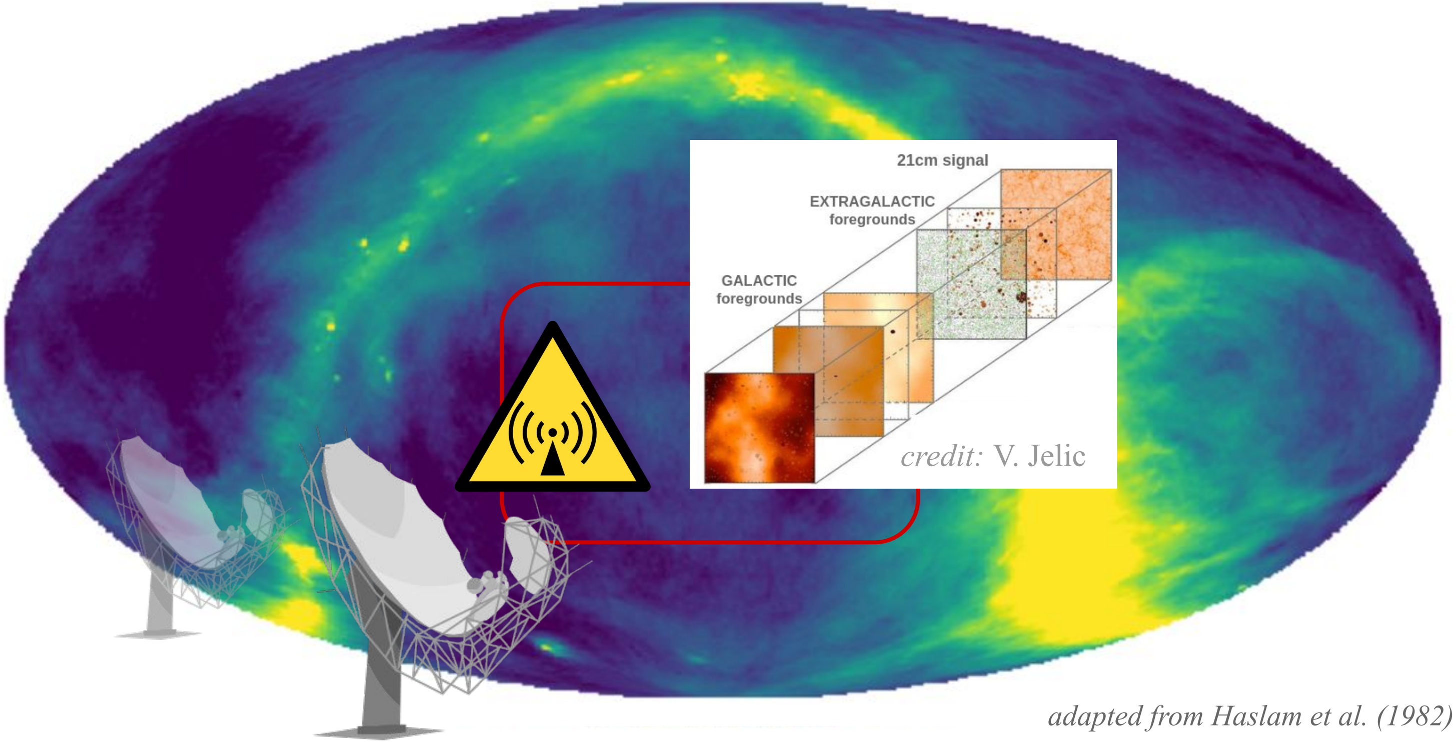
adapted from Haslam et al. (1982)

The challenge of foregrounds



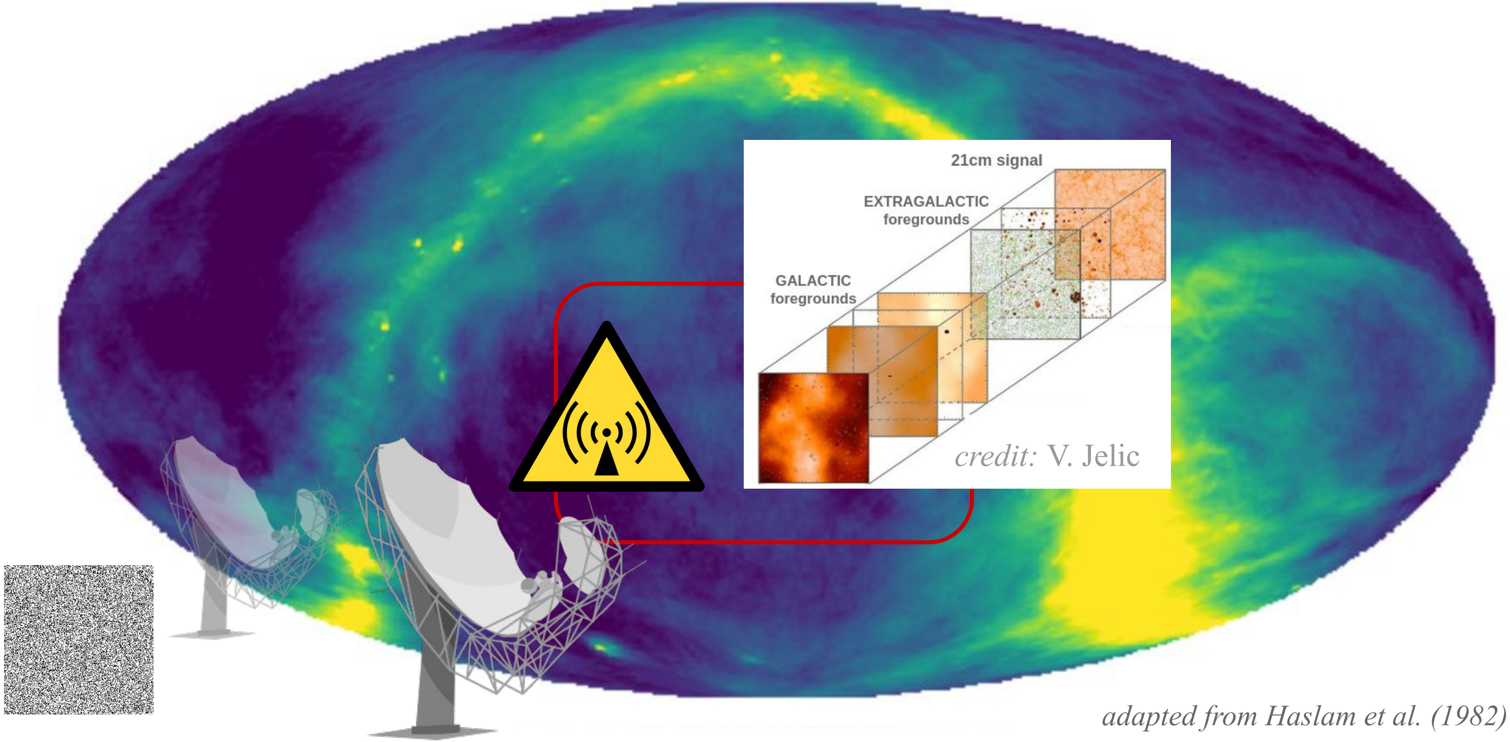
adapted from Haslam et al. (1982)

The challenge of foregrounds



adapted from Haslam et al. (1982)

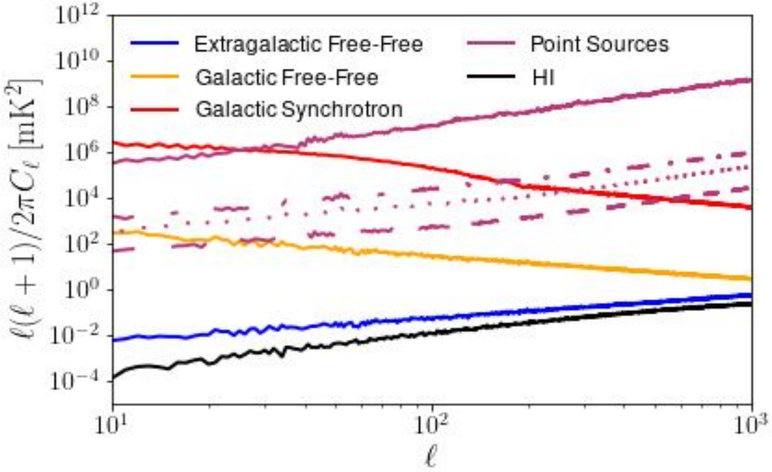
The challenge of foregrounds



adapted from Haslam et al. (1982)

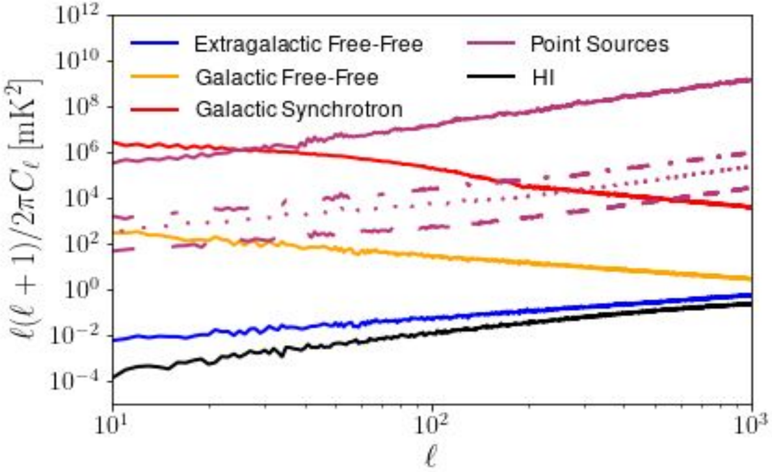
Properties of the foregrounds

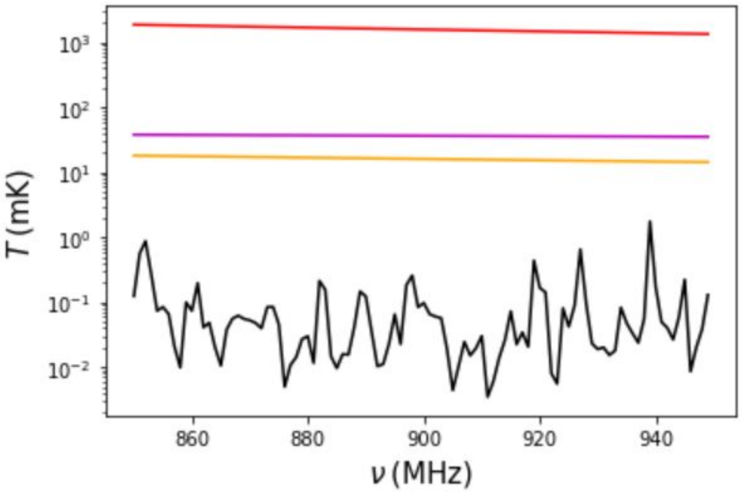
Matshawule, MS et al. 2021



Properties of the foregrounds

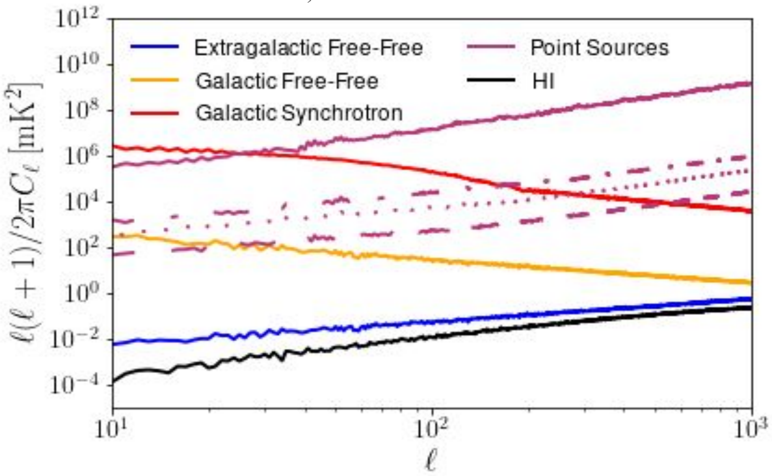
Matshawule, MS et al. 2021

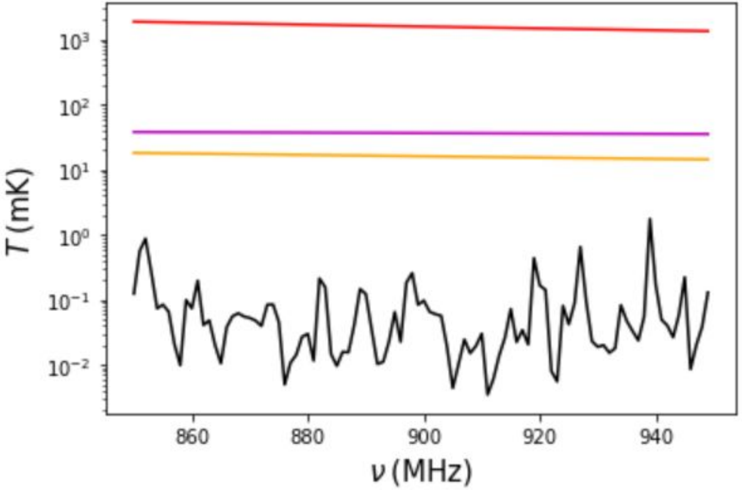




Properties of the foregrounds

Matshawule, MS et al. 2021

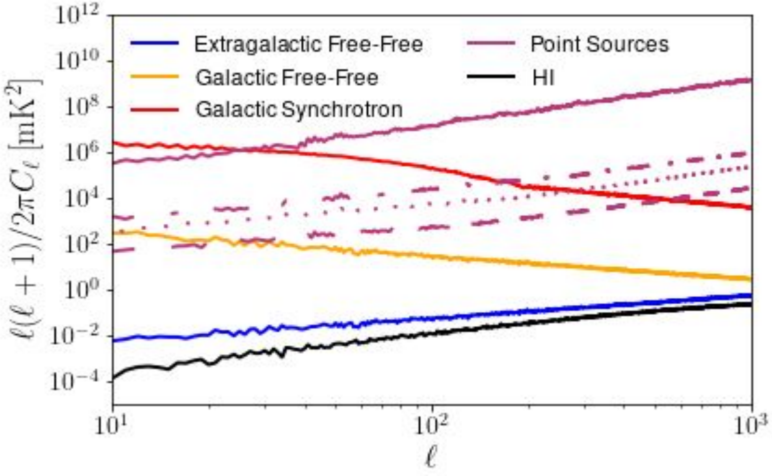


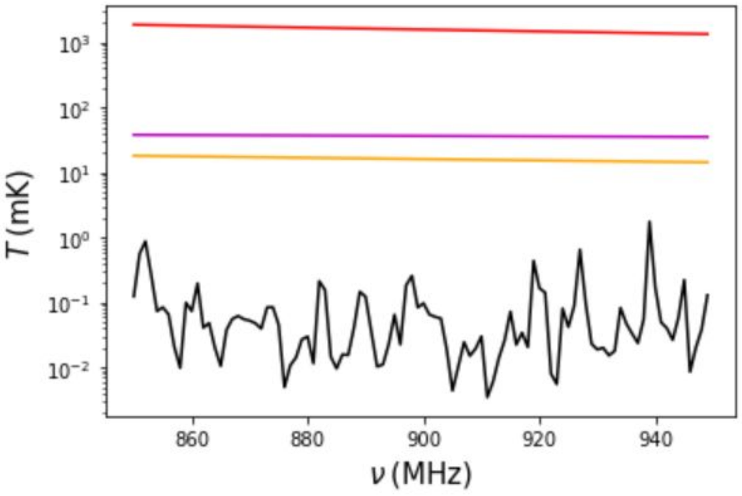


- ❏ foregrounds are orders of magnitude stronger than the 21cm signal
- ❏ their frequency behaviour is smooth (highly correlated)

Properties of the foregrounds

Matshawule, MS et al. 2021





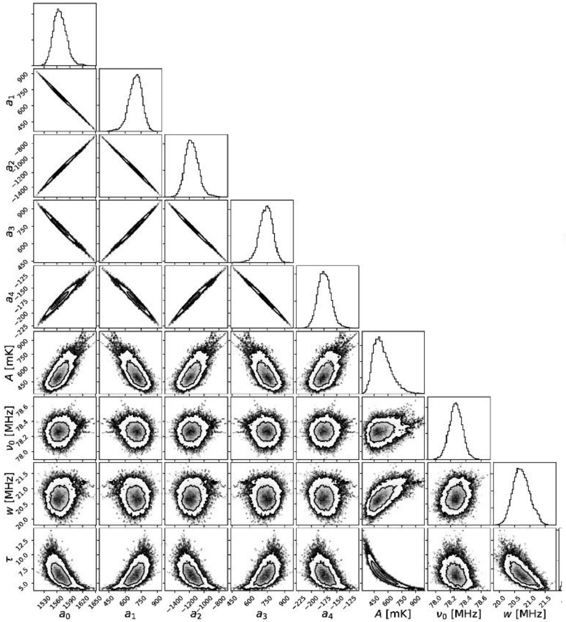
- foregrounds are orders of magnitude stronger than the 21cm signal
- their frequency behaviour is smooth (highly correlated)

Questions:

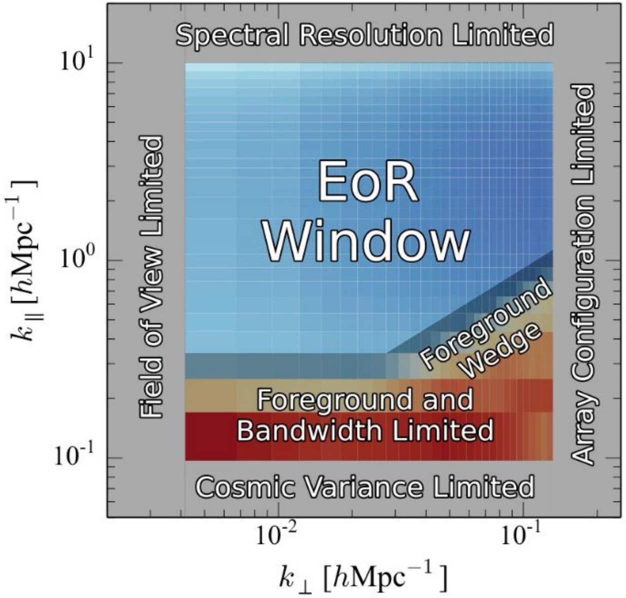
- Can the properties of the foregrounds be used to separate them from the pristine **21cm signal**?
- **Even if we add some realism to our simulations?** (foregrounds,beam response,noise,RFI,..)

Dealing with foregrounds

Various strategies: e.g. **modelling**, **avoidance** and **separation/cleaning**



Bowman et al. (2018)



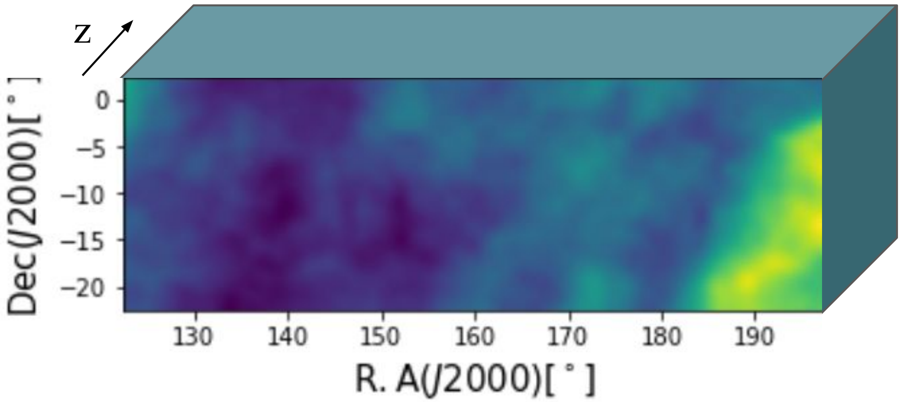
Liu et al. (2014)

separation/cleaning:

PCA, kPCA, FastICA, GMCA, GPR, ML-GPR

A cleaning example

Mock observation “**cube**”



Simulation includes:

- ❑ 100 channels around redshift 0.5
- ❑ Foreground contamination: *Synchrotron, Free-free, point sources*
- ❑ Gaussian beam
- ❑ White noise

T = **A**s+n+c

A mixing matrix of the foreground sources

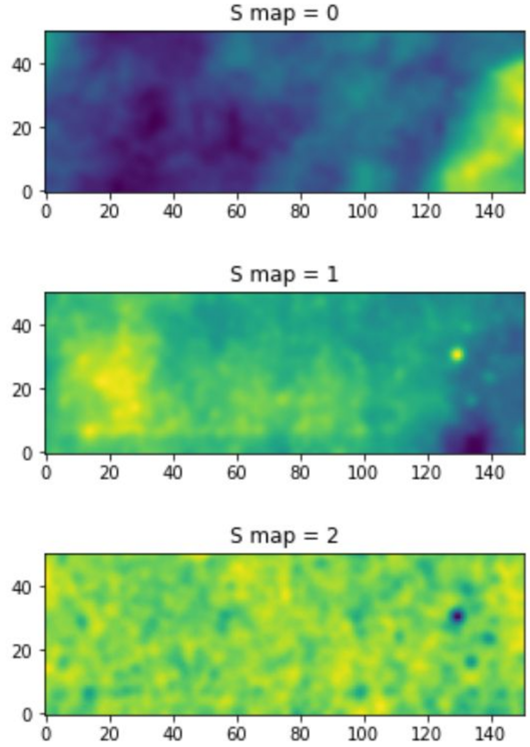
noise

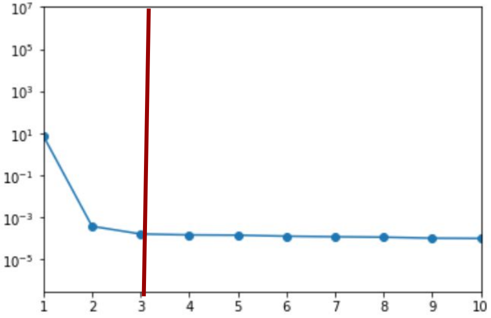
Cosmological signal

How many sources?

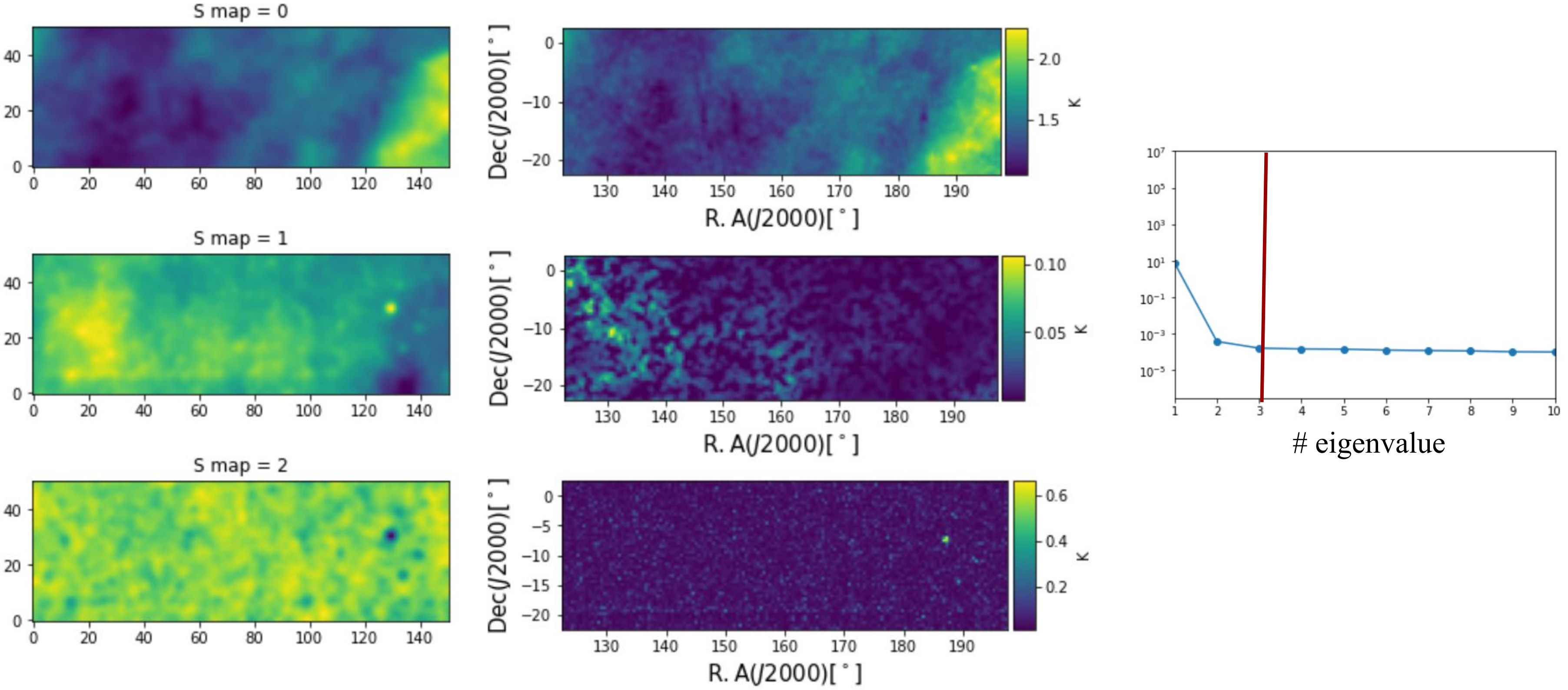
$\mathbf{N_{fg}}$ need to be estimated/guessed

A cleaning example





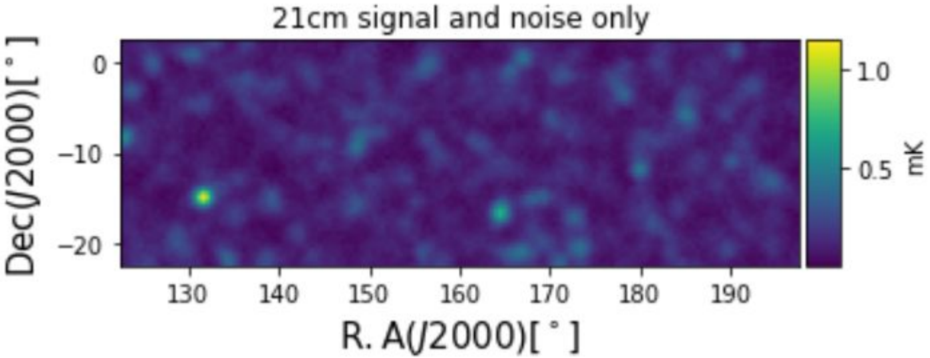
A cleaning example

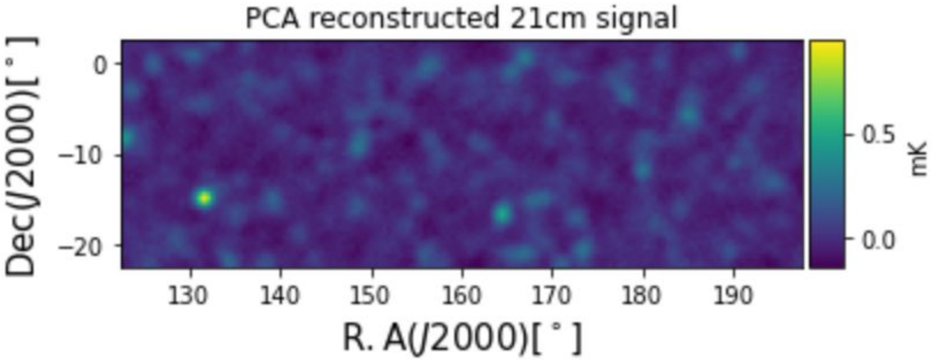


A cleaning example

c+n = T - **As**

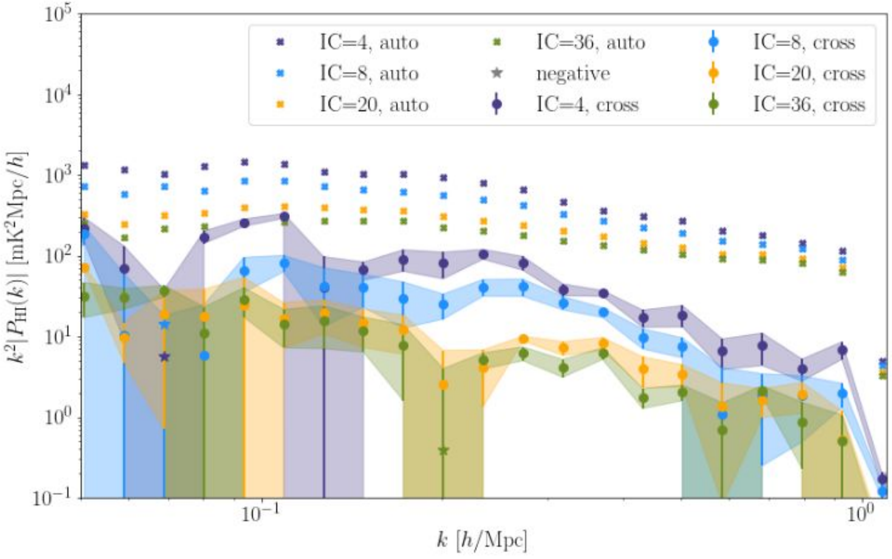
A mixing matrix including only the first Nfg components





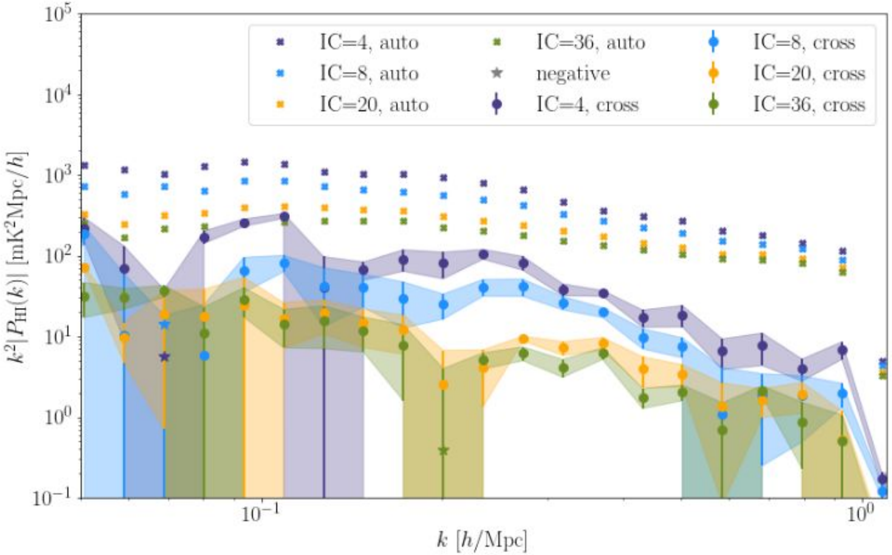
With GBT data

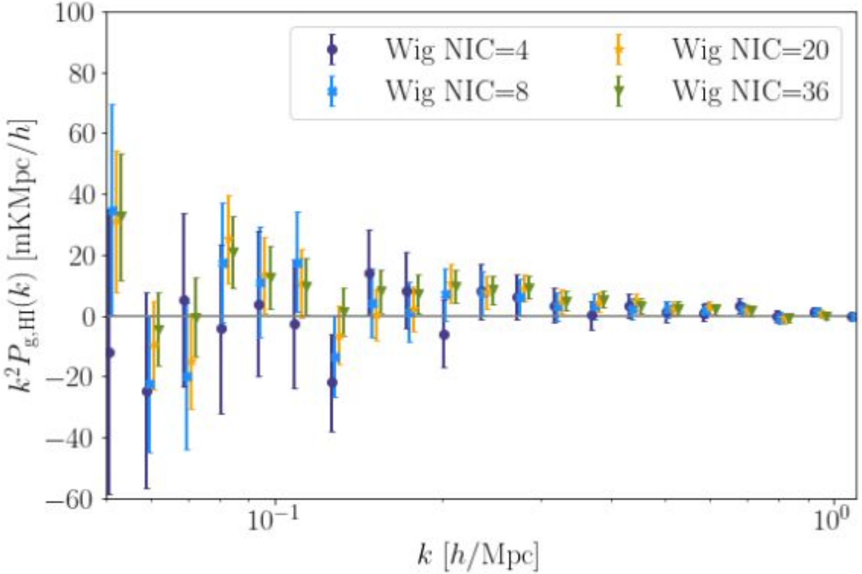
Wolz et al. 2022



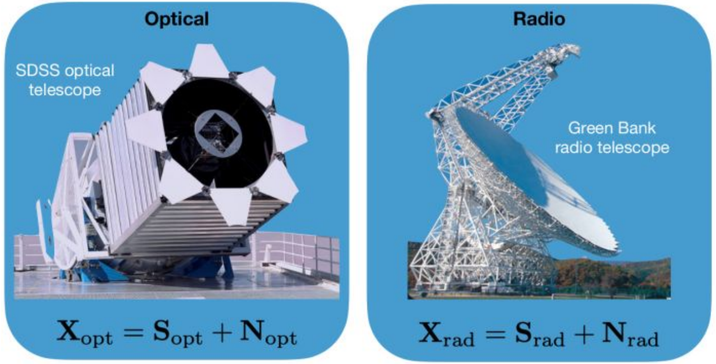
With GBT data

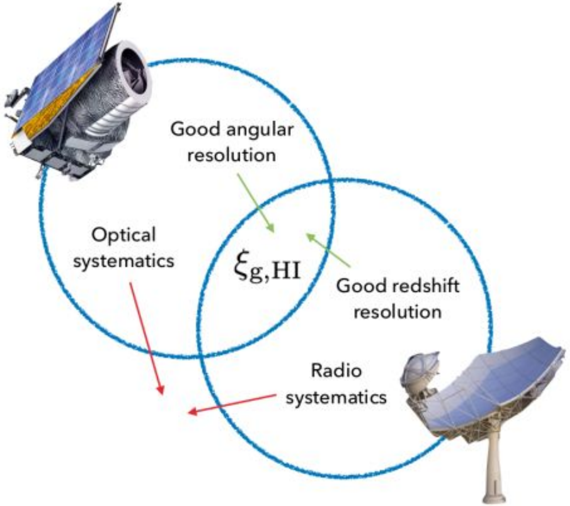
Wolz et al. 2022





Mitigation of systematics with cross-correlation





Auto Correlation:

uncorrelated

$$\langle \mathbf{X}_{\rm opt}\mathbf{X}_{\rm opt}\rangle = \langle \mathbf{S}_{\rm opt}\mathbf{S}_{\rm opt}\rangle + 2\langle \mathbf{S}_{\rm opt}\mathbf{N}_{\rm opt}\rangle + \langle \mathbf{N}_{\rm opt}\mathbf{N}_{\rm opt}\rangle$$

$$\langle \mathbf{X}_{\rm opt}\mathbf{X}_{\rm opt}\rangle = \langle \mathbf{S}_{\rm opt}\mathbf{S}_{\rm opt}\rangle + \langle \mathbf{N}_{\rm opt}\mathbf{N}_{\rm opt}\rangle$$

signal you want *noise you don't want*

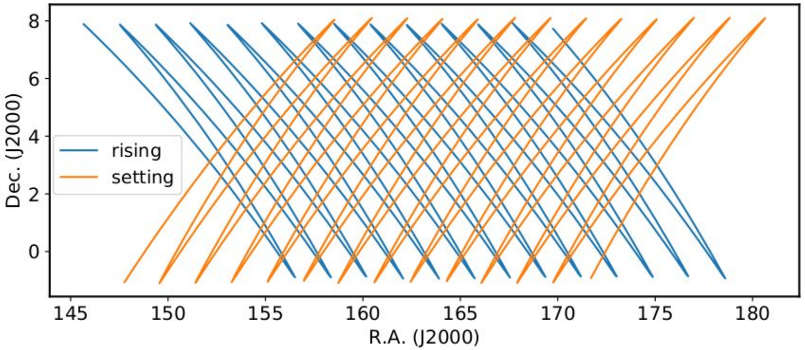
Cross Correlation:

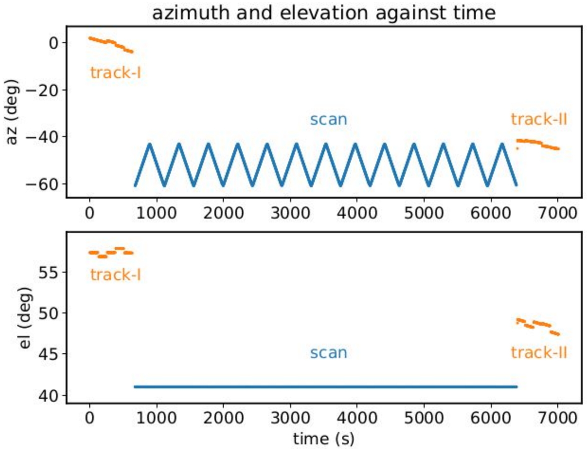
$$\langle \mathbf{X}_{\rm opt}\mathbf{X}_{\rm rad}\rangle = \langle \mathbf{S}_{\rm opt}\mathbf{S}_{\rm rad}\rangle + \langle \mathbf{S}_{\rm opt}\mathbf{N}_{\rm rad}\rangle + \langle \mathbf{S}_{\rm rad}\mathbf{N}_{\rm opt}\rangle + \langle \mathbf{N}_{\rm opt}\mathbf{N}_{\rm rad}\rangle$$

courtesy of Steve Cunnington

Intensity Mapping with MeerKAT

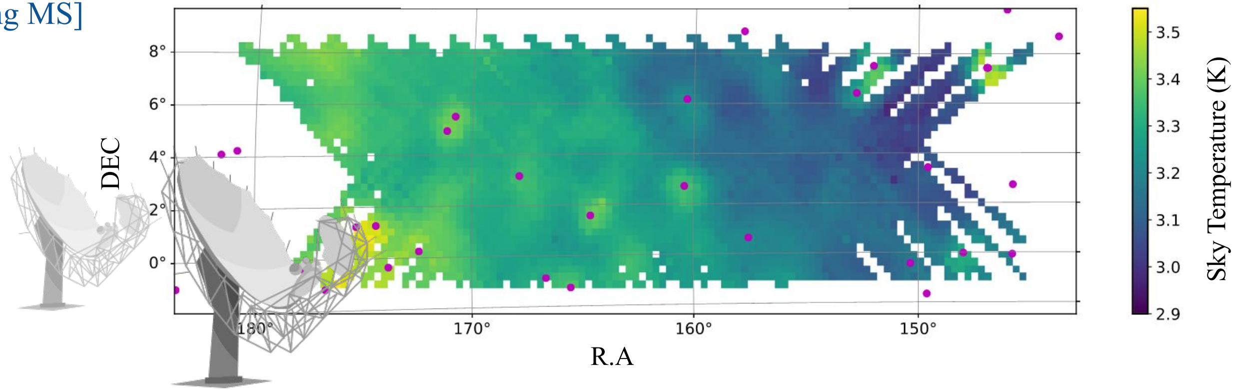
Antennas	All 64 MeerKAT dishes
Observation mode	Single-dish
Frequency range	0.856-1.712 GHz
Frequency resolution	0.2 MHz
Time resolution	2s
Exposure time	1.5hr x 7 scans
Target field	WiggleZ 11hr field ($10^\circ \times 30^\circ$)





MeerKAT observations

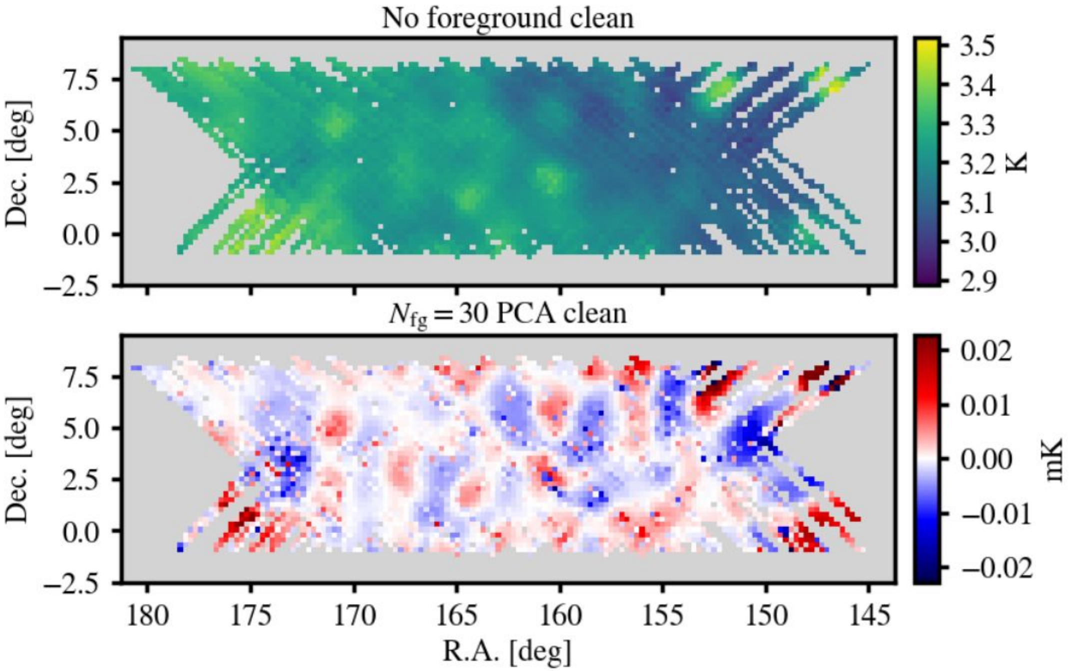
Wang et al. 2021
[including MS]



MeerKLASS: 64 MeerKAT antennas used in **single-dish mode**

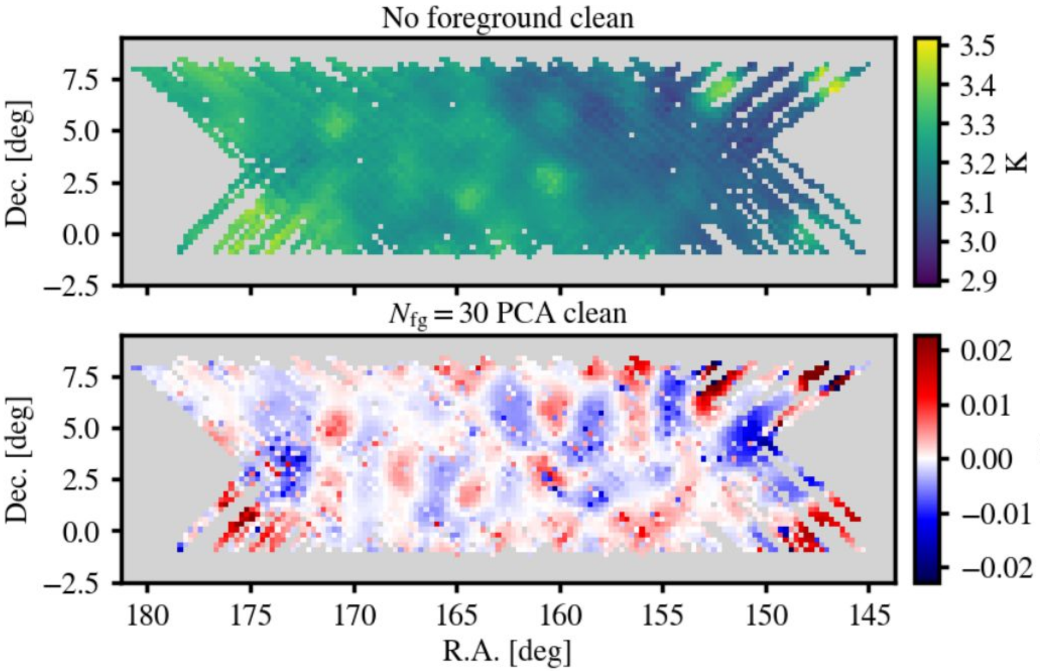
- ❑ first successful calibration of **intensity mapping data from MeerKAT**
- ❑ L-band: 850-1700 MHz (4096 channels)

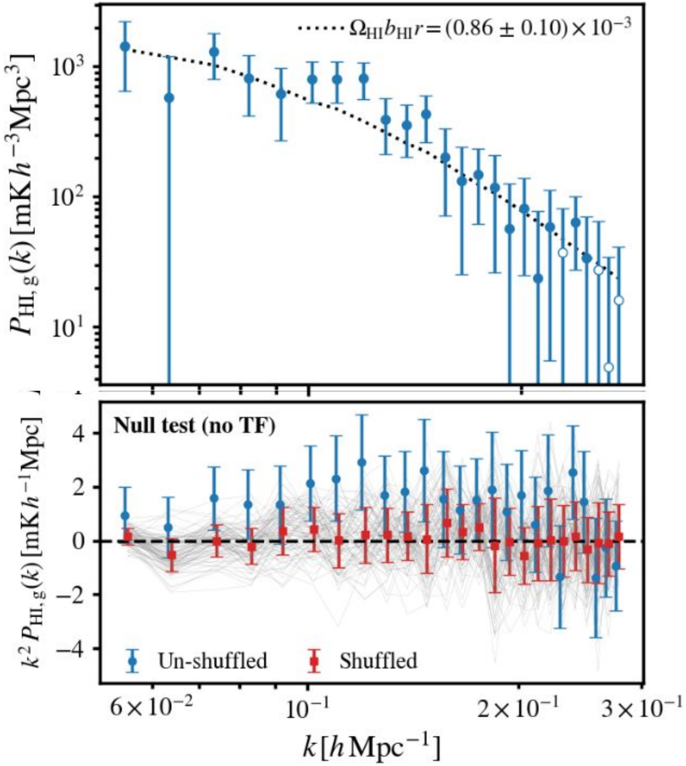
MeerKLASS results



Cunnington et al. 2022 [including MS]

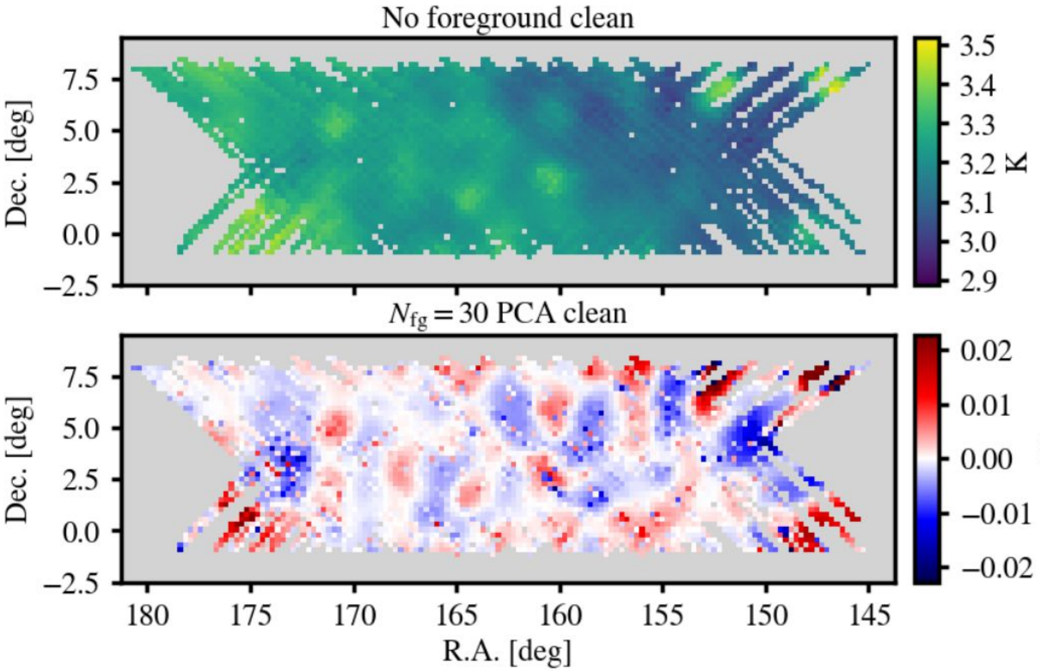
MeerKLASS results

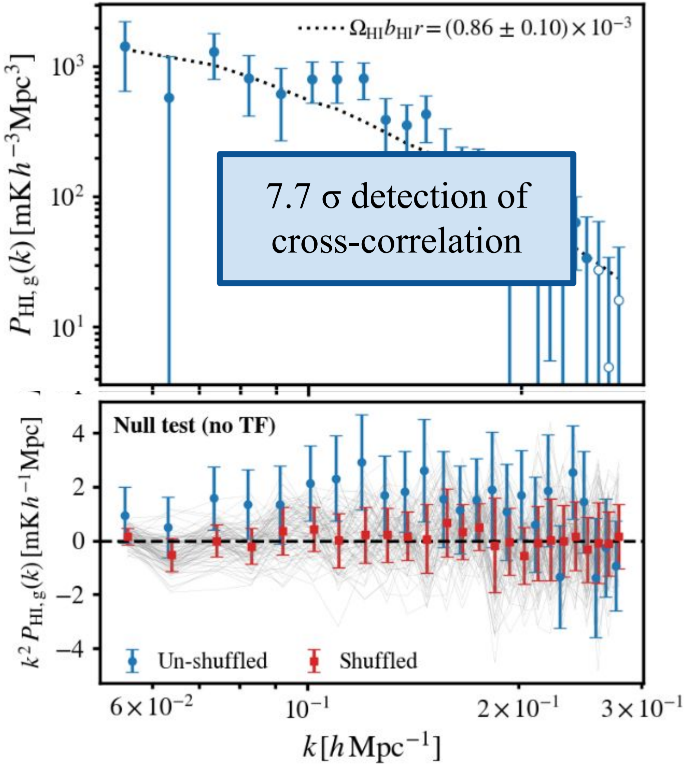




Cunnington et al. 2022 [including MS]

MeerKLASS results





7.7 σ detection of cross-correlation

Cunnington et al. 2022 [including MS]

Towards the SKA Observatory

We have:

21cm intensity mapping data difficult to clean (signal only in cross-correlation)

Simulations that are still not a realistic representation of the actual data

Cleaning methods that have still to be extensive tested with realistic simulations

We would like:

More and better data

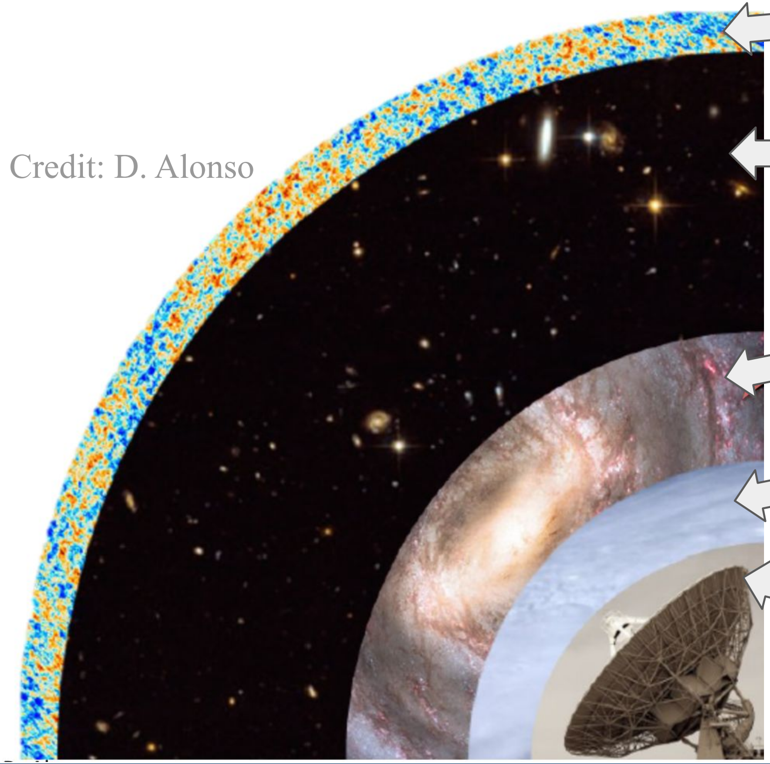
More realistic simulations mimicking the data

More sophisticated cleaning methods tested on more realistic simulations

Final aim:

A 21cm **(auto) power spectrum detection** validated with realistic simulations and tested with various and robust cleaning methods

Ingredients for the simulations



Credit: D. Alonso

mock 21cm intensity maps

extra-galactic foregrounds

Galactic foregrounds

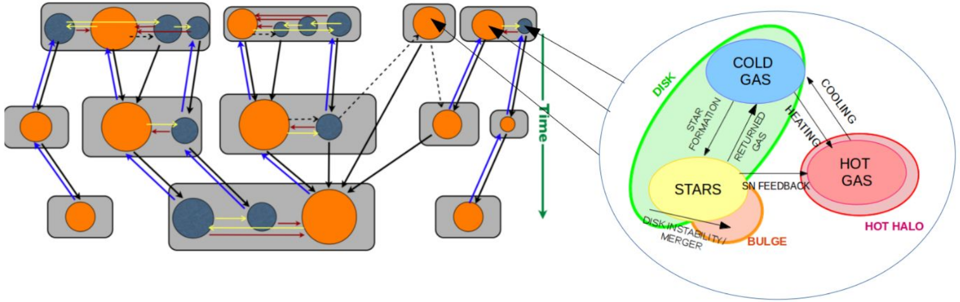
atmosphere and ground contamination

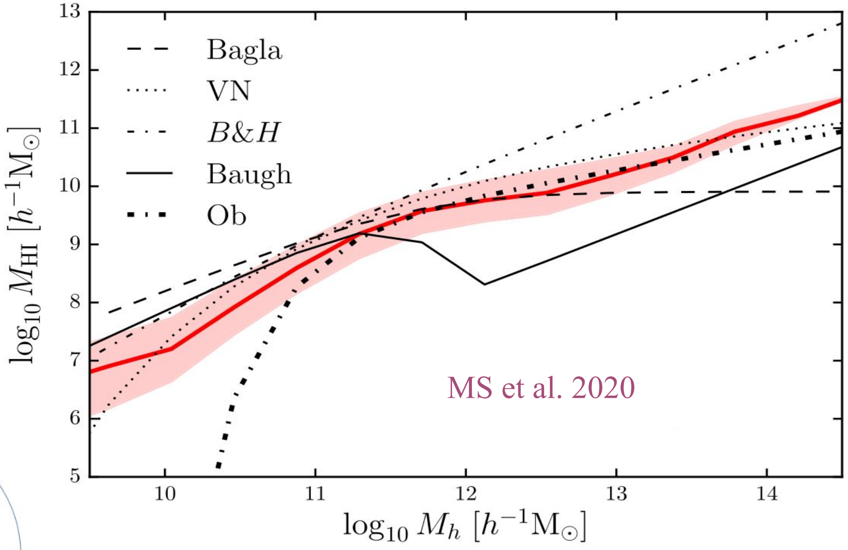
RFI and instrumental effects

Mock 21cm maps

HI properties

- ❏ HI is a key ingredient for galaxy evolution
- ❏ estimation of HI bias properties and of the **MHI-Mhalo relation** using the semi-analytical model GAEA (the GAlaxy Evolution and Assembly model, De Lucia et al. 2016, Xie et al. 2018)

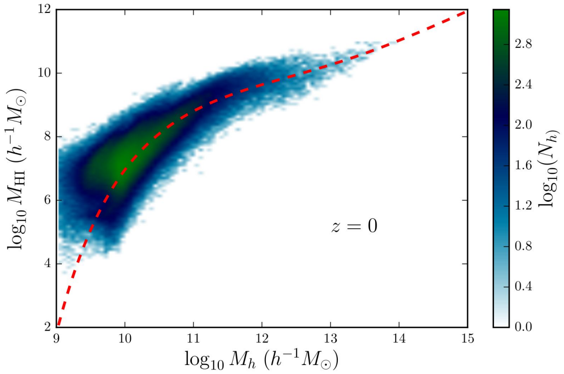




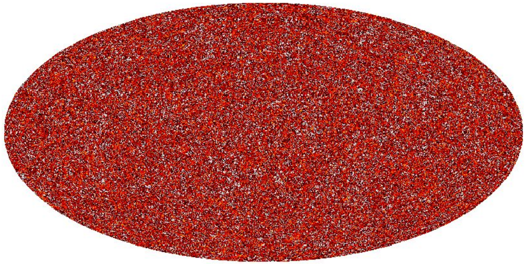
Mock 21cm maps

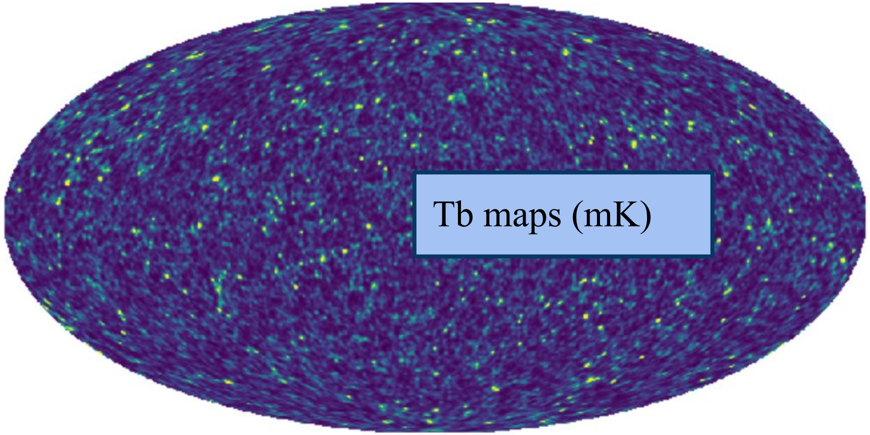
Fast 21cm intensity map generation

- ❑ fundamental for **cosmological volume** and **end-to-end simulations**
- ❑ HOD methods on fast halo catalogues
  MS et al. 2020,2022



Mhalo





Foregrounds

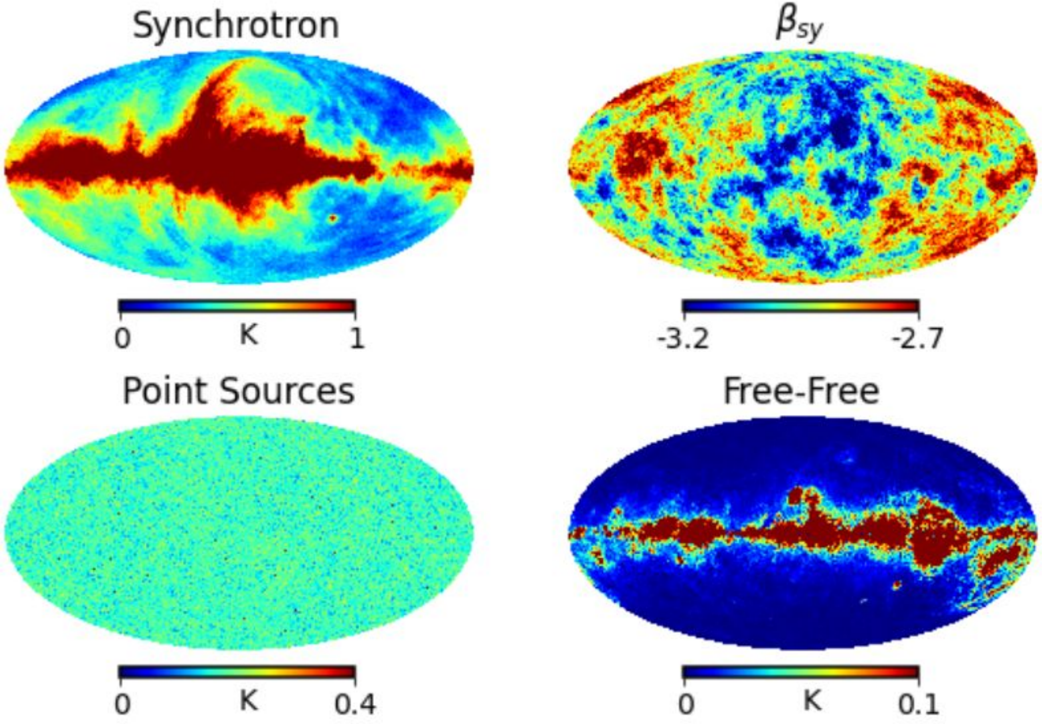
Typical modeling:

Haslam 408 MHz
Ramazeilles et al. (2015)

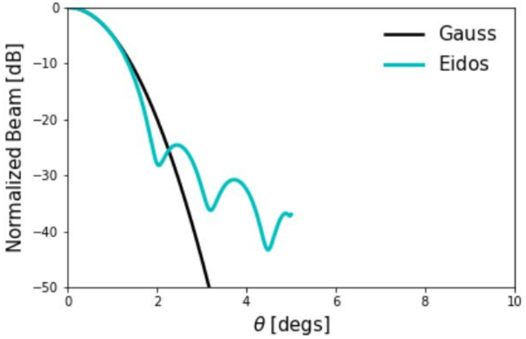
Spatially varying synch spectral index
Miville-Deschenes et al. (2008)

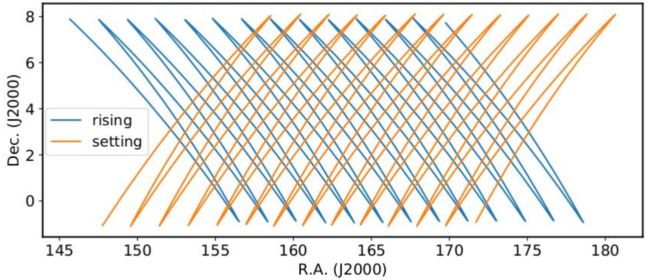
Free-Free from Planck Sky Model
Delabruille et al (2013)

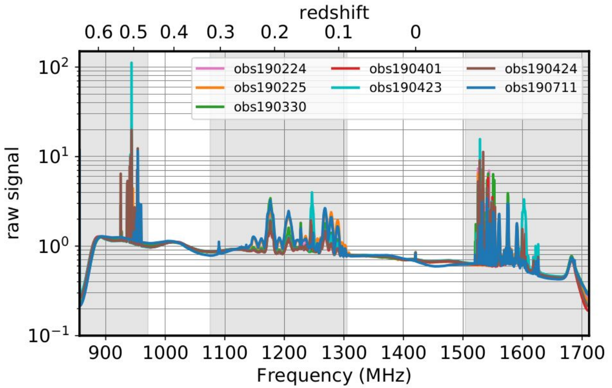
Extragalactic PS
Olivari (2018), Matshwule et al. (2021)



Instrumental effects







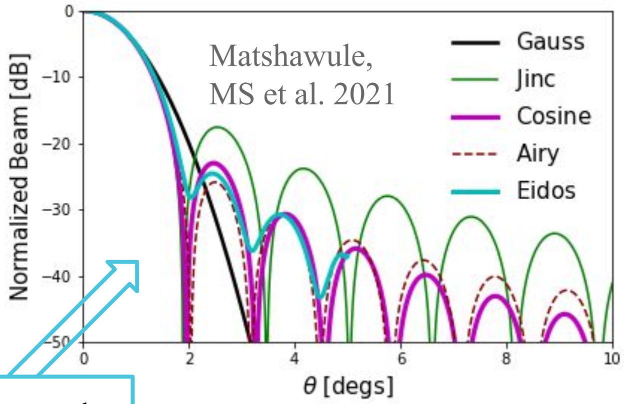
Need a realistic beam modeling
side-lobes, frequency evolution,
more accurate deconvolution

Scanning strategy
non homogeneous noise, need
for real space convolution,
polarization leakage

Radio Frequency Interference (RFI)
impact on cleaning,
impact on signal interpretation

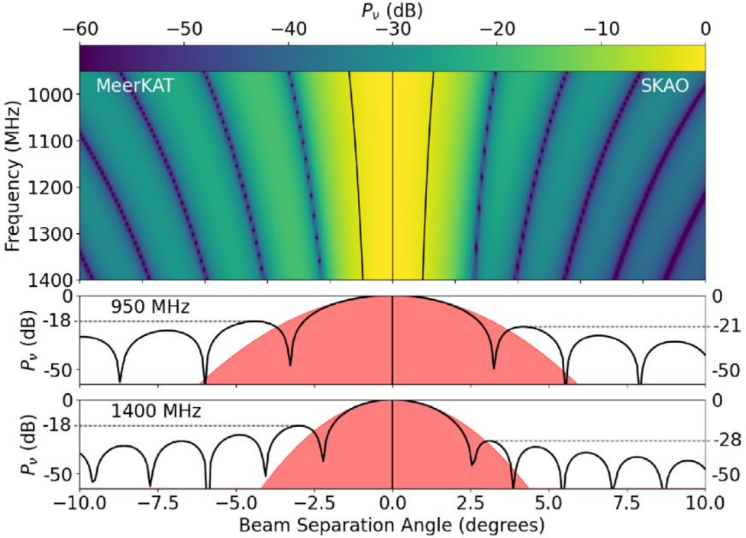
Instrumental effects: the beam

- MeerKAT beam has **side-lobes** (same for SKA-MID)
- a strong point source in the side-lobes contaminates the signal and can complicate the foreground subtraction



Eidos: measured
Asad et al. 2020

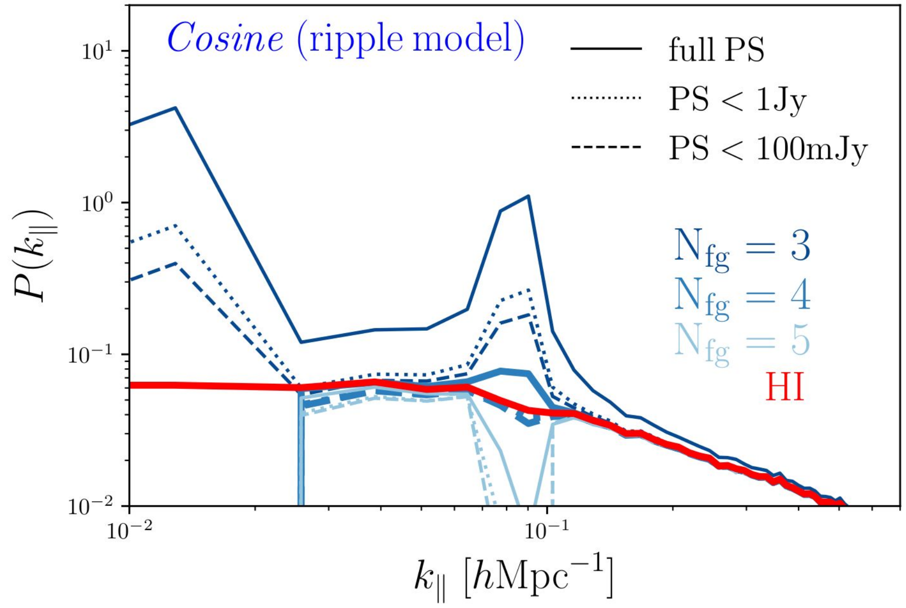
Airy beam
Harper et al. (2018)



MS et al. 2022

The beam evolves with frequency

Effect of the telescope beam



Matshawule, MS et al. 2021

a realistic **MeerKAT** beam**:** side-lobes (cosine) and a non-trivial frequency evolution (ripple)

- **point sources** and synchrotron spatial structures coupled with the beam **complicate the cleaning**
- Careful **beam-deconvolution** alleviates the problem but need to be careful for precision cosmology

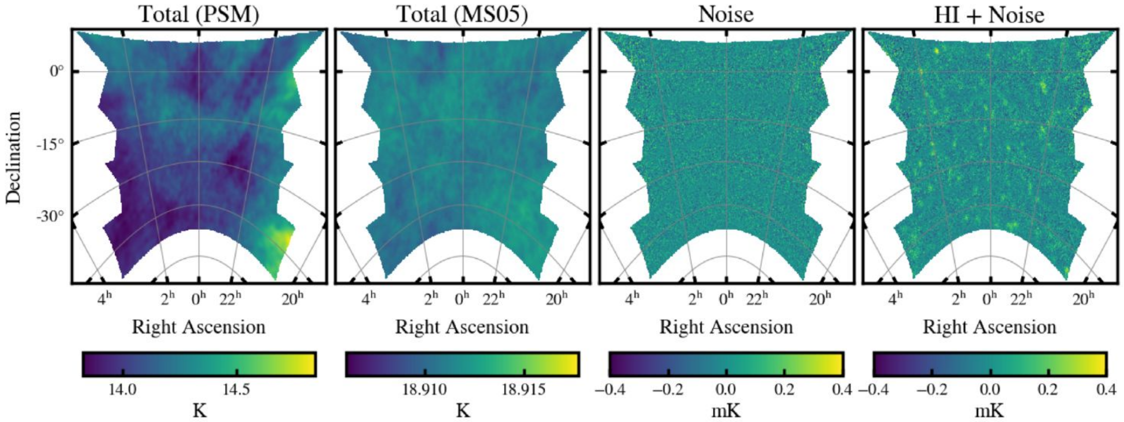
Foreground subtraction challenge

(subset of) **SKA IM Focus Group**

Project setup:

- various foreground models and realistic HI maps
- **instrumental modeling MeerKAT-like and SKAO-like**
- 9 different foreground removal methods (PCA, FastICA, …)

Isabella Carucci, Steve Cunnington, Ze Fonseca, Stuart Harper, Mel Irfan, Alkistis Pourtsidou, Marta Spinelli, Laura Wolz



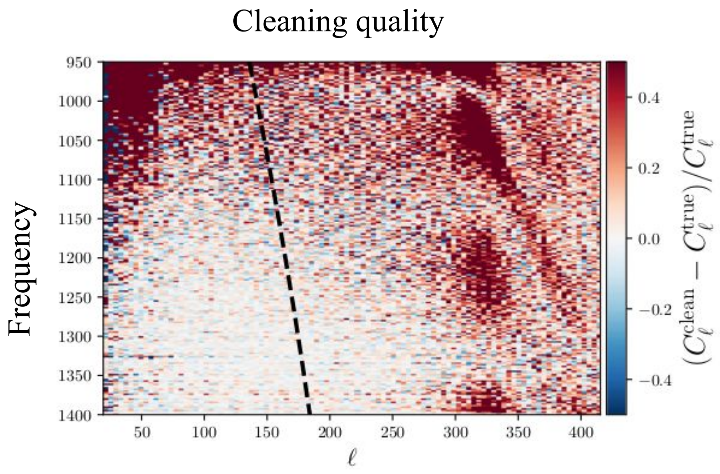
Blind challenge to discover weaknesses and strengths of the various methods

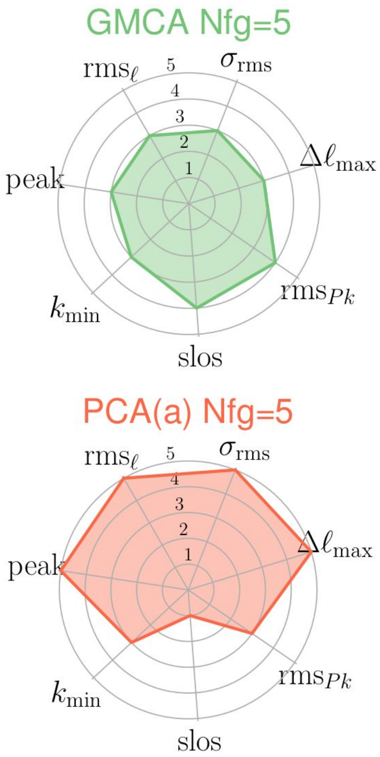
given IM data now, would your favorite method extract the cosmological signal?

Foreground subtraction challenge

- How much can **instrument/foregrounds** coupling impact the signal reconstruction?
- definition of statistics and metrics to evaluate the relative performances

Realistic instrumental effects inevitably complicate the foreground cleaning





MS et al. 2022

Moving forward

Data:

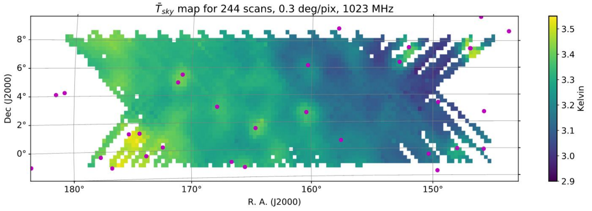
Keep working with pathfinder data (MeerKLASS) to understand the instrument and improve the pipelines

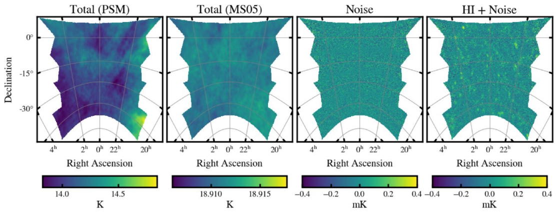
Simulations:

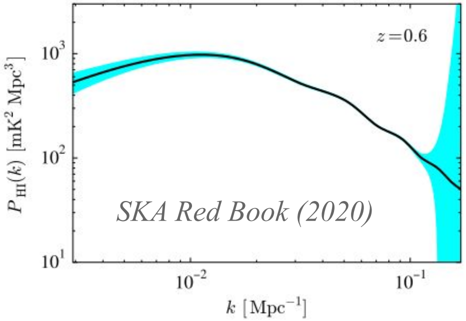
Improve and refine end-to-end simulations

Final aim:

A 21cm **(auto) power spectrum detection** validated with realistic simulations







- ❏ new L band data under analysis (41 x 1.5 h)
- ❏ UHF band available (could go to higher redshift)

Conclusions

- **21cm Cosmology** still have to prove its full potential but offers an incredible window into the evolution of the Universe
- Intensity Mapping surveys are taking data (and new instrument are planned)
- **detection in cross-correlation** from MeerKLASS survey x galaxy survey (7.7 σ)
- analysing new data: effort in understanding the instrument and developing better analysis pipelines
- Keep improving the simulations: both signal, foregrounds and instrumental effects
- **Prepare for the SKAO era and its potential contribution to the knowledge of large-scale structures**