Observing the Epoch of Reionization with LOFAR
Progress and challenges

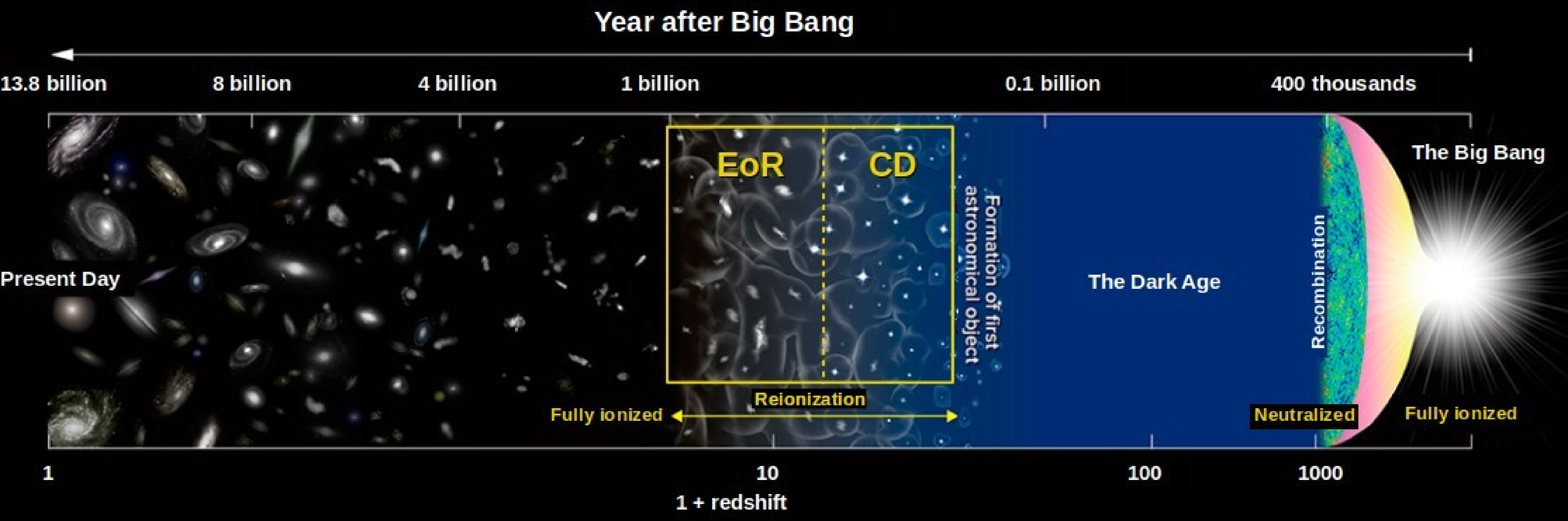
CoDA II — Pierre Ocvirk

Florent Mertens (LERMA / Kapteyn Astronomical Institute)

The LOFAR EoR KSP Team

Michiel Brentjens (ASTRON)
Wim Brouw (Kapteyn)
Emma Chapman (Imperial)
Benedetta Ciardi (MPA)
Keri Dixon (Sussex)
Simon Gazagnes (Kapteyn)
Hyoyin Gan (Kapteyn)
B.K. Geholot (ASU)
Abhik Ghosh (SKAO-SA)
Ilian Iliev (Sussex)
Vibor Jelic (IRB)
Hannes Jensen (Imperial)
Koki Kakiichi (MPG)
Robin Kooistra (Tokyo)
Léon Koopmans (Kapteyn)
F. Krause (Kapteyn/UCL)
Suman Majumdar (Imperial)
Garrelt Mellema (Stockholm)
Maaijke Mevius (ASTRON)
Florent Mertens (Kapteyn)
Modhurita Mitra (Kapteyn)
Rajesh Mondal (Sussex)
Kahn Asad (SKAO-SA)
André Offringa (ASTRON)
V.N Pandey (K./ASTRON)
Marta Silva (Oslo)
Joop Schaye (Leiden)
M. Sardarabadi (Kapteyn)
H. Vedantham (ASTRON)
Stefan Wijnholds (ASTRON)
Sarod Yatawatta (ASTRON)
Saleem Zaroubi (K./Haifa)

Cosmic Dawn / Epoch of Reionization



Credit: NAOJ

Epoch of Reionization

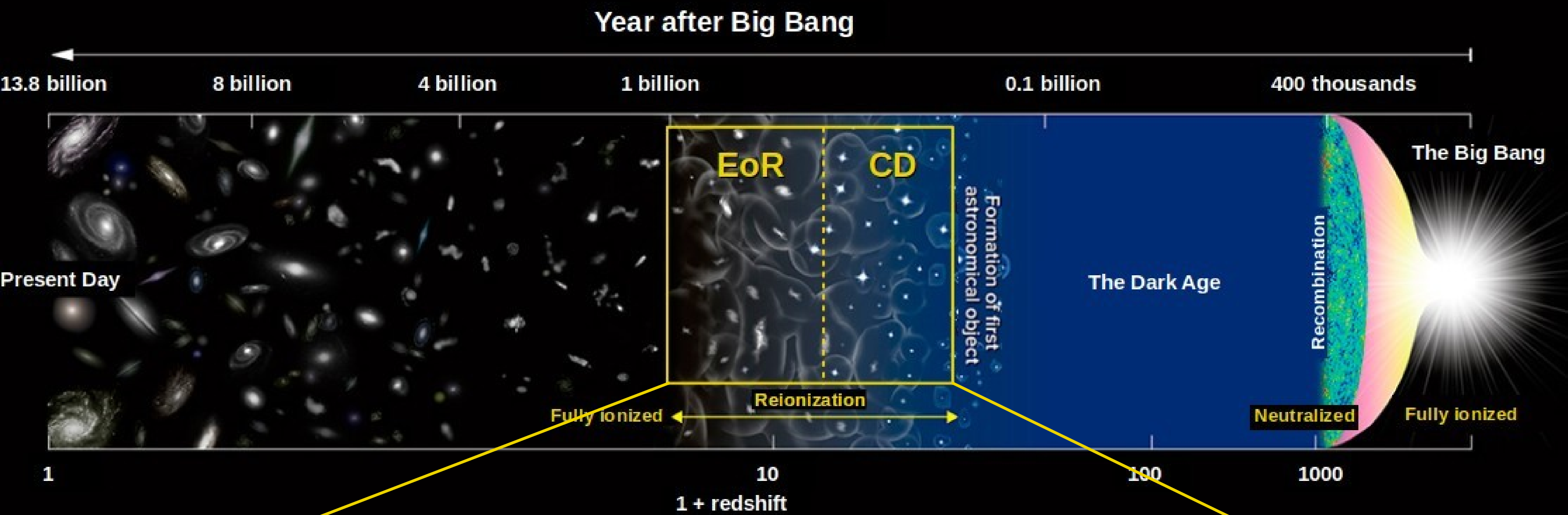
- Reionization by stars & mini-quasars
- IGM feedback (e.g. metals)
- PopIII - PopII transition
- Emergence of the visible universe

Cosmic Dawn

- Appearance of first stars/Bhs (PopIII?)
- Ly-α radiation field
- Impact of Baryonic Bulk Flows
- First X-ray heating sources

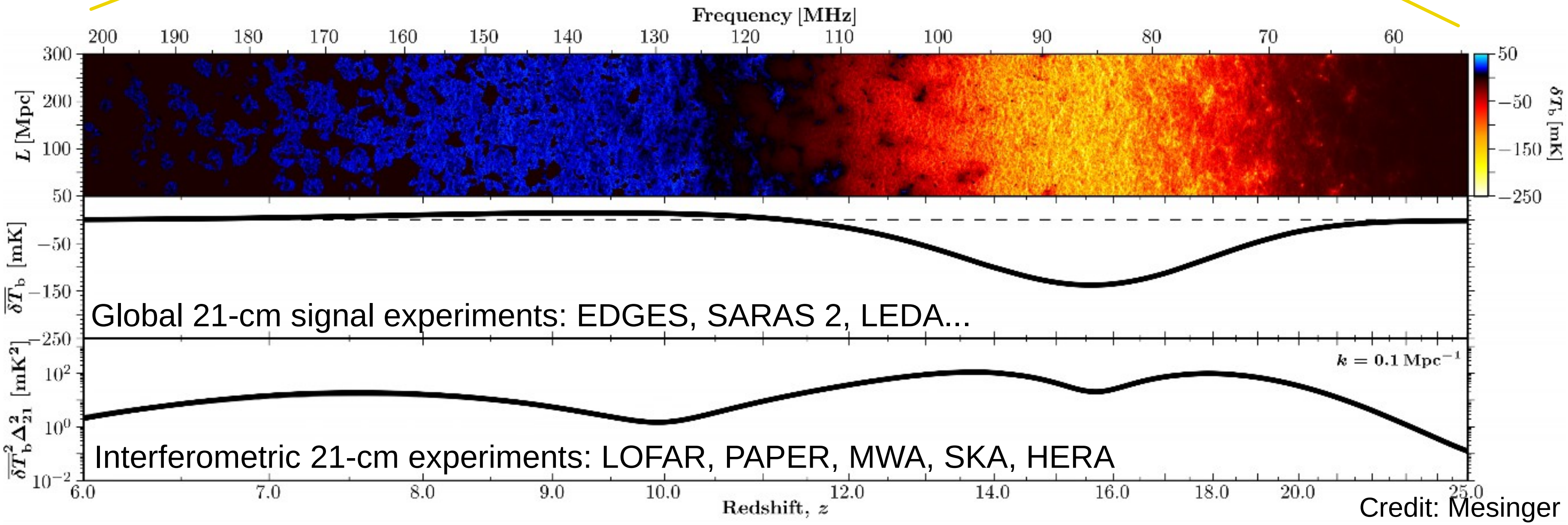
- When did the first galaxies/stars/black hole form?
- How did reionization proceed?
- How do galaxies form and evolve?

Cosmic Dawn / Epoch of Reionization



High-z HI 21-cm signal unique probe of the CD/EoR

Credit: NAOJ

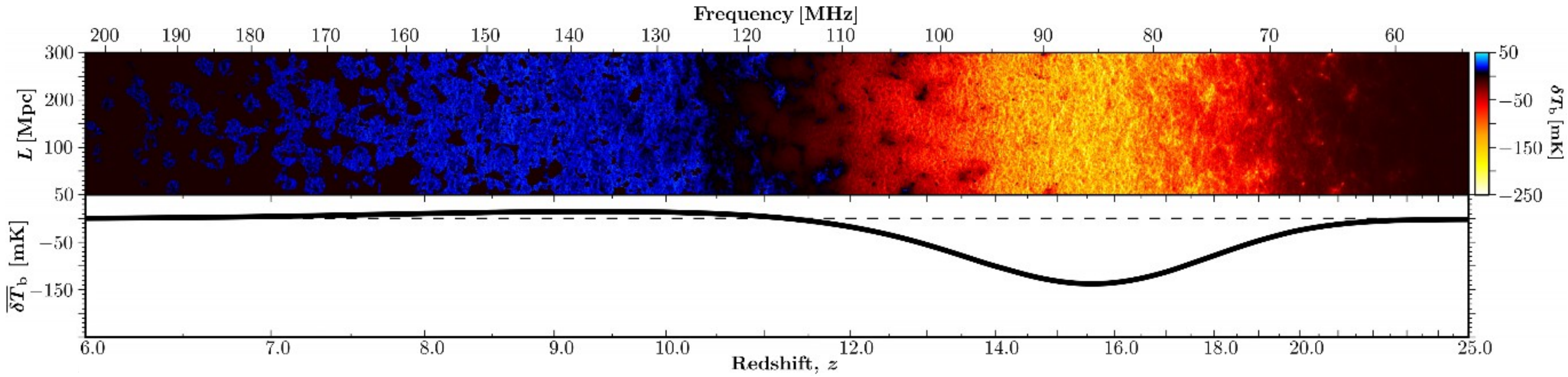


Global 21-cm signal experiments: EDGES, SARAS 2, LEDA...

Interferometric 21-cm experiments: LOFAR, PAPER, MWA, SKA, HERA

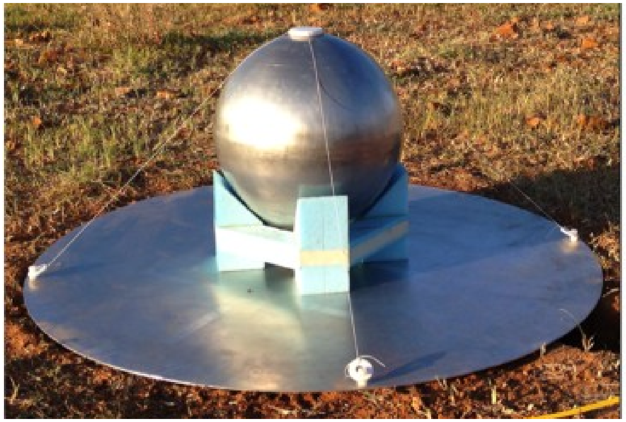
Credit: Mesinger & Greig

The Global experiments



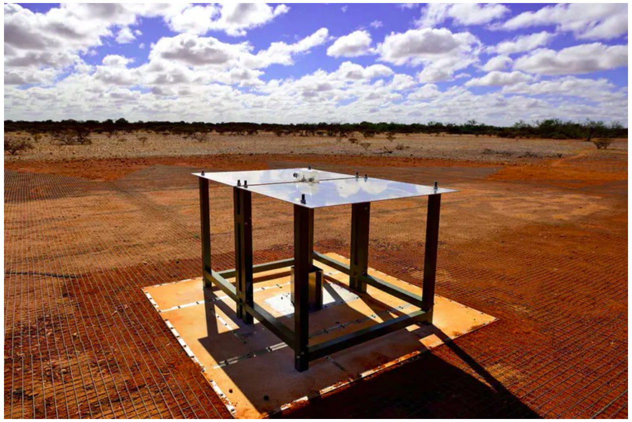
PRIZM
30-200 MHz
Marion Island

Peterson, Sievers, Chiang ++



SARAS
50-100, 100-200 MHz
India (Himalayas)

Singh et al. 2017

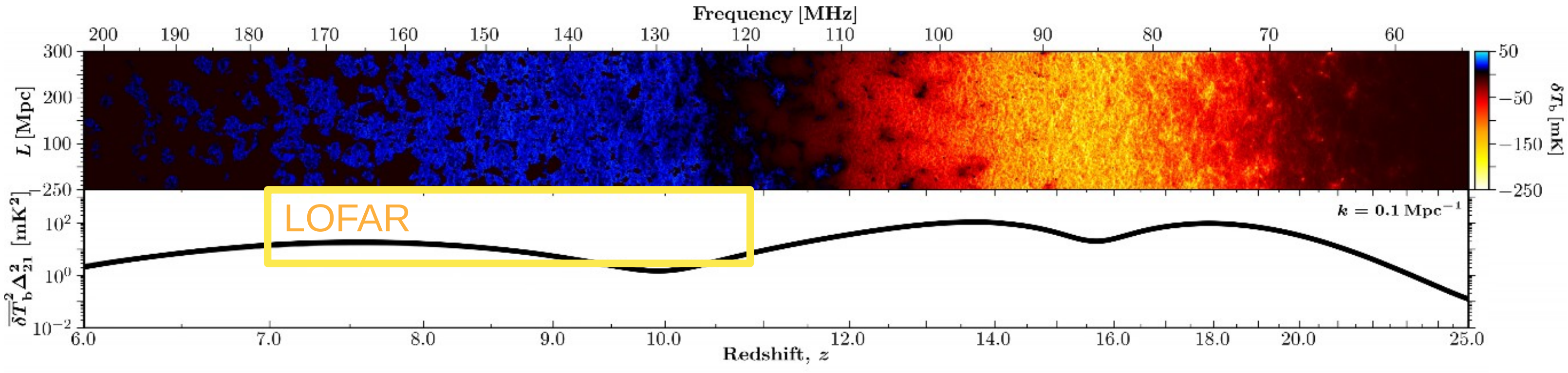


EDGES
50-100, 100-200 MHz
Western Australia

Rogers & Bowman 2008, 2012;
Bowman et al 2018

+ Many more

The Interferometric experiments

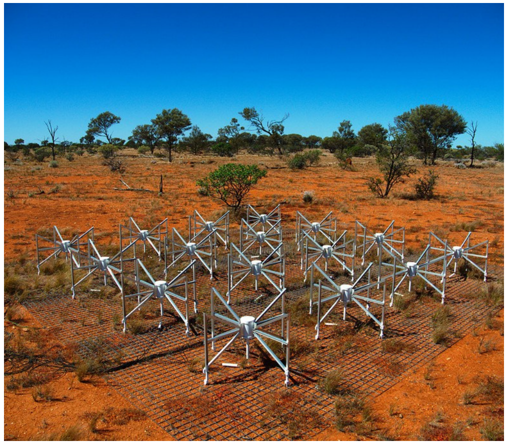


GMRT
India

40 h @ z ~ 8.5
Paciga et al. 2013

PAPER
South Africa

1148 h @ z=8.4
~~Ali et al. 2015~~
Kolopanis et al. 2019



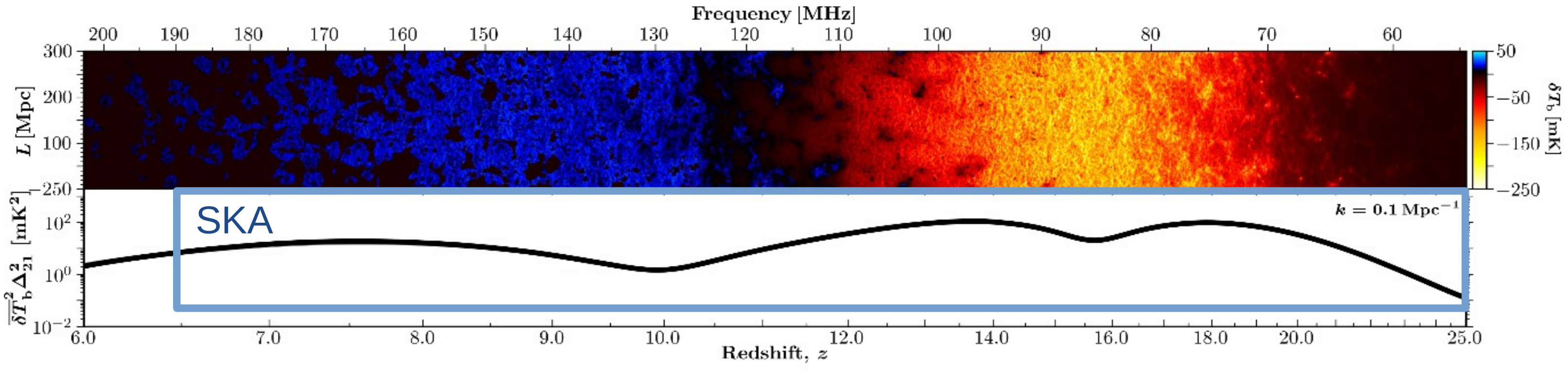
MWA
Western Australia

z ~ 6 – 10
~ 32 h published
Beardsley et al. 2016
Barry et al. 2019
MWA phase 2

LOFAR
The Netherlands

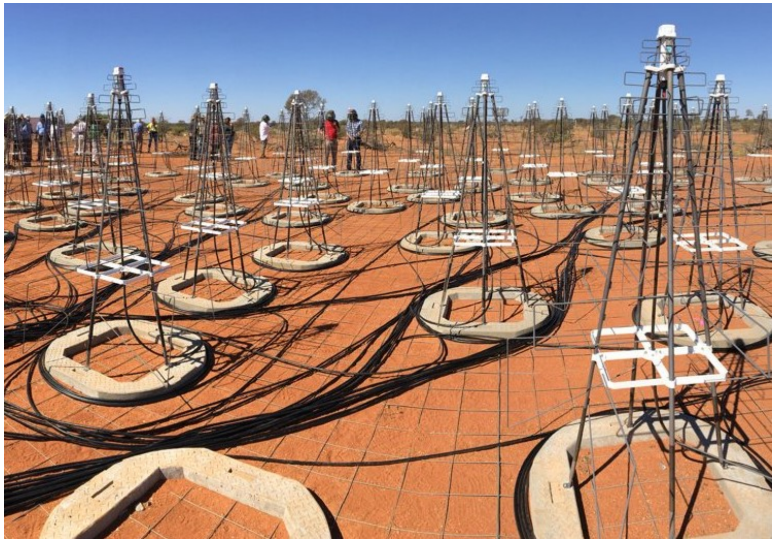
z ~ 7 – 11
+ 2000 h observed
13h published
Patil et al. 2017
140h in prep.

The Interferometric experiments



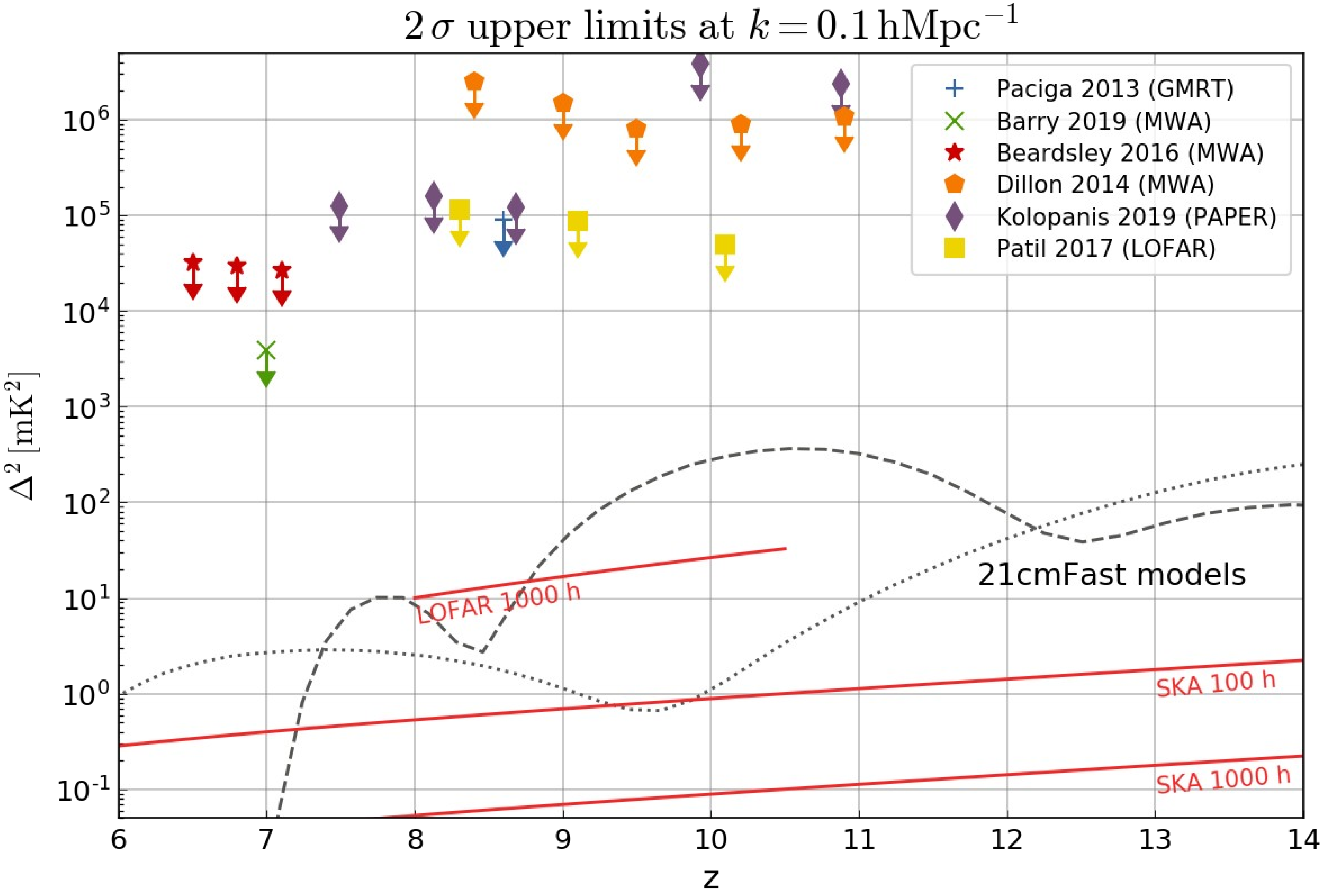
Second generation experiments in near and far future

HERA
South Africa
z ~ 6 – 25
240 dishes of 14 m (by ~ 2020)
In (partial) commissioning

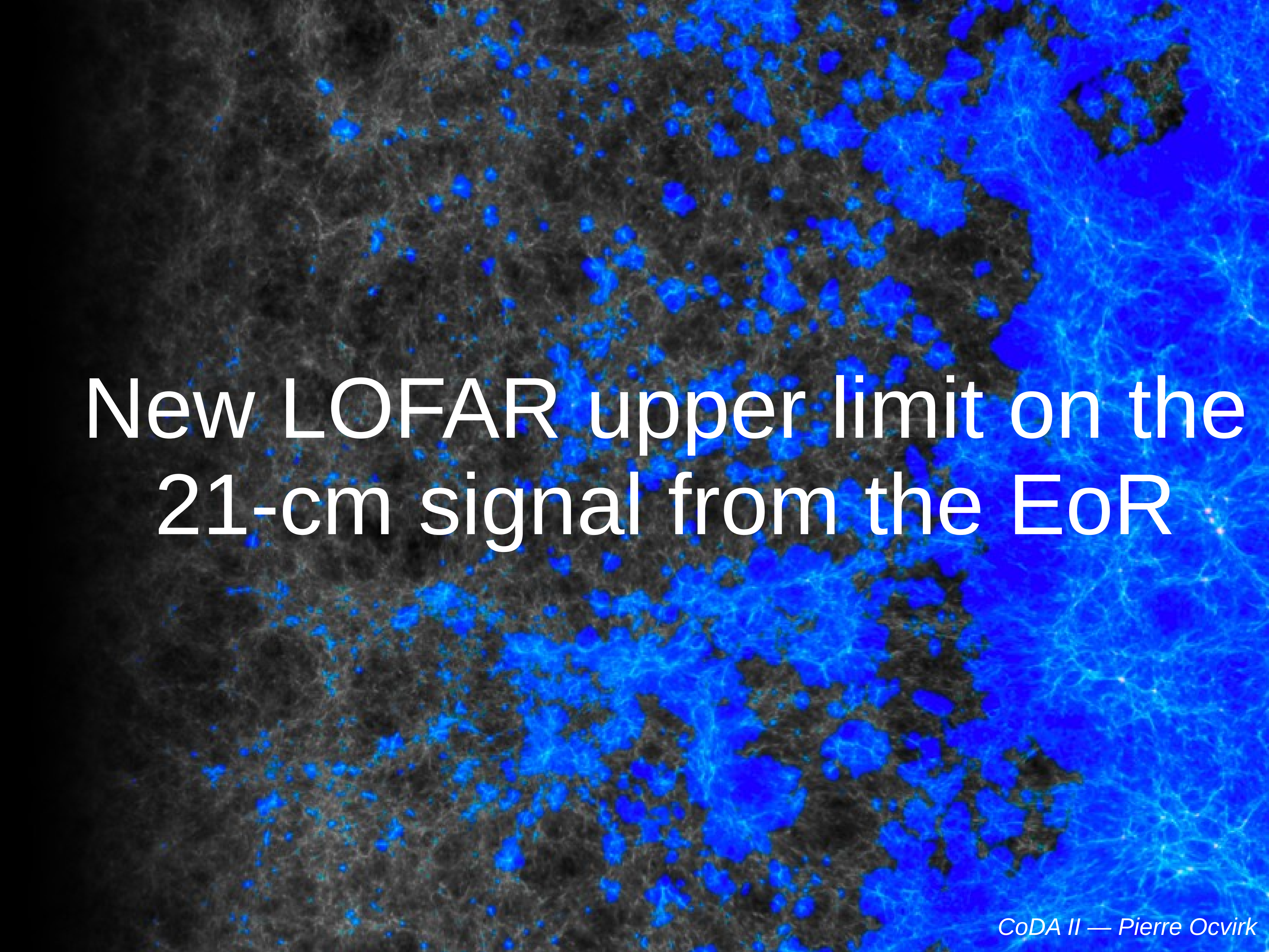


SKA
Western Australia
Low band (z ~ 6 – 25)
Construction 2020-2025

Where do we stand ?

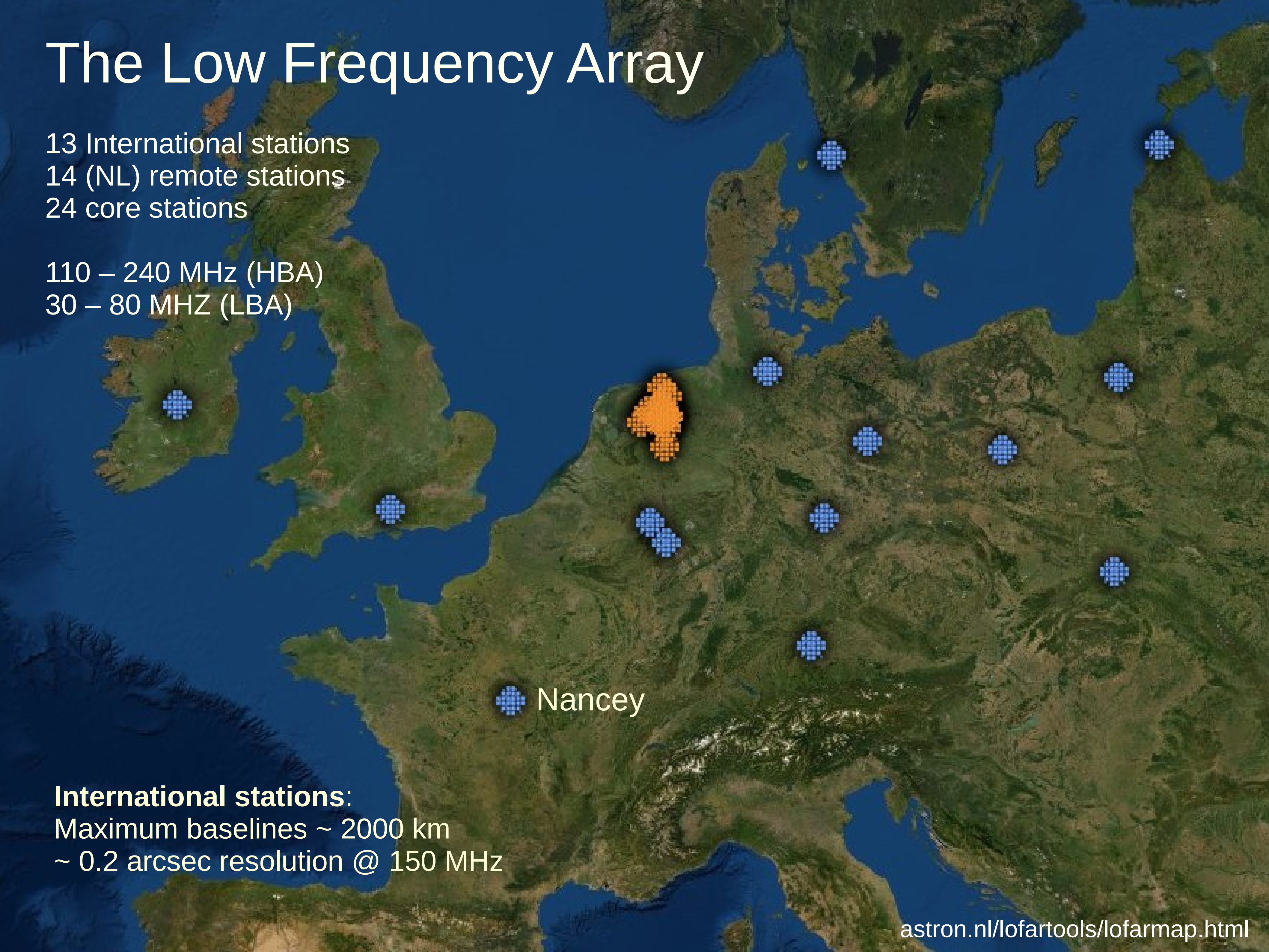


New LOFAR upper limit on the 21-cm signal from the EoR



CoDA II — Pierre Ocvirk

The Low Frequency Array



13 International stations
14 (NL) remote stations
24 core stations

110 – 240 MHz (HBA)
30 – 80 MHZ (LBA)

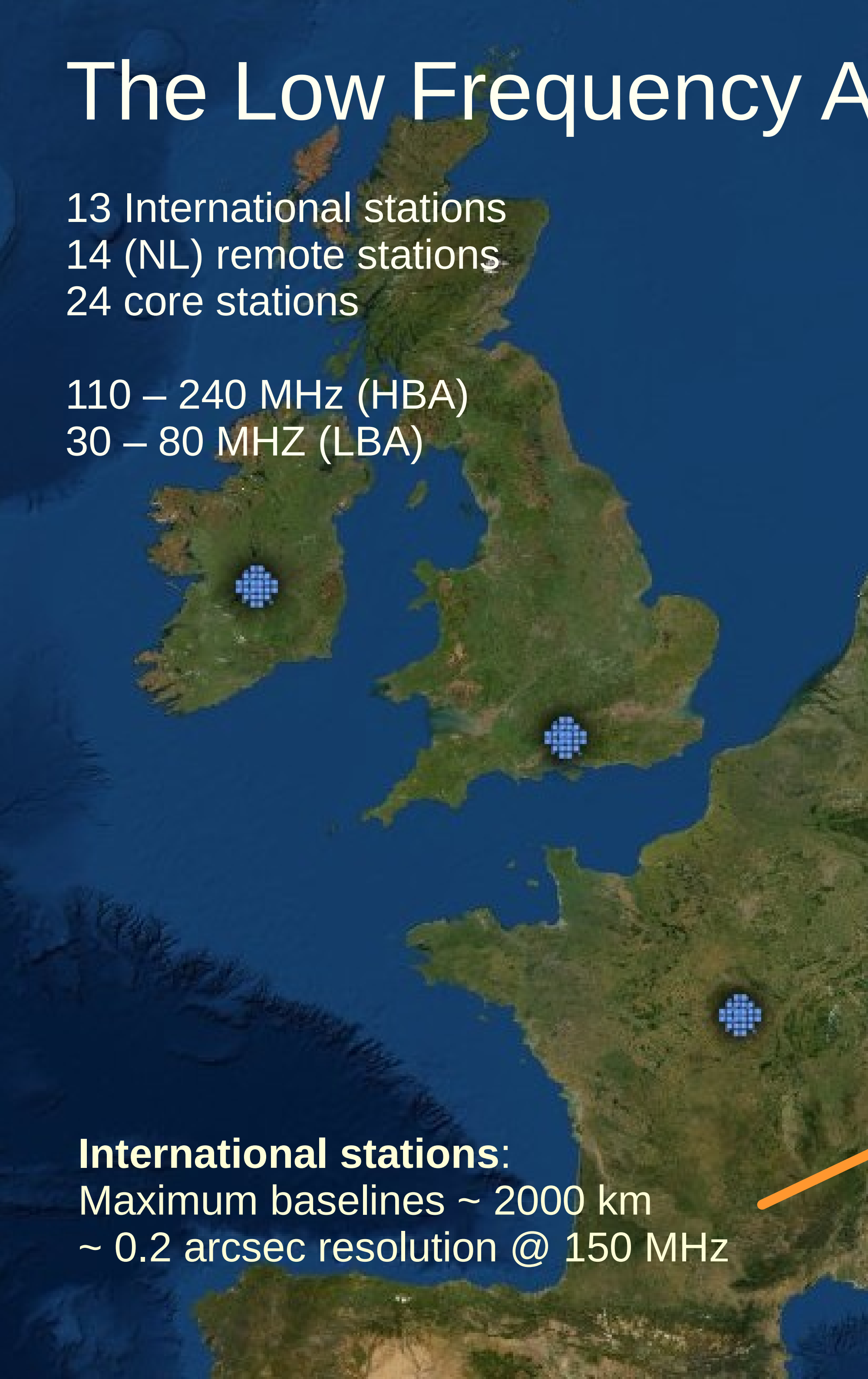
International stations:
Maximum baselines ~ 2000 km
~ 0.2 arcsec resolution @ 150 MHz

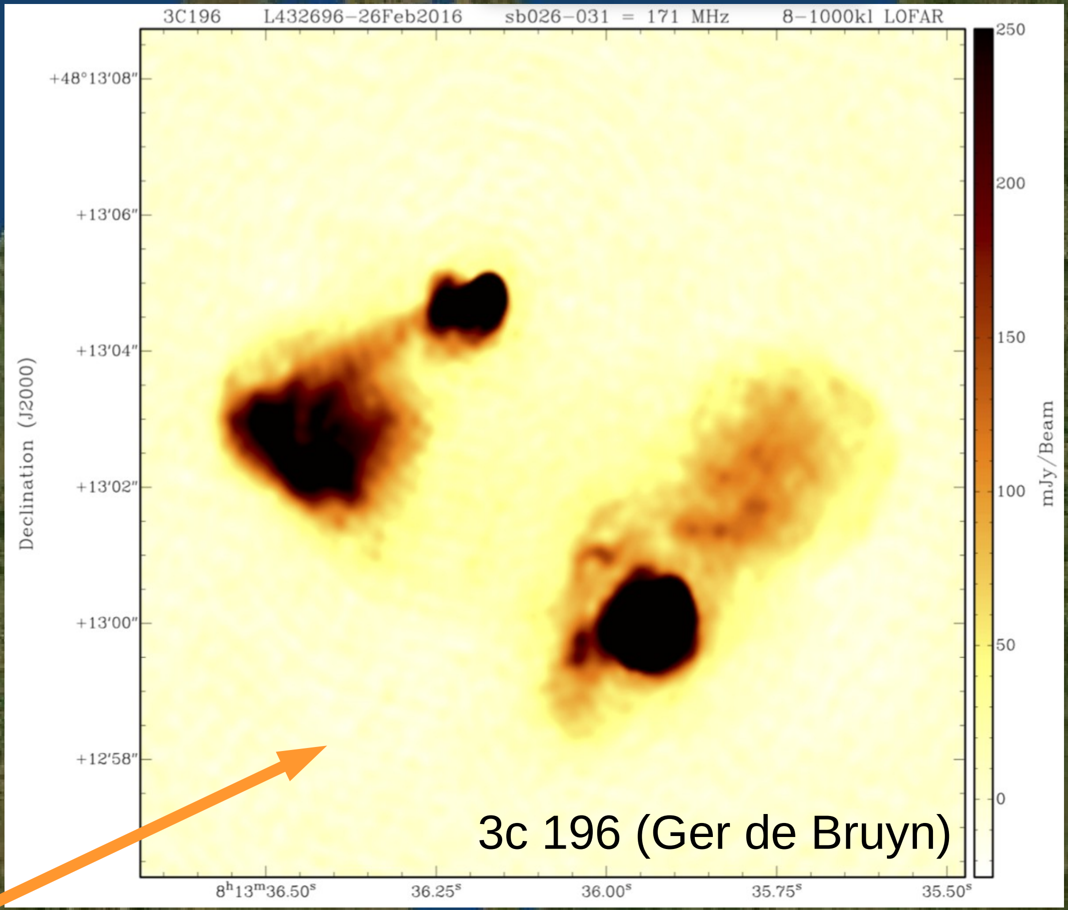
astron.nl/lofartools/lofarmap.html

The Low Frequency Array

13 International stations
14 (NL) remote stations
24 core stations

110 – 240 MHz (HBA)
30 – 80 MHZ (LBA)



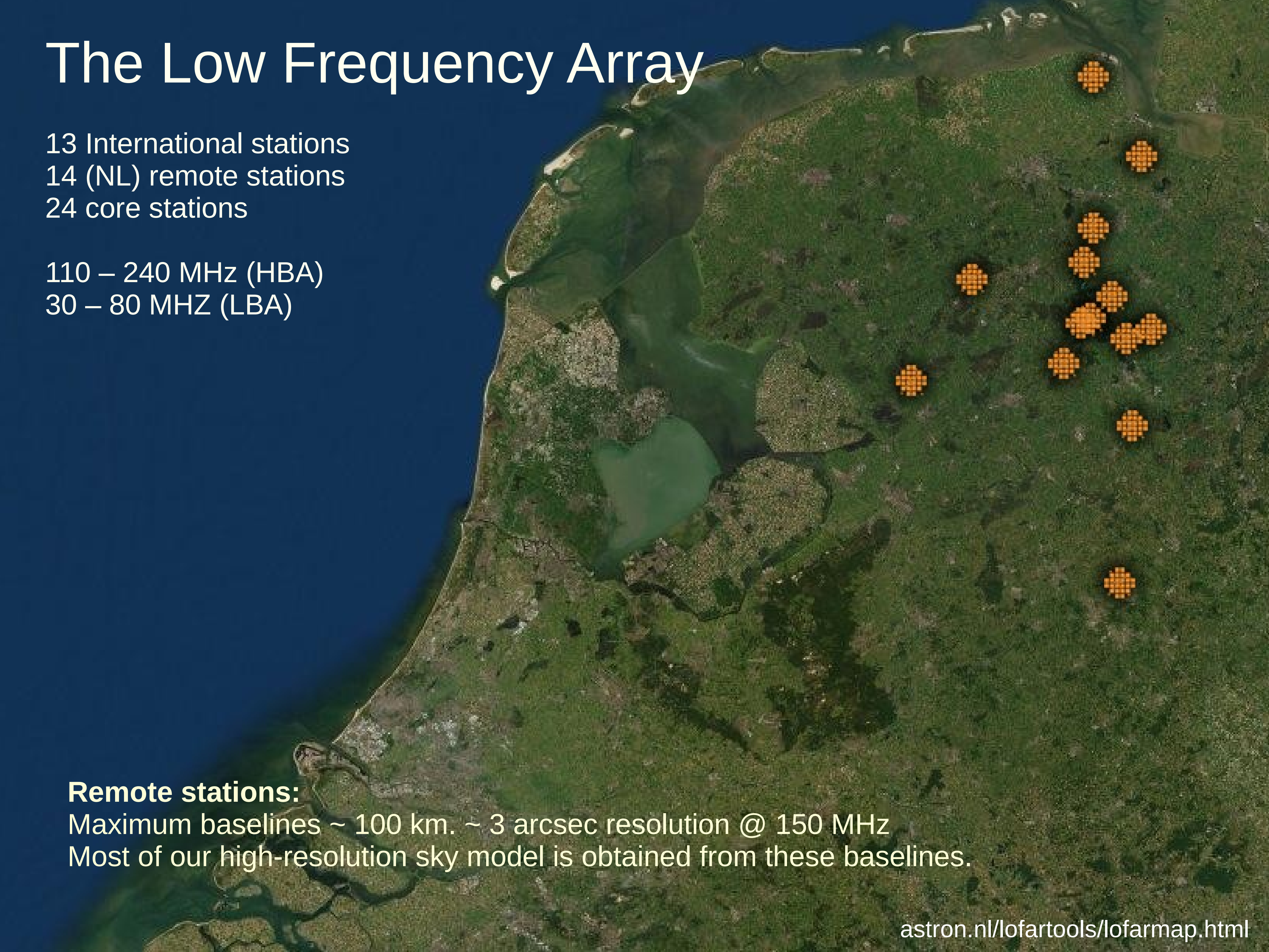


3c 196 (Ger de Bruyn)

International stations:
Maximum baselines ~ 2000 km
~ 0.2 arcsec resolution @ 150 MHz

astron.nl/lofartools/lofarmap.html

The Low Frequency Array



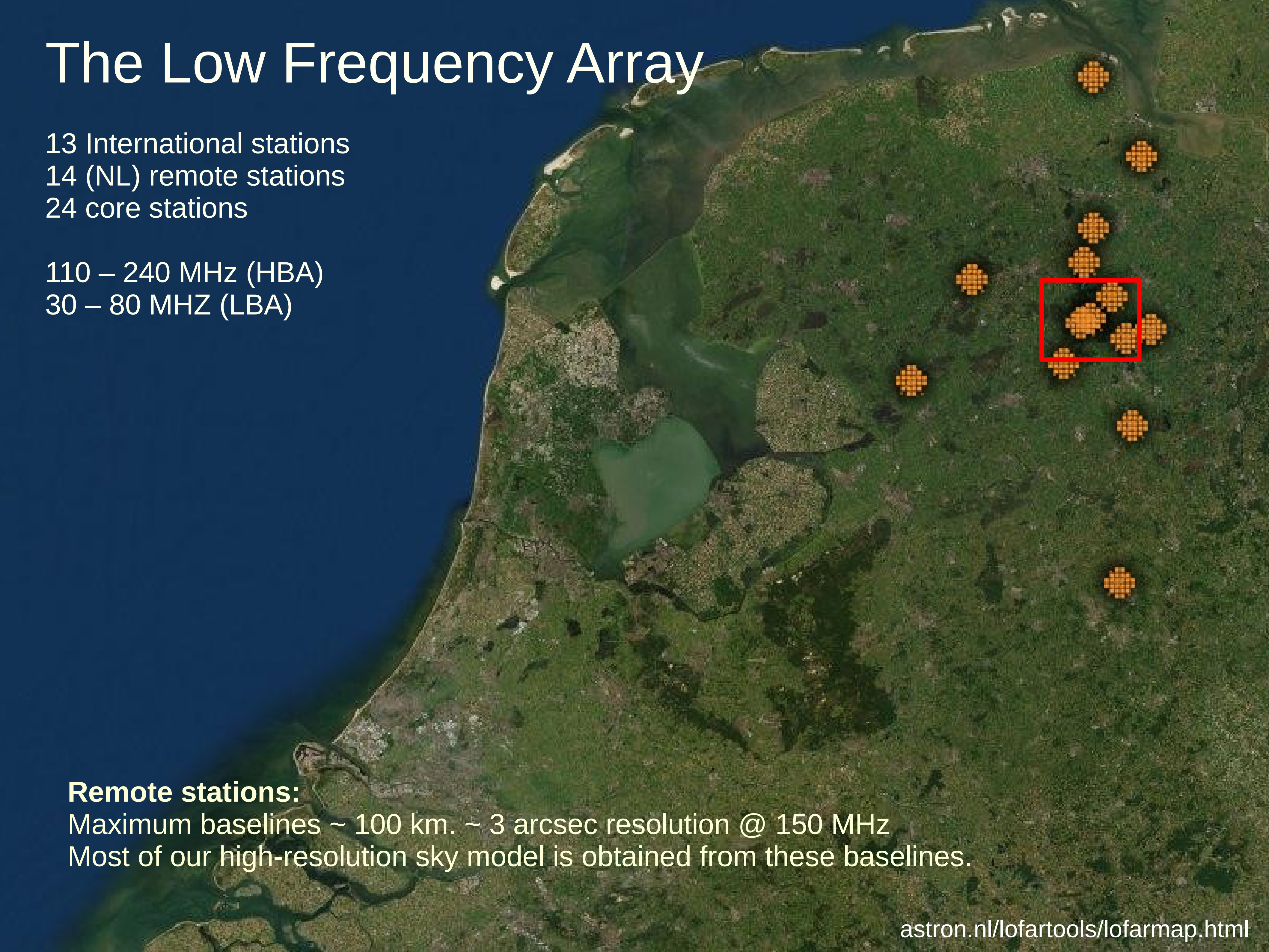
13 International stations
14 (NL) remote stations
24 core stations

110 – 240 MHz (HBA)
30 – 80 MHZ (LBA)

Remote stations:
Maximum baselines ~ 100 km. ~ 3 arcsec resolution @ 150 MHz
Most of our high-resolution sky model is obtained from these baselines.

astron.nl/lofartools/lofarmap.html

The Low Frequency Array



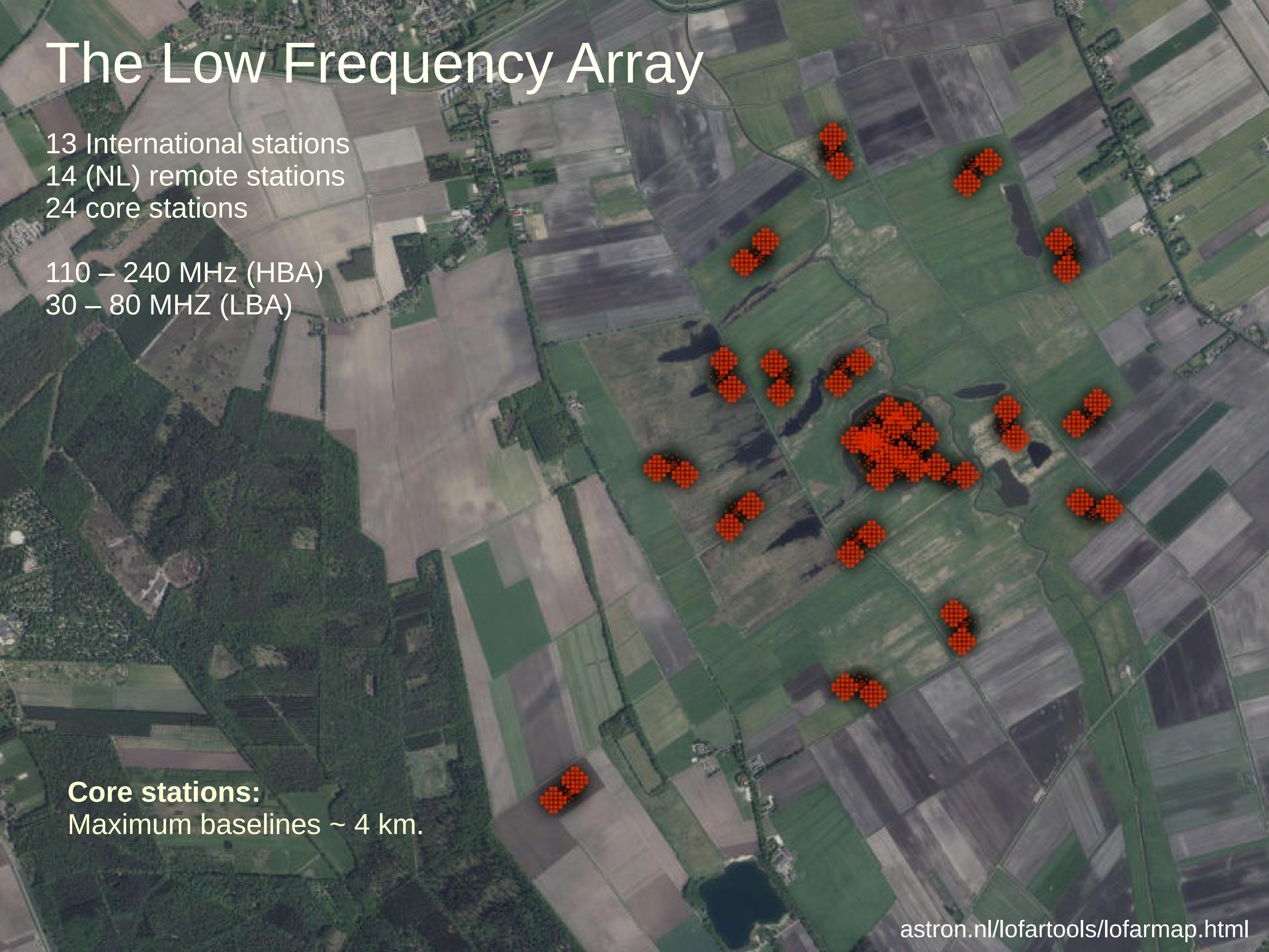
13 International stations
14 (NL) remote stations
24 core stations

110 – 240 MHz (HBA)
30 – 80 MHZ (LBA)

Remote stations:
Maximum baselines ~ 100 km. ~ 3 arcsec resolution @ 150 MHz
Most of our high-resolution sky model is obtained from these baselines.

astron.nl/lofartools/lofarmap.html

The Low Frequency Array



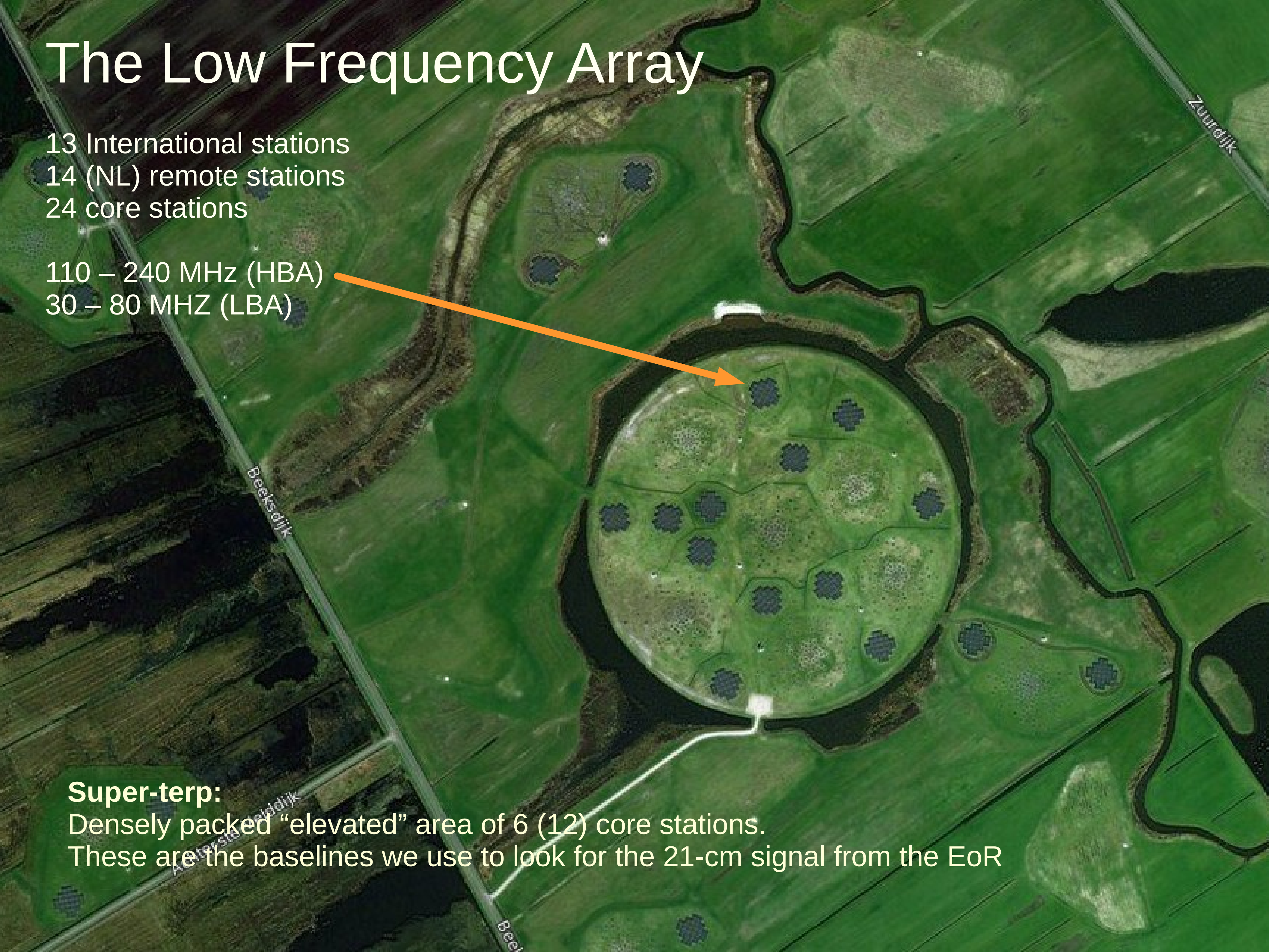
13 International stations
14 (NL) remote stations
24 core stations

110 – 240 MHz (HBA)
30 – 80 MHZ (LBA)

Core stations:
Maximum baselines ~ 4 km.

astron.nl/lofartools/lofarmap.html

The Low Frequency Array



13 International stations
14 (NL) remote stations
24 core stations

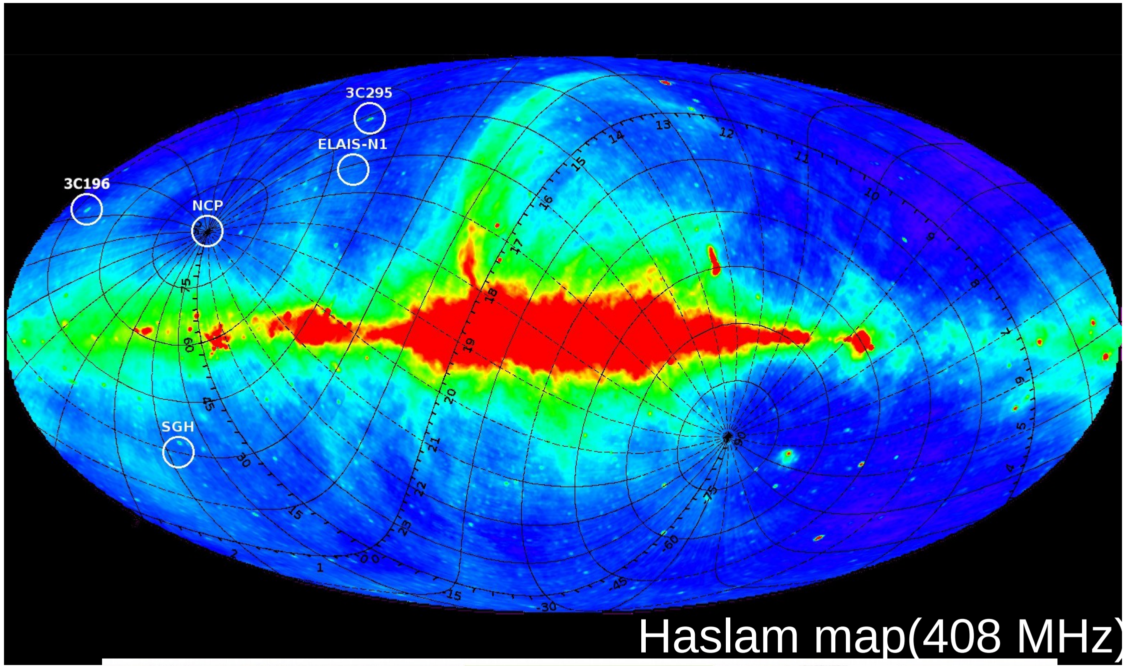
110 – 240 MHz (HBA)
30 – 80 MHZ (LBA)

Super-terp:
Densely packed “elevated” area of 6 (12) core stations.
These are the baselines we use to look for the 21-cm signal from the EoR

The LOFAR-EoR KSP

2 main targets

- North Celestial Pole
  - ✓ Constant Beam, all year observable
  - ✓ + 2200 hours observed
- 3C 196
  - ✓ Bright calibrator
  - ✓ + 1100 hours observed



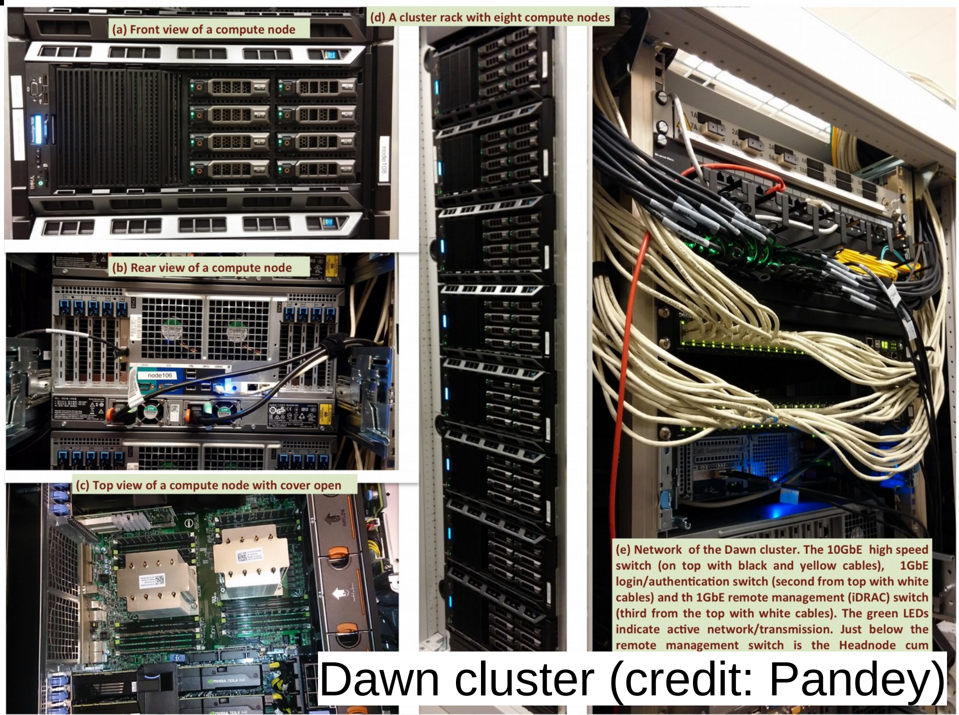
Haslam map(408 MHz)

2-3 other windows for various other projects

Raw data volume: 20-70 TB / night
Archived data: > 5 PB

Dawn cluster:

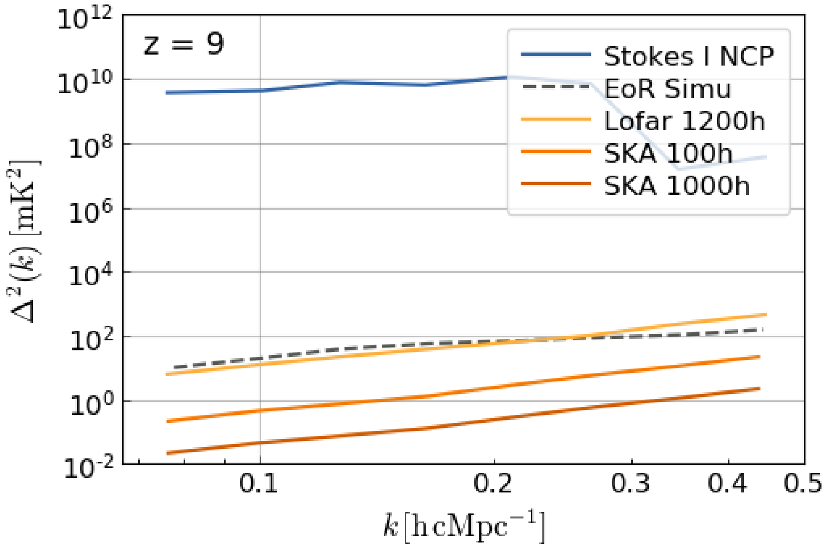
- ✓ 32 nodes
- ✓ each with 48 CPU cores + 4 GPU



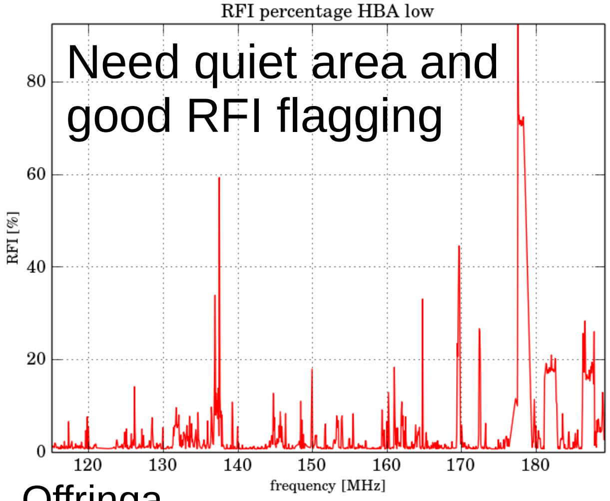
Dawn cluster (credit: Pandey)

What make this experiment so challenging ?

Foregrounds



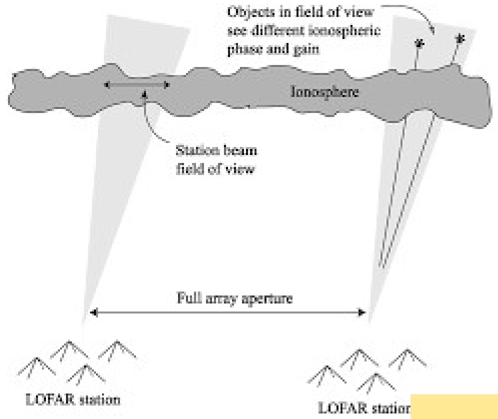
Radio Frequency Interferance (RFI)



Need quiet area and good RFI flagging

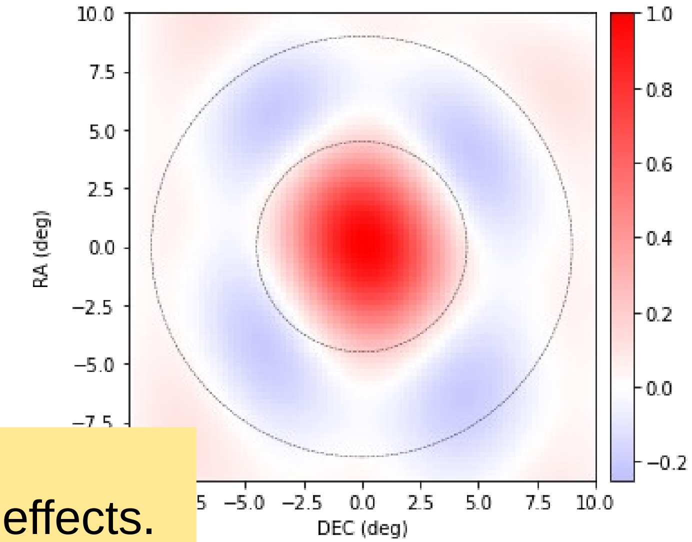
Credit: A. Offringa

Ionosphere



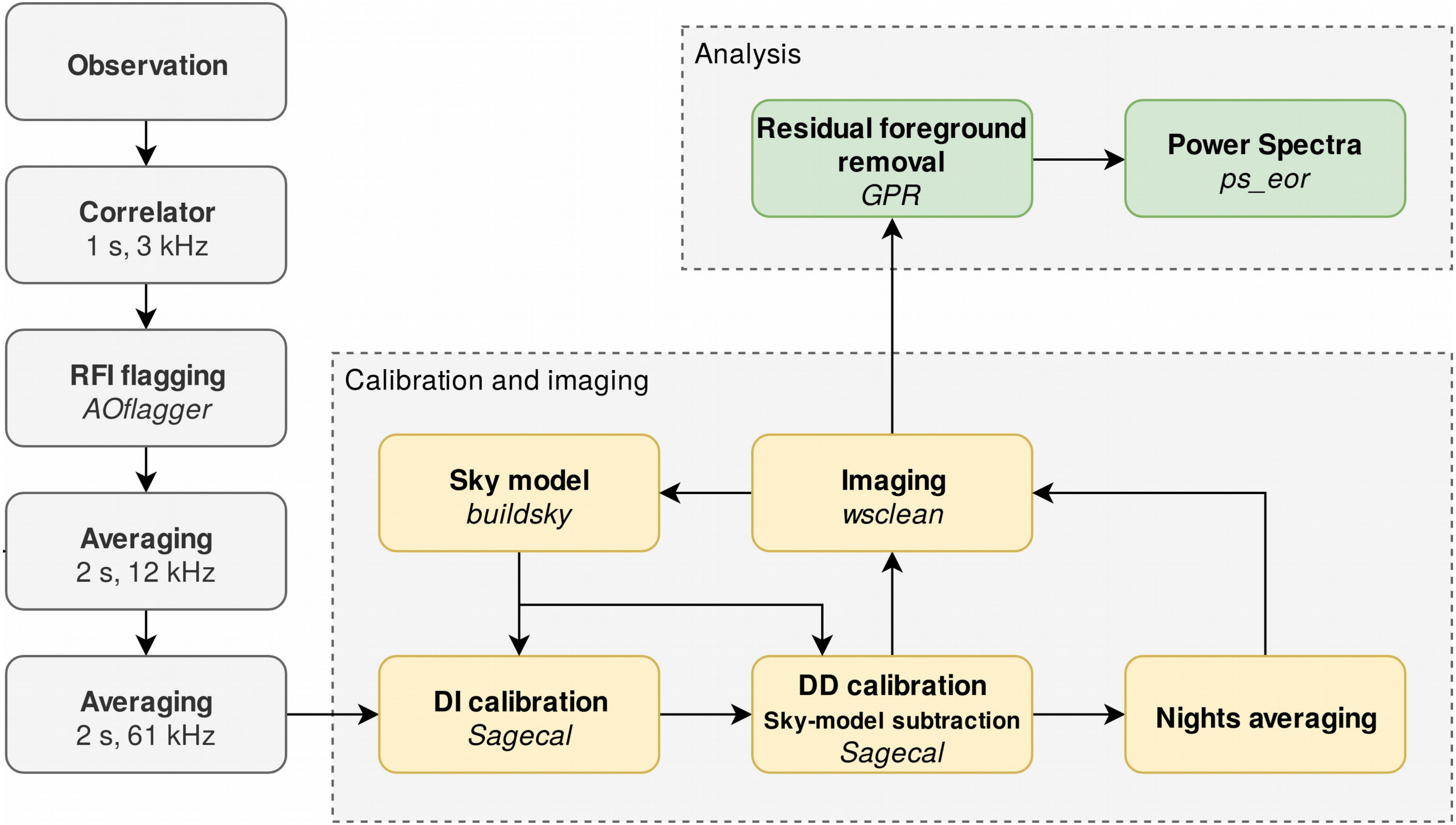
Credit: S. Van der Tol

Primary beam



+ polarization leakage.
All direction dependent effects.

(Simplified) Processing Pipeline

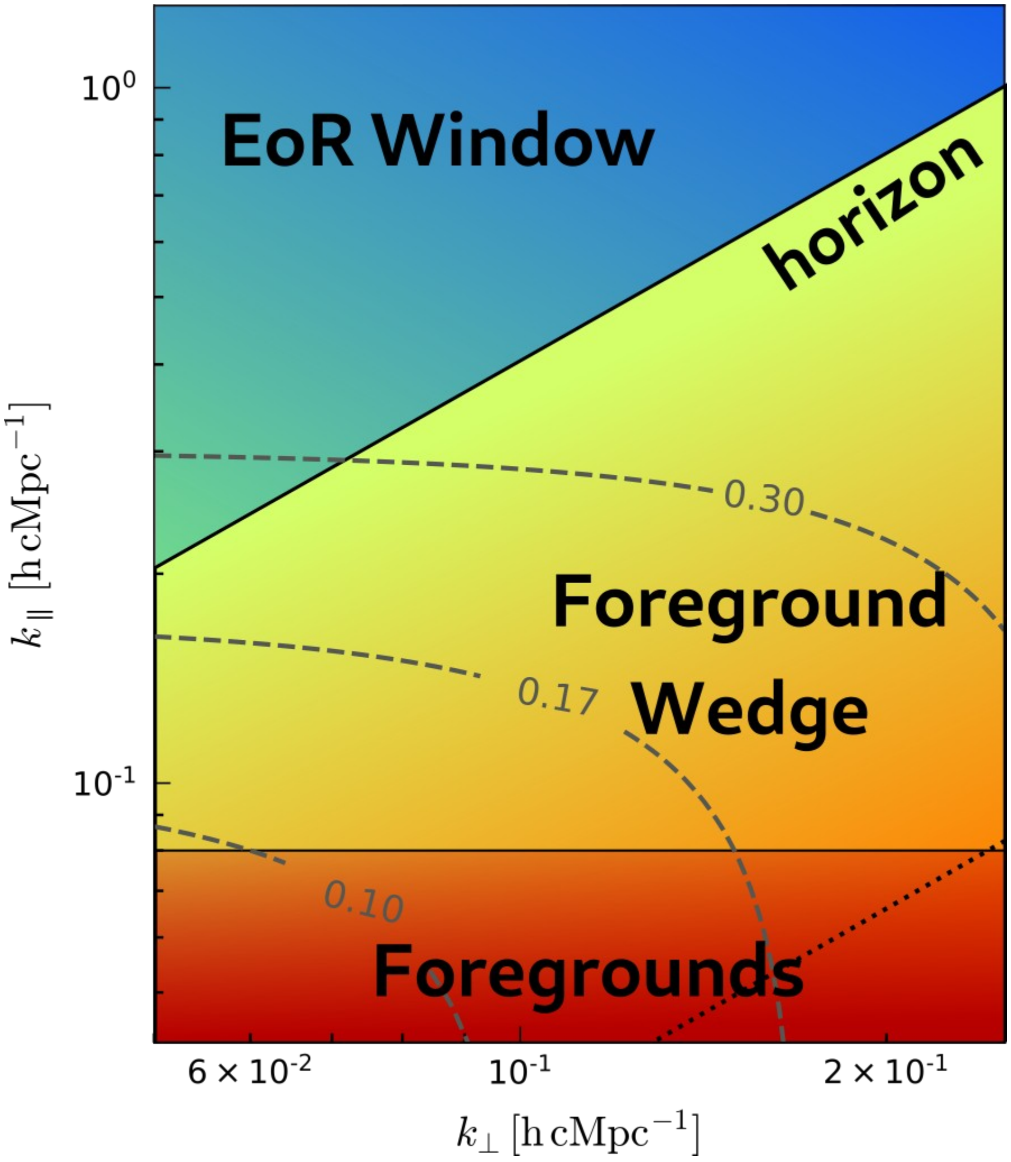


The challenge of the foregrounds

21-cm signal:
- Uncorrelated ~ MHz
- Isotropic

Foreground emission:
- Mainly synchrotron and free-free emission
- Smooth in frequency

Foreground Wedge:
- ➔ Chromatic instrument (beam/uv-coverage)
- ➔ Ionosphere
- ➔ Calibration error
- ➔ Polarization leakage



Spatial vs line-of-sight power-spectra

Removing the foregrounds

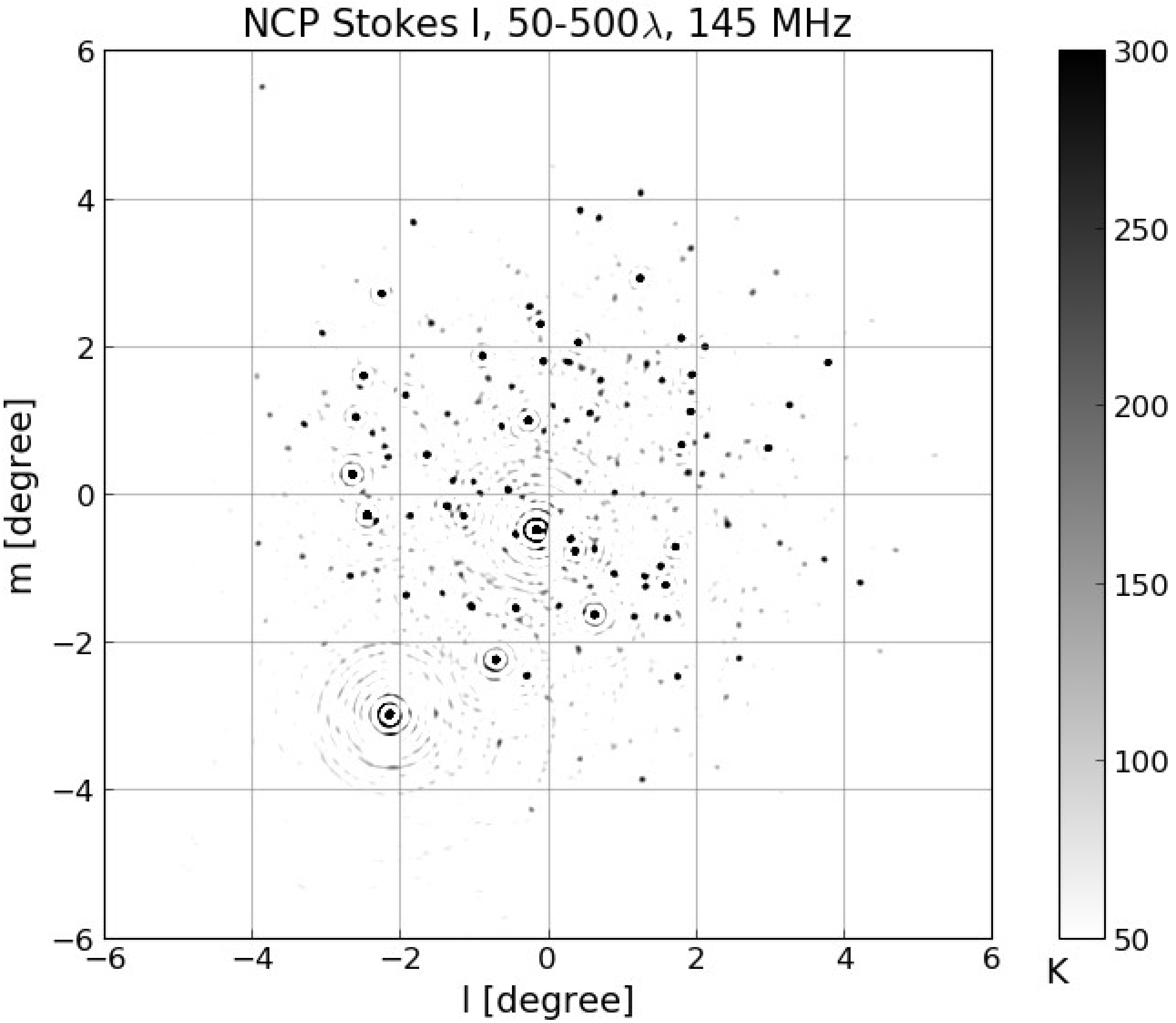
Step 1:
Point-sources subtraction

- ➔ Need accurate sky-model
- ➔ Solve for instruments gains in direction of sources

Direction Dependent (DD) calibration using Sagecal-CO (Yattawatta et al. 2013, 1015, ...)

Step 2:
Residual spectrally-smooth foregrounds subtraction

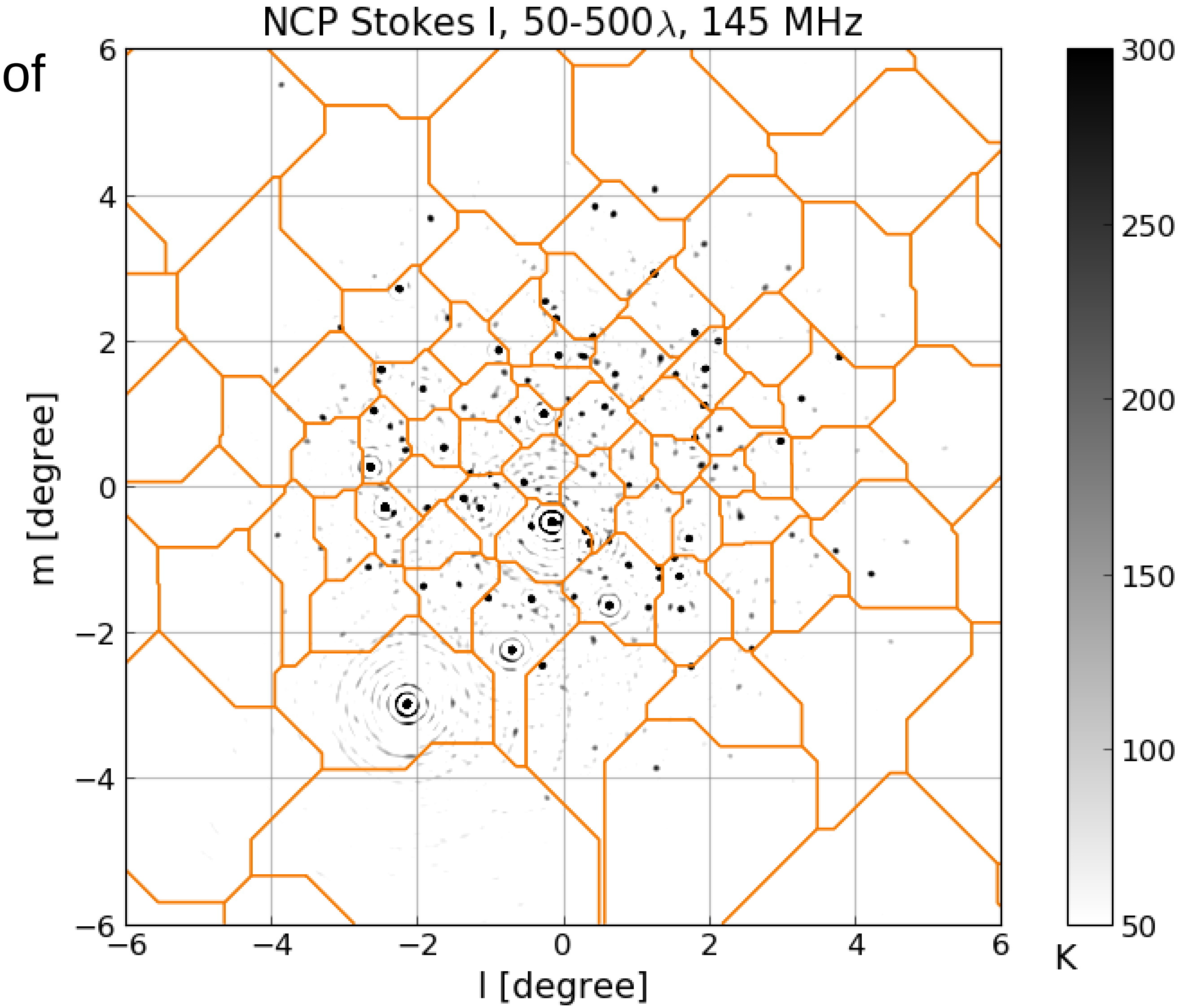
Using e.g. Gaussian Process Regression (GPR) (Mertens et al. 2018)



Direction Dependent calibration

Need to reduce the number of degree of freedom:

- ➔ Clustering (NCP ~ 120 clusters)



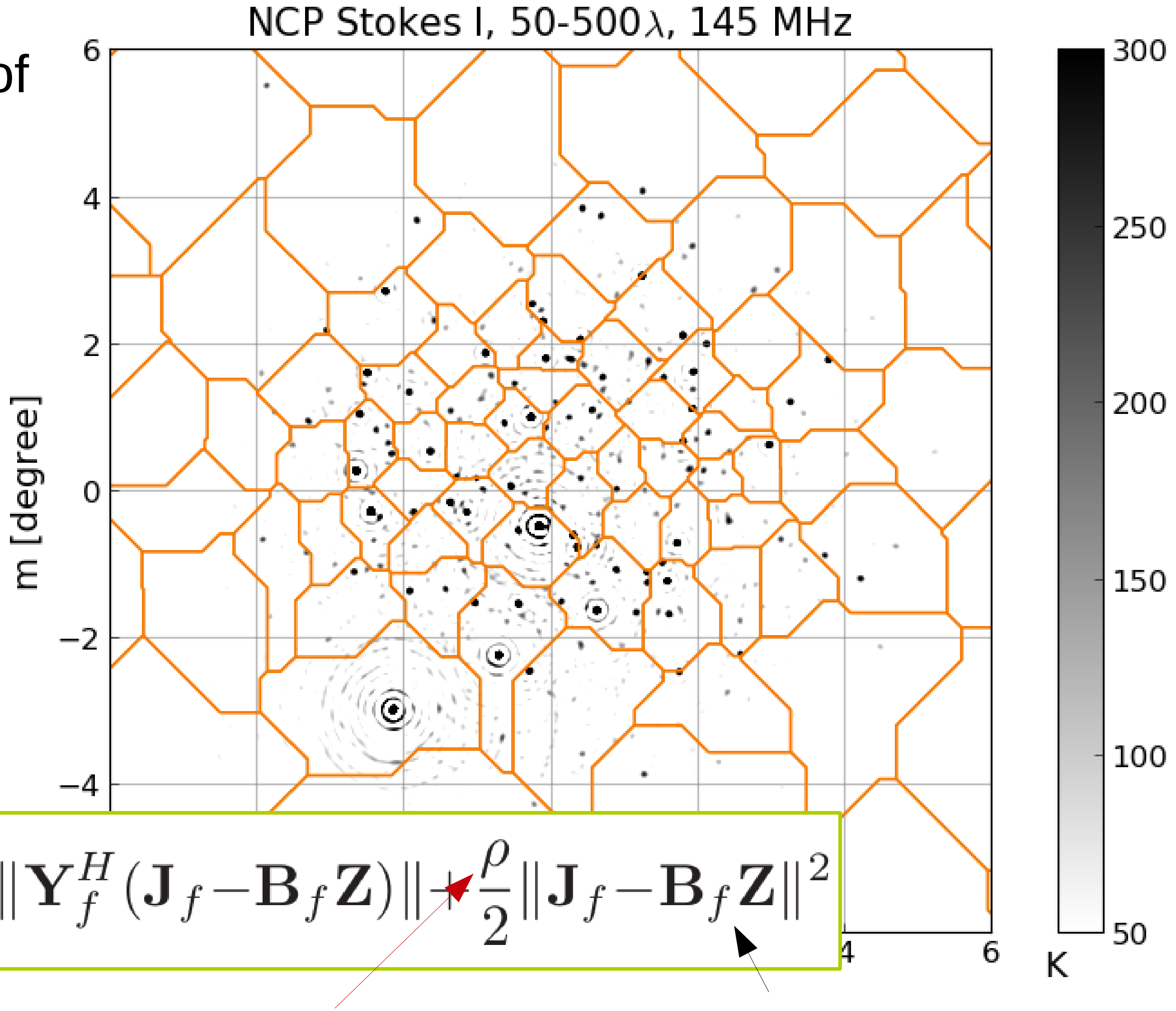
(Yatawatta et al. 2013, 2015)

Direction Dependent calibration

Need to reduce the number of degree of freedom:

- ➔ Clustering (NCP ~ 120 clusters)
- ➔ Force spectrally-smooth instrumental response

Sagecal-CO: distributed calibration, solve augmented Lagrangian:



$$L_f(\mathbf{J}_f,\mathbf{Z},\mathbf{Y}_f) = g_f(\mathbf{J}_f) + \|\mathbf{Y}_f^H(\mathbf{J}_f-\mathbf{B}_f\mathbf{Z})\| + \frac{\rho}{2}\|\mathbf{J}_f-\mathbf{B}_f\mathbf{Z}\|^2$$

Gains

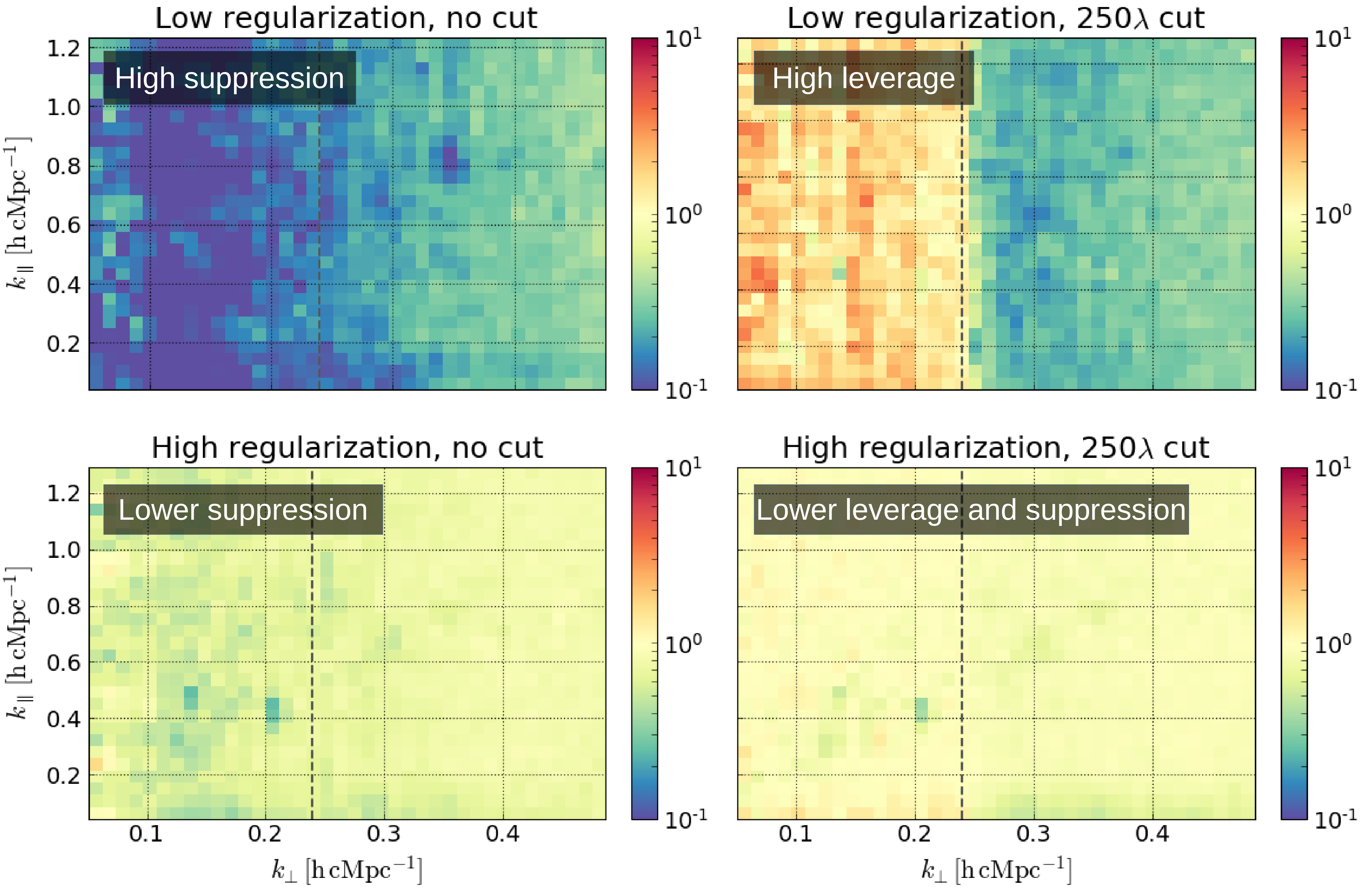
Original cost function

Regularization parameter

Spectrally-smooth constraint

(Yatawatta et al. 2013, 2015)

DD calibration: effect of enforcing smoothness

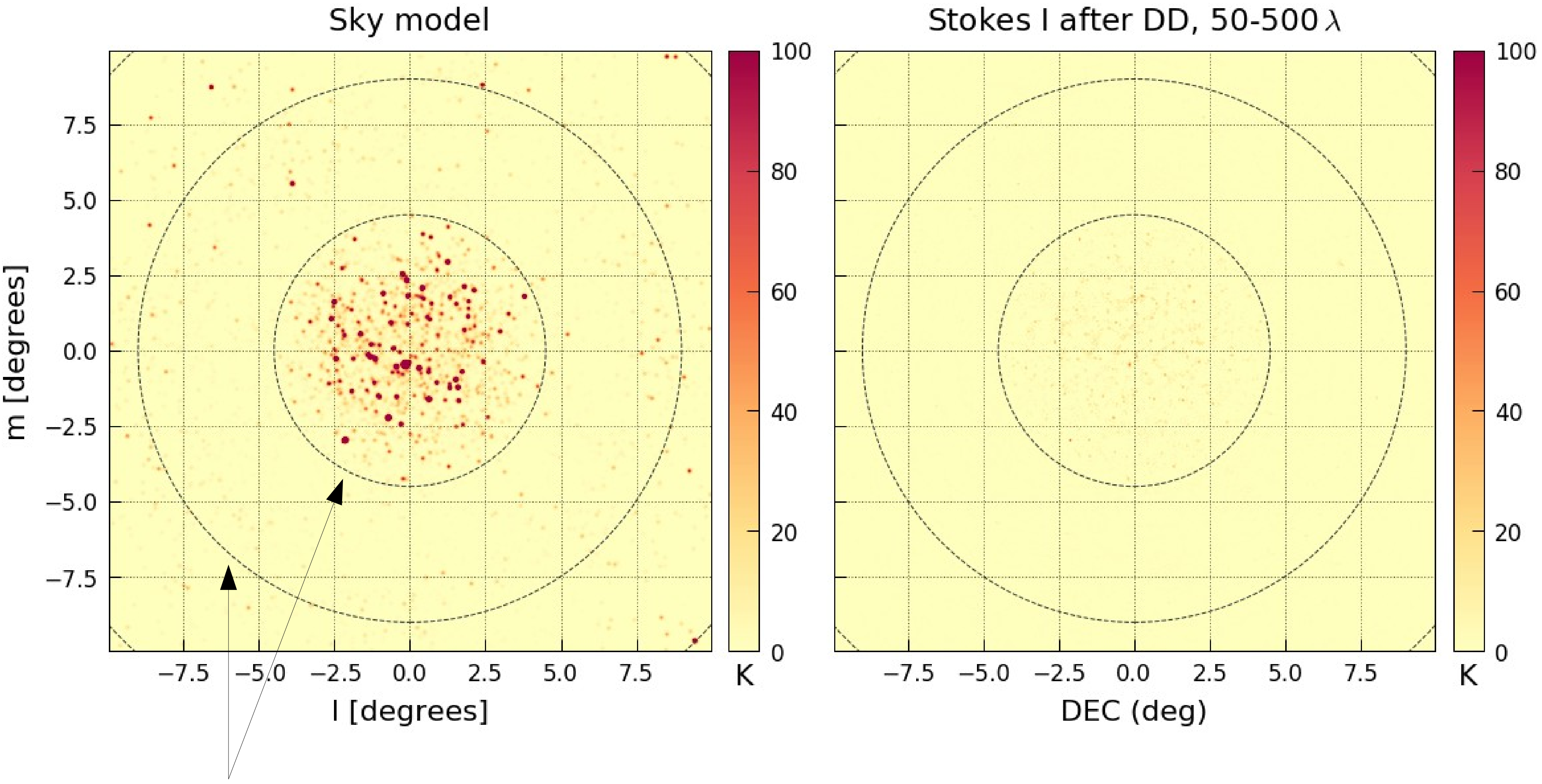


Ratio of Stokes V before/after DD show the effect of enforcing frequency-smoothness

(Mevius, Mertens, et al in prep.)
See also Sardarabadi et al. 2018

DD calibration results

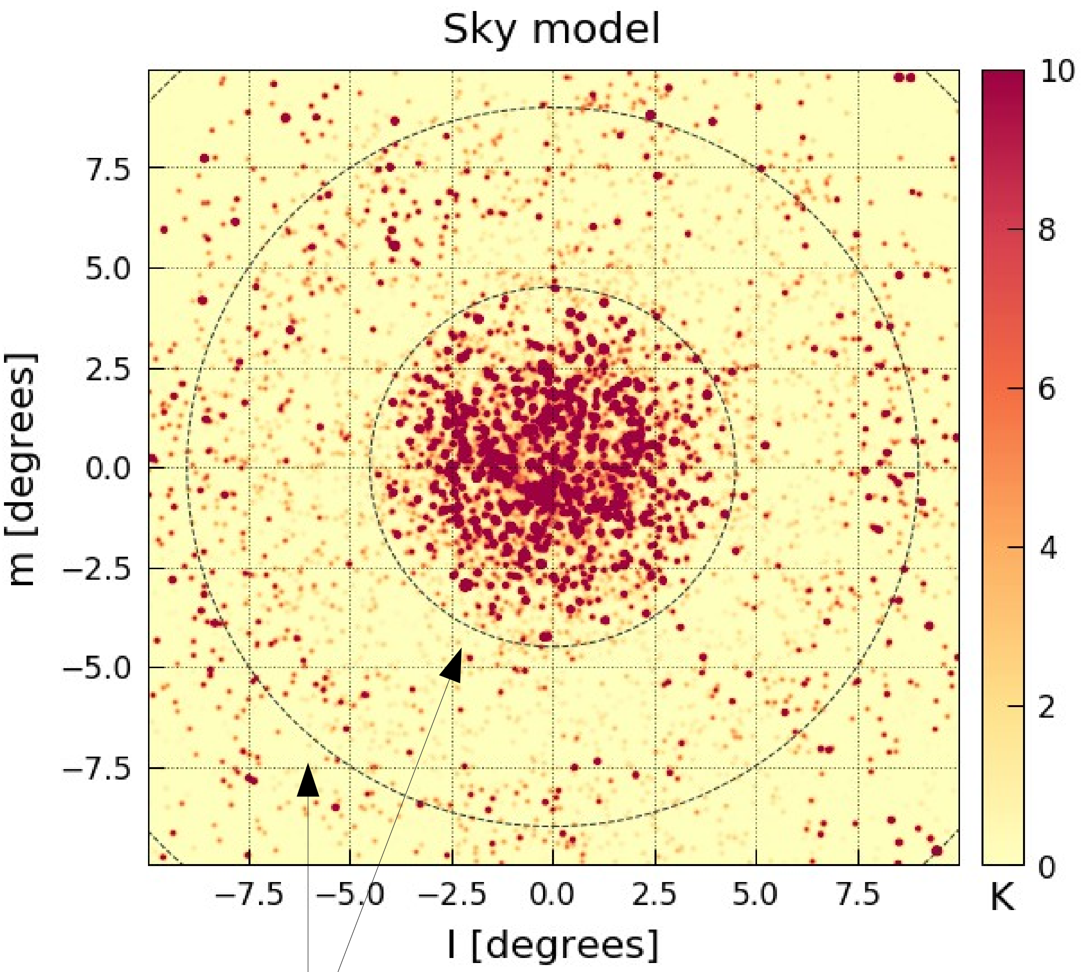
NCP field, 140 hours, 134-146 MHz, z ~ 9.1

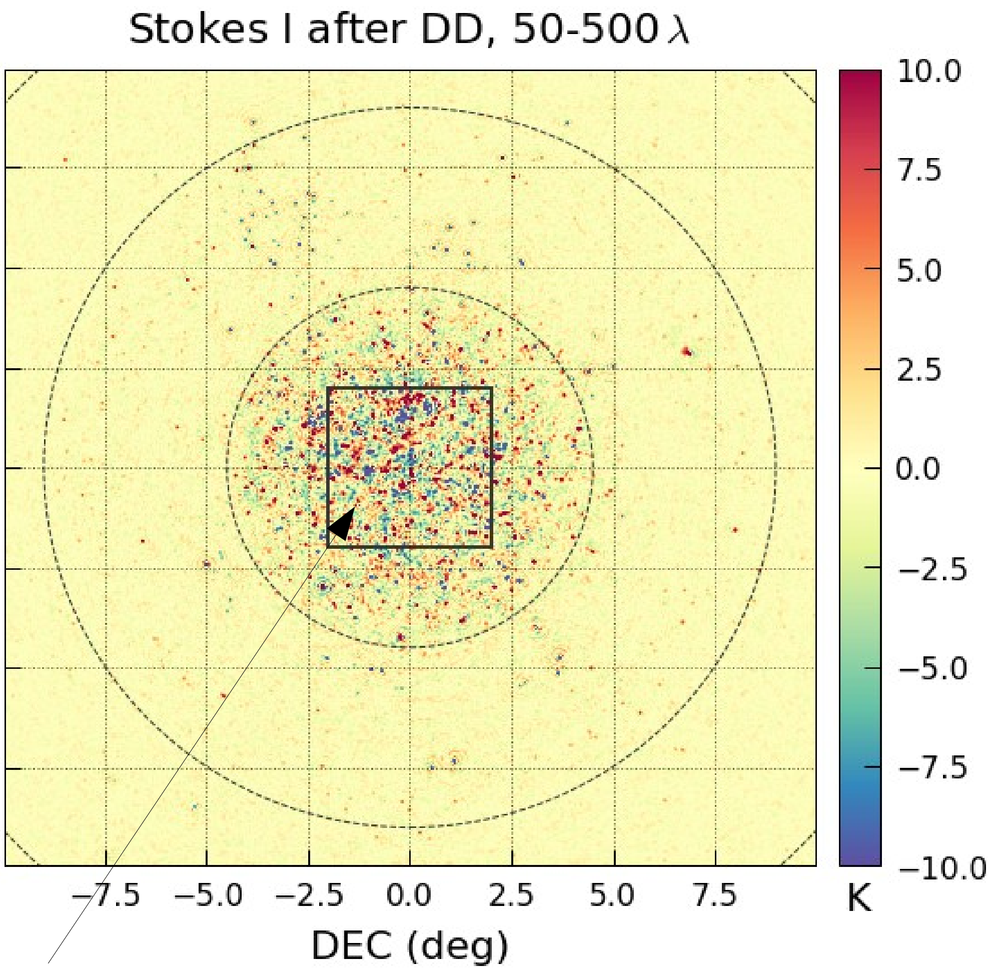


First and second null
of the Primary Beam

DD calibration results

NCP field, 140 hours, 134-146 MHz, z ~ 9.1





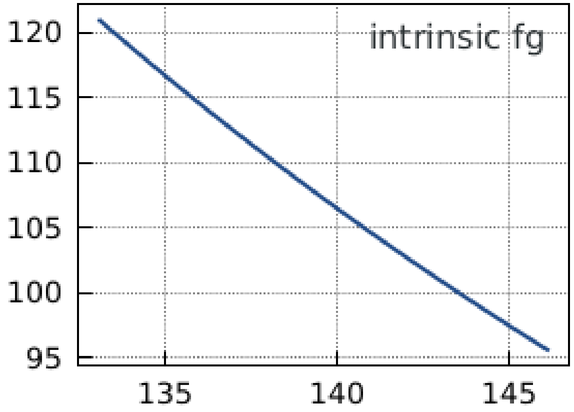
Inner 4x4 where we look for the signal

First and second null of the Primary Beam

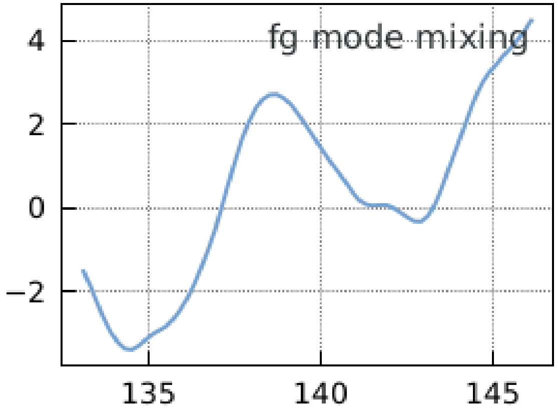
Next step: Remove confusion-limited foregrounds

GPR modeling for 21-cm experiments

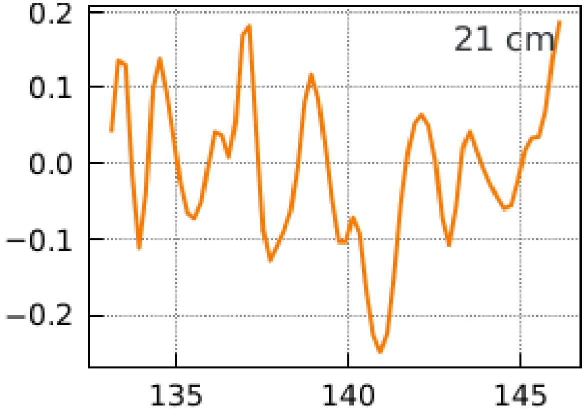
Residual data can be decomposed in three main components:



Residual astrophysical sources:
Smooth in frequency



Mode mixing:
Less frequency smooth



21-cm signal:
Uncorrelated ~ MHz

GPR: uses Gaussian Process (GP) as prior information $\mathbf{f} \sim \mathcal{N}(0, K)$

$$E(\mathbf{f}_{\mathrm{fg}}) = K_{\mathrm{fg}} \left[K_{\mathrm{fg}} + K_{21} + \sigma_n^2 I\right]^{-1} \mathbf{d}$$

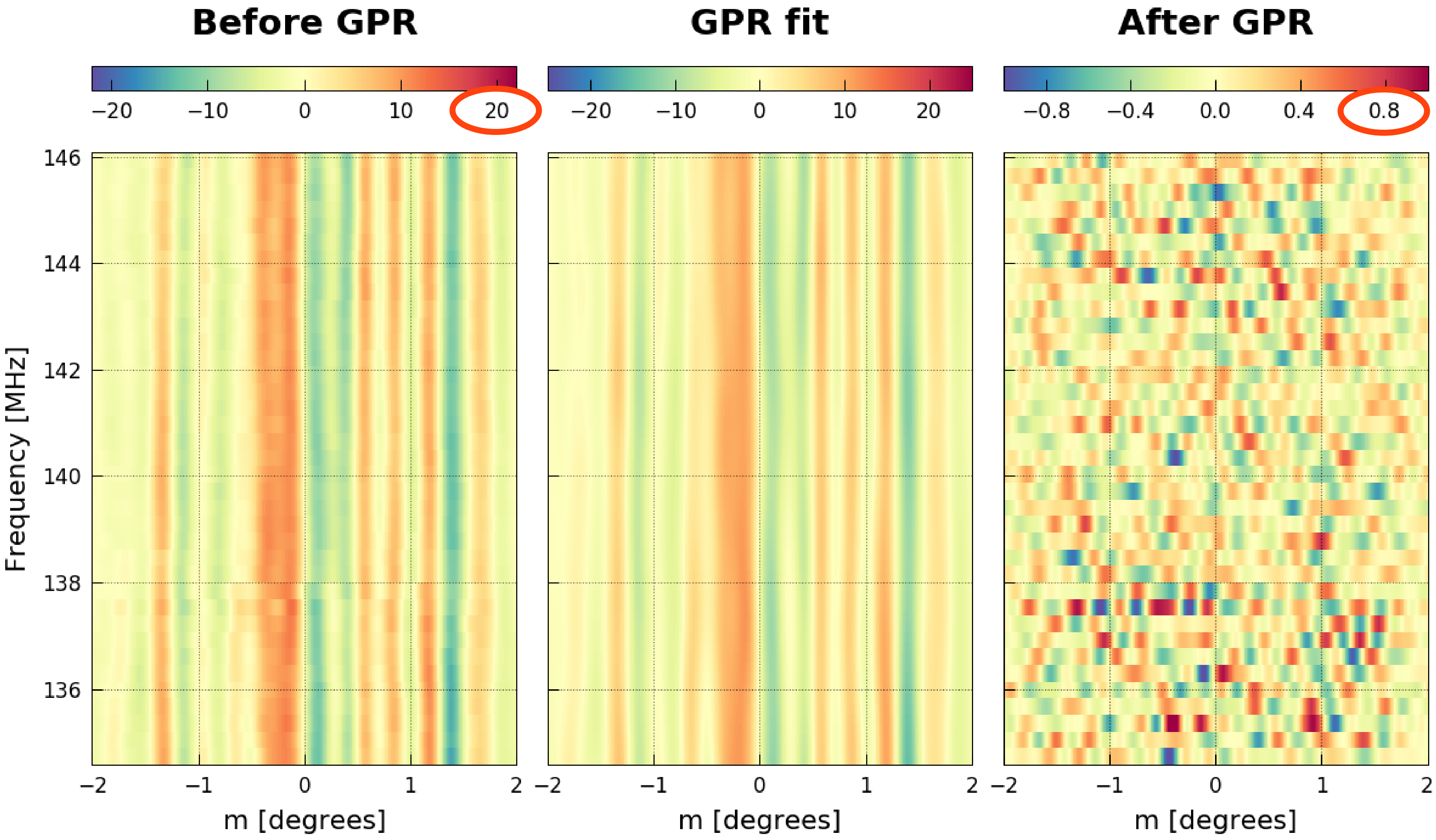
$$\mathrm{cov}(\mathbf{f}_{\mathrm{fg}}) = K_{\mathrm{fg}} - K_{\mathrm{fg}} \left[K_{\mathrm{fg}} + K_{21} + \sigma_n^2 I\right]^{-1} K_{\mathrm{fg}}$$

- ➔ Parametric Covariance optimized by maximizing the marginal likelihood (i.e. Bayesian evidence).
- ➔ Including prior information on the covariance contribution of the signal is key to avoid signal suppression!

(Mertens et al. 2018)

GPR on LOFAR data

NCP field, 140 hours, 134-146 MHz, z ~ 9.1

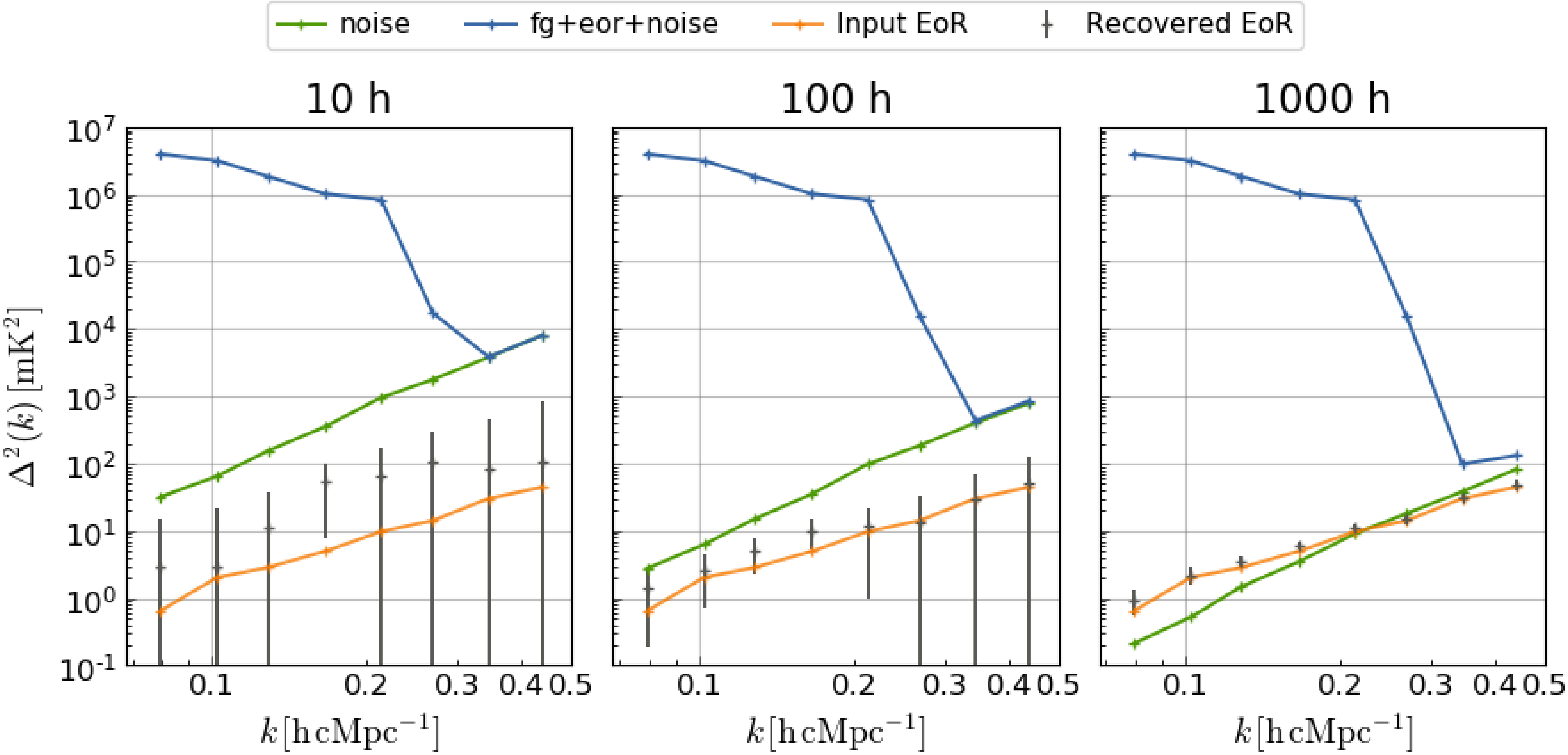


GPR remove frequency-coherent structure
Residual power level close to thermal noise

GPR on SKA simulation

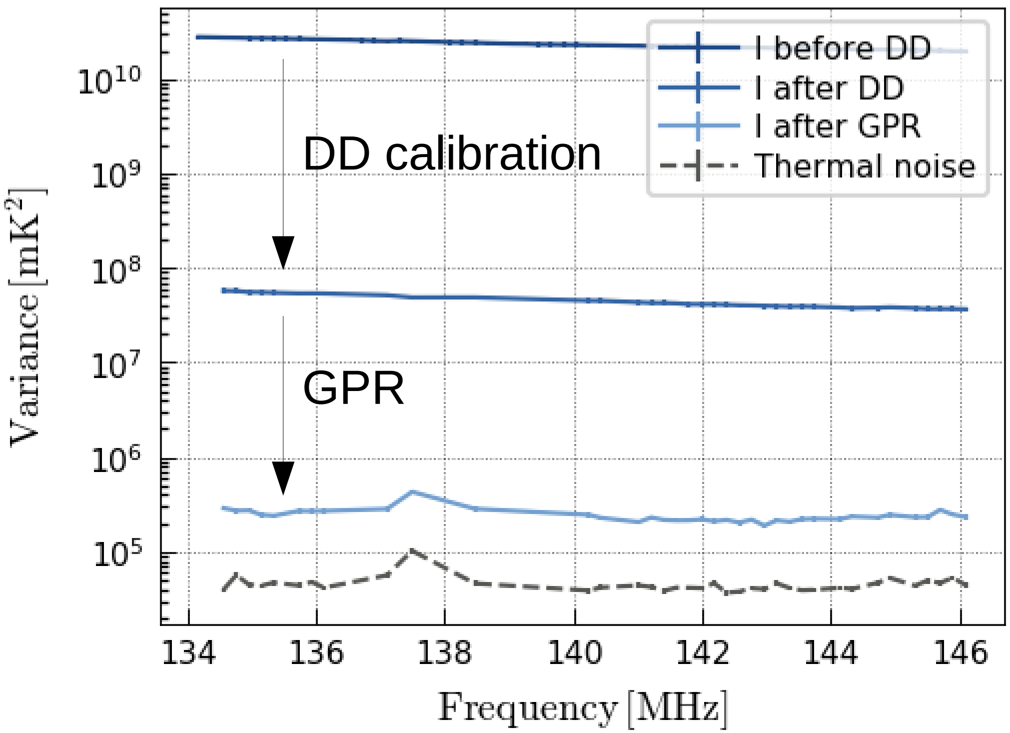
Simulation (from Modhurita Mitra for the SKA CD/EoR blind challenge):

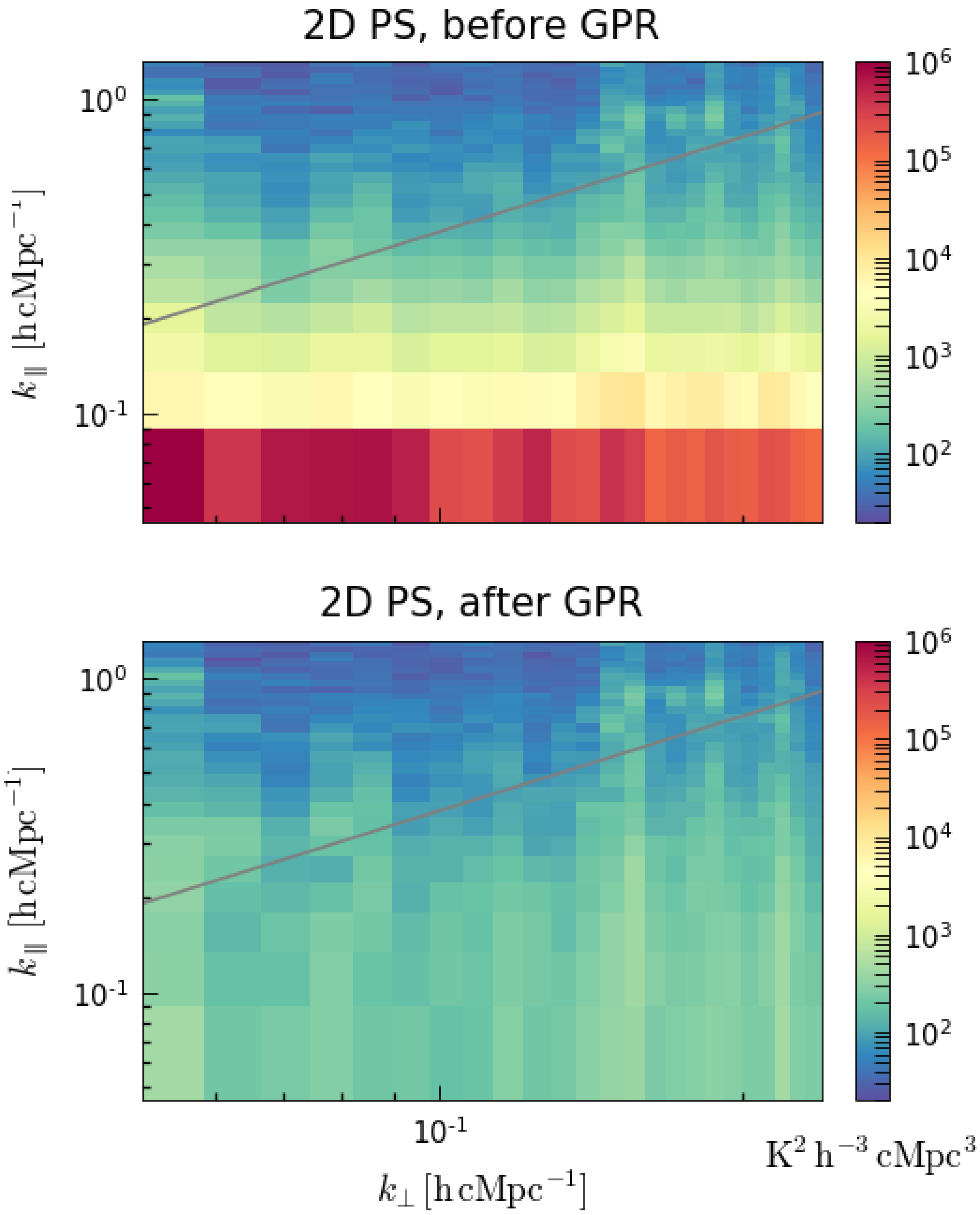
- Intrinsic foregrounds: galactic diffuse emission, 10 degree FoV
- 21-cm input signal: simulated from 21cmFast
- noise: equivalent to 10-100-1000 hours of SKA observation
- Visibility simulated using OSKAR



GPR on LOFAR data

NCP field, 140 hours, 134-146 MHz, z ~ 9.1

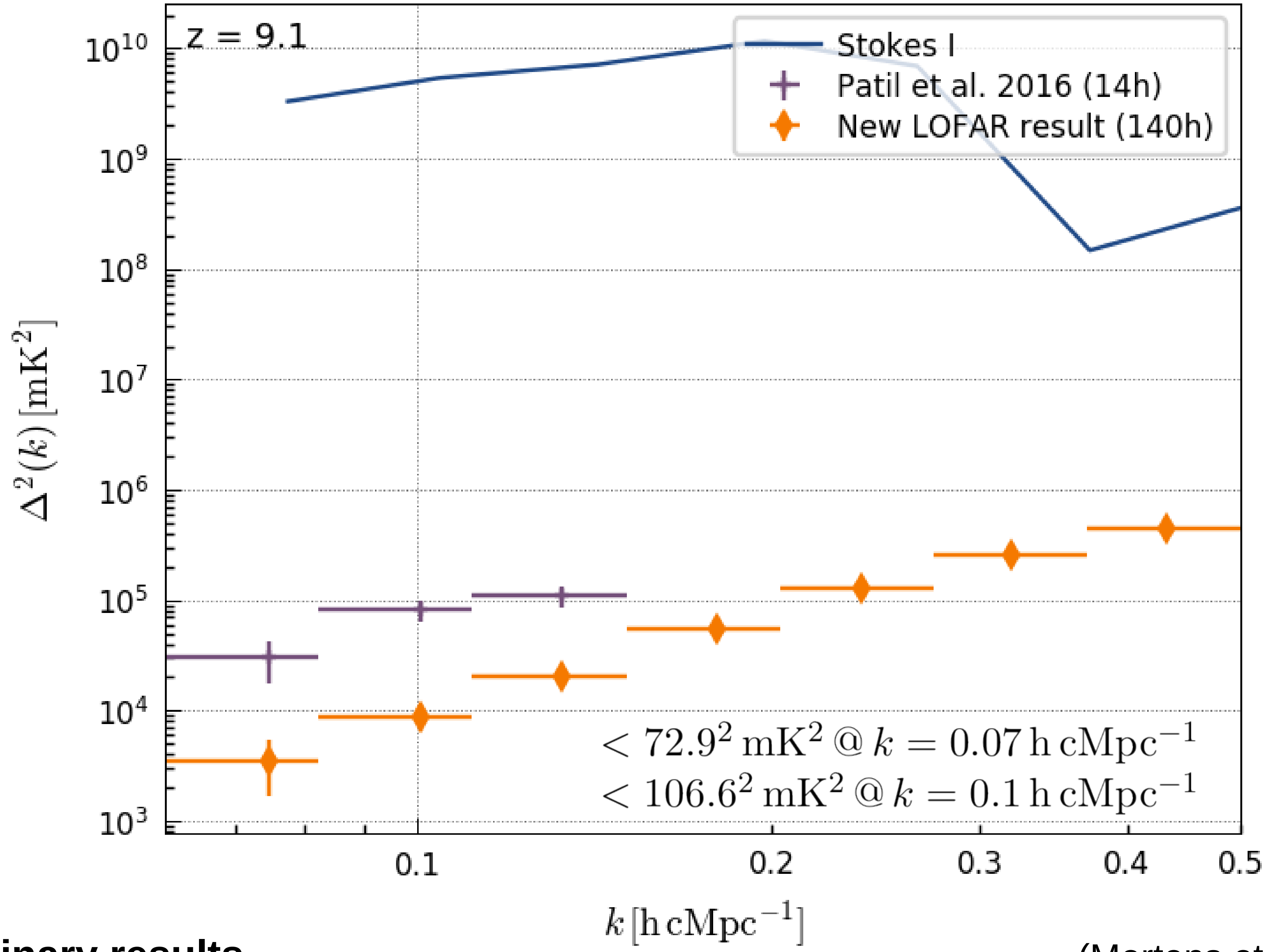




- DD calibration reduce foregrounds power by an order of magnitude down to confusion limit
- GPR remove residual foregrounds down to (very close to) noise level
- Residual power mostly incoherent between nights

New upper limit !

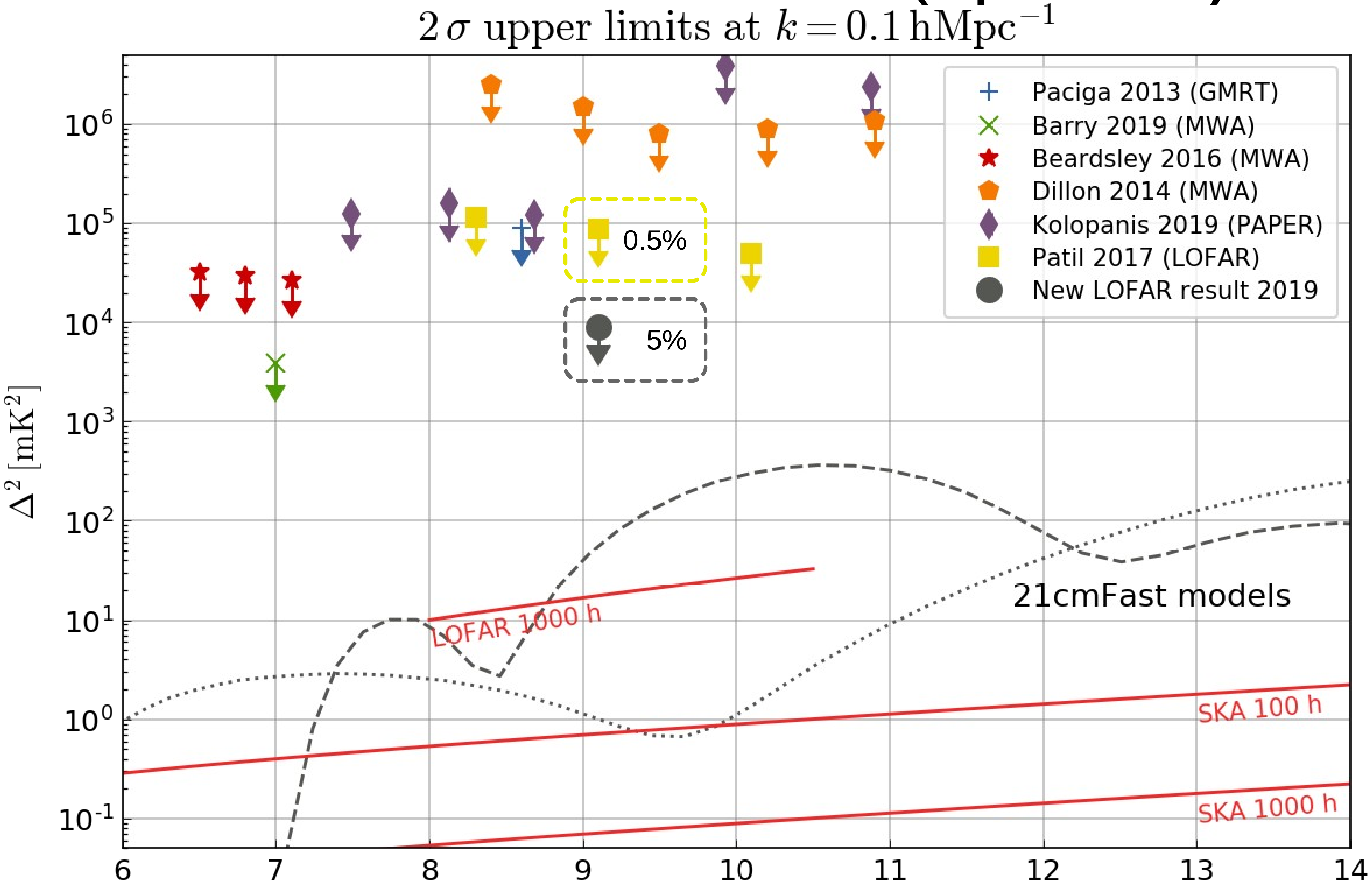
NCP field, 140 hours, 134-146 MHz, z ~ 9.1



Preliminary results

(Mertens et al. In prep.)

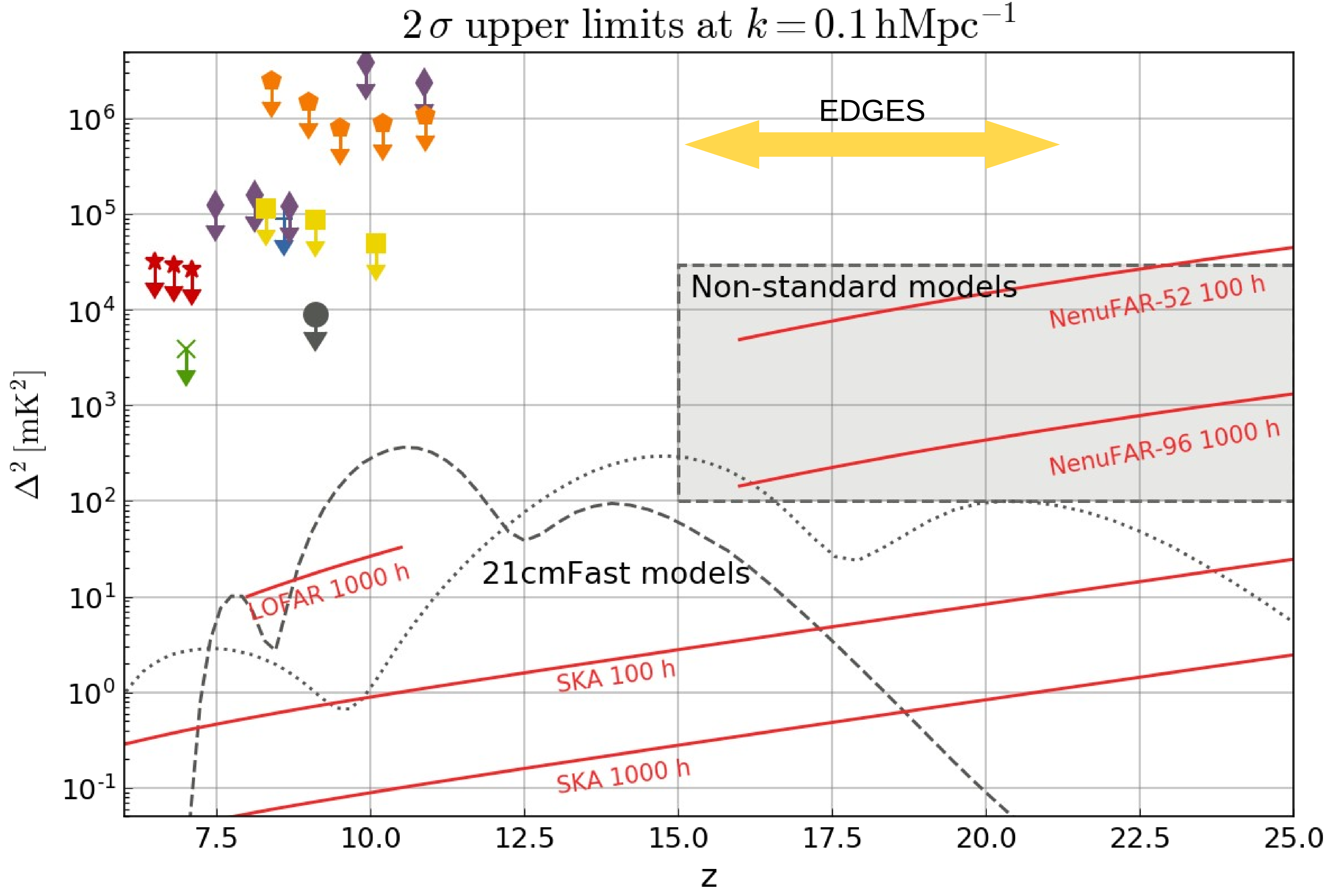
Where do we stand ? (updated)



Preliminary results

(Mertens et al. In prep.)

Perspective: NenuFAR, SKA

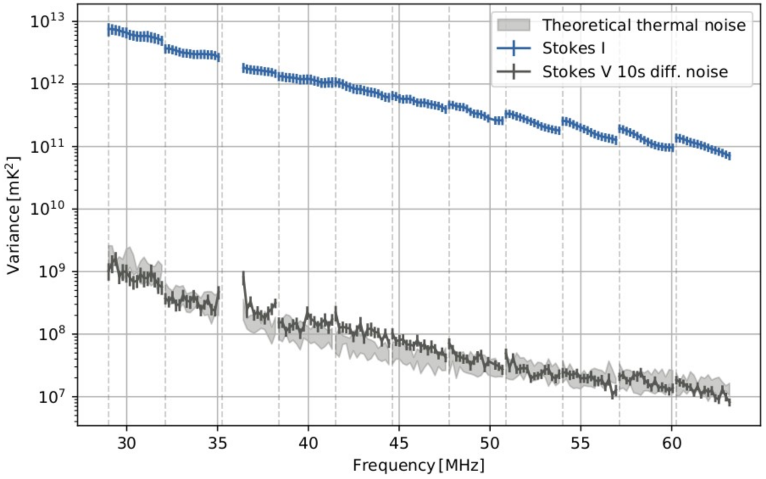


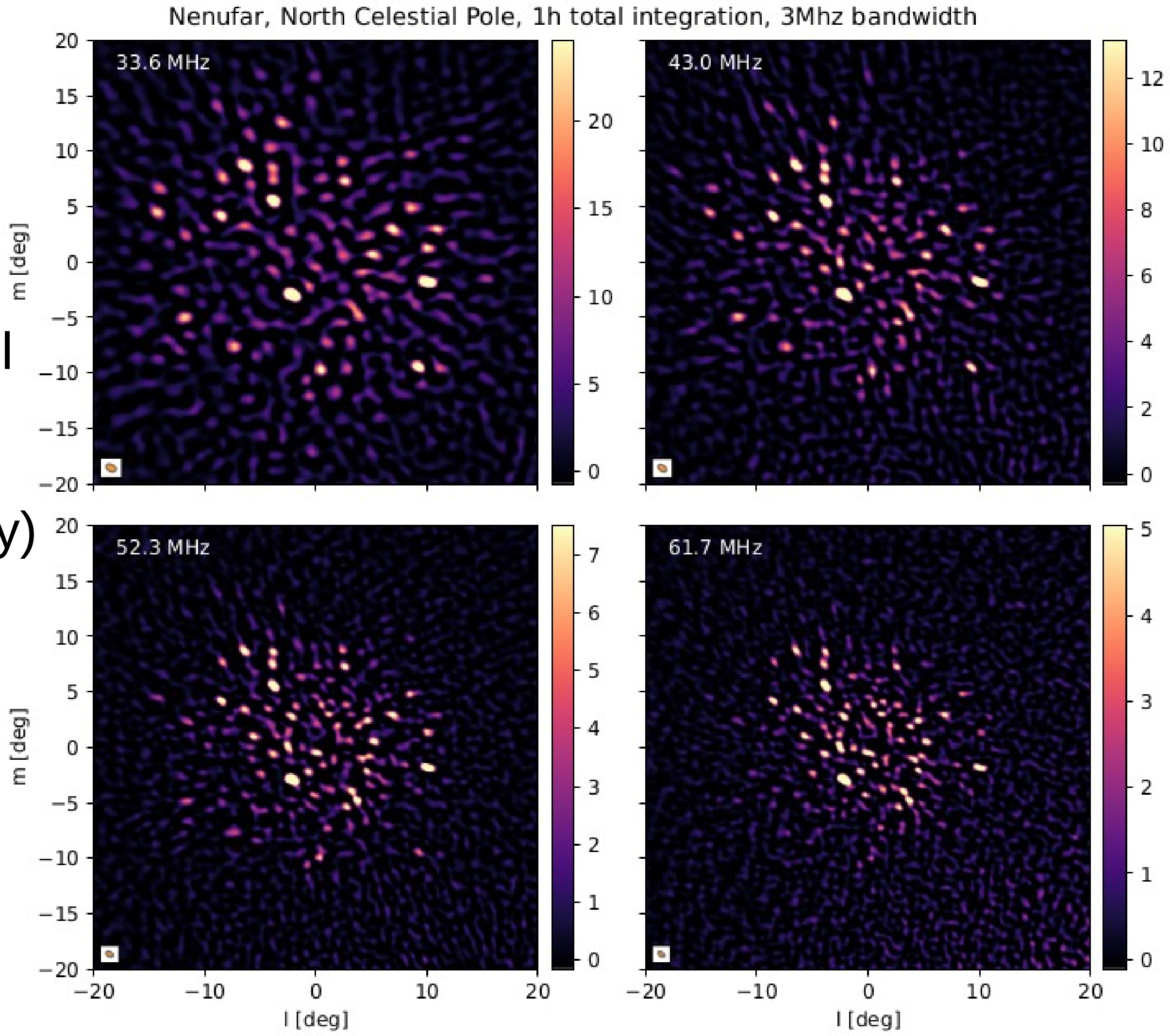
First NenuFAR results

Nenufar Cosmic Dawn KSP started observing 13/07/2019

- **First phase:**
  - Limited bandwidth and frequency/spatial resolution
  - ~ 170 / 330 hours observed
  - Goal 30-85 MHz NCP sky model

- **Second phase:**
  - Beginning 2020 (correlator ready)
  - + 1000 hours on NCP
  - 30 – 85 MHz (z ~ 46 – 15)





Summary

- The 21-cm signal from the Dark Ages, Cosmic Dawn and Reionization promises a new and unique probe of the first billion year of the Universe.

- Many ongoing/planned global and interferometric experiments, but very difficult experiments.

- Dealing with the foregrounds is one of the major challenges of CD/EoR experiments.

- **Current Status of the LOFAR-EoR project:**
  - ➔ Preliminary LOFAR deepest upper limits (based on ~5% of data): $\Delta^2 < (100\ \mathrm{mK})^2$ @ $k=0.1\ \mathrm{cMpc}^{-1}$, $z \sim 9$

- **Perspectives:**
  - ➔ Very interesting upper limit is still at reach with LOFAR.
  - ➔ Near future: NenuFAR exploring the Cosmic Dawn.
  - ➔ Far future: SKA promising tomography of the 21-cm signal.