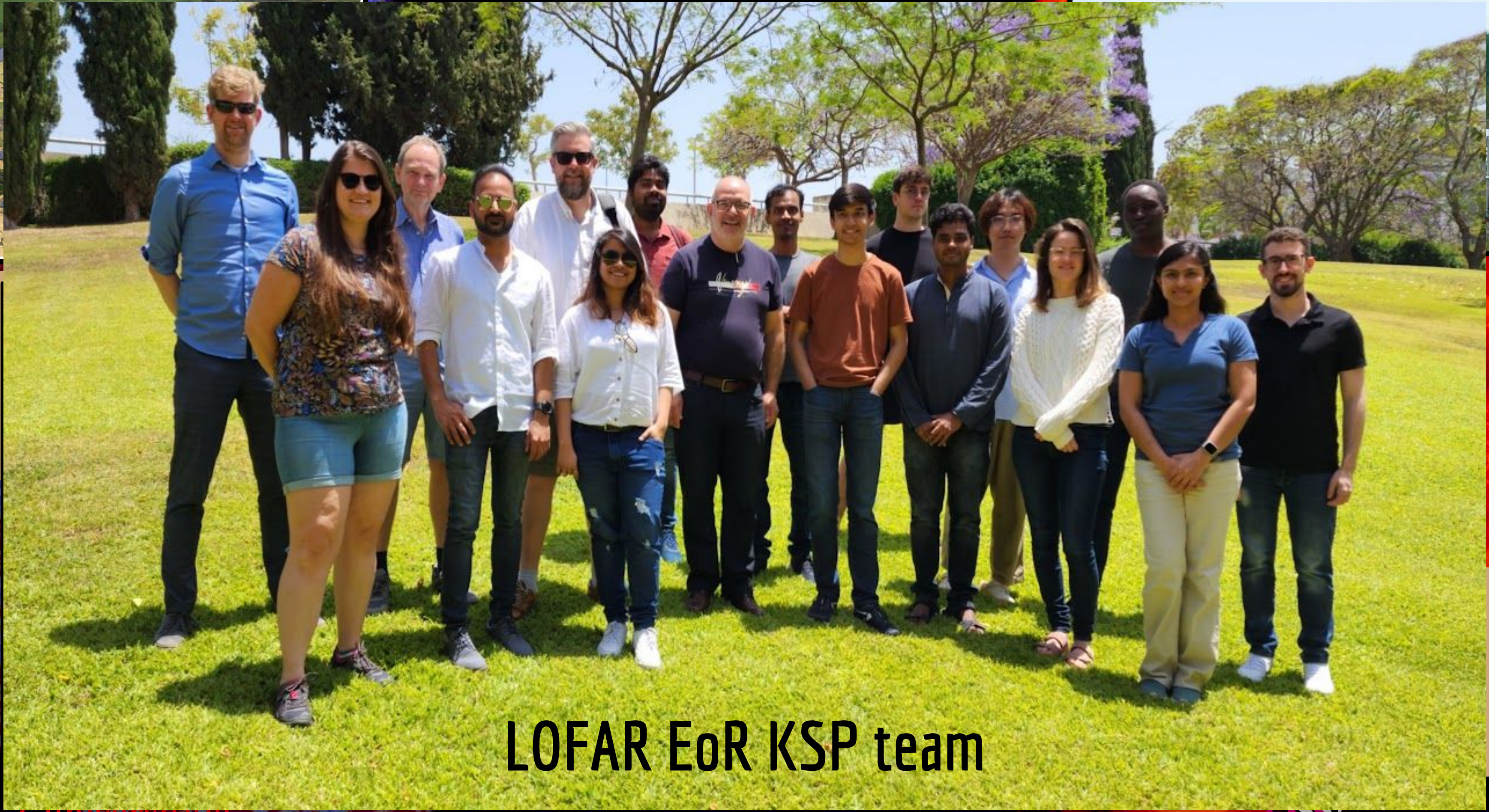
Kariuki Chege
and the LOFAR EoR KSP team

The 21-cm signal from the Epoch of Reionization and the Cosmic Dawn

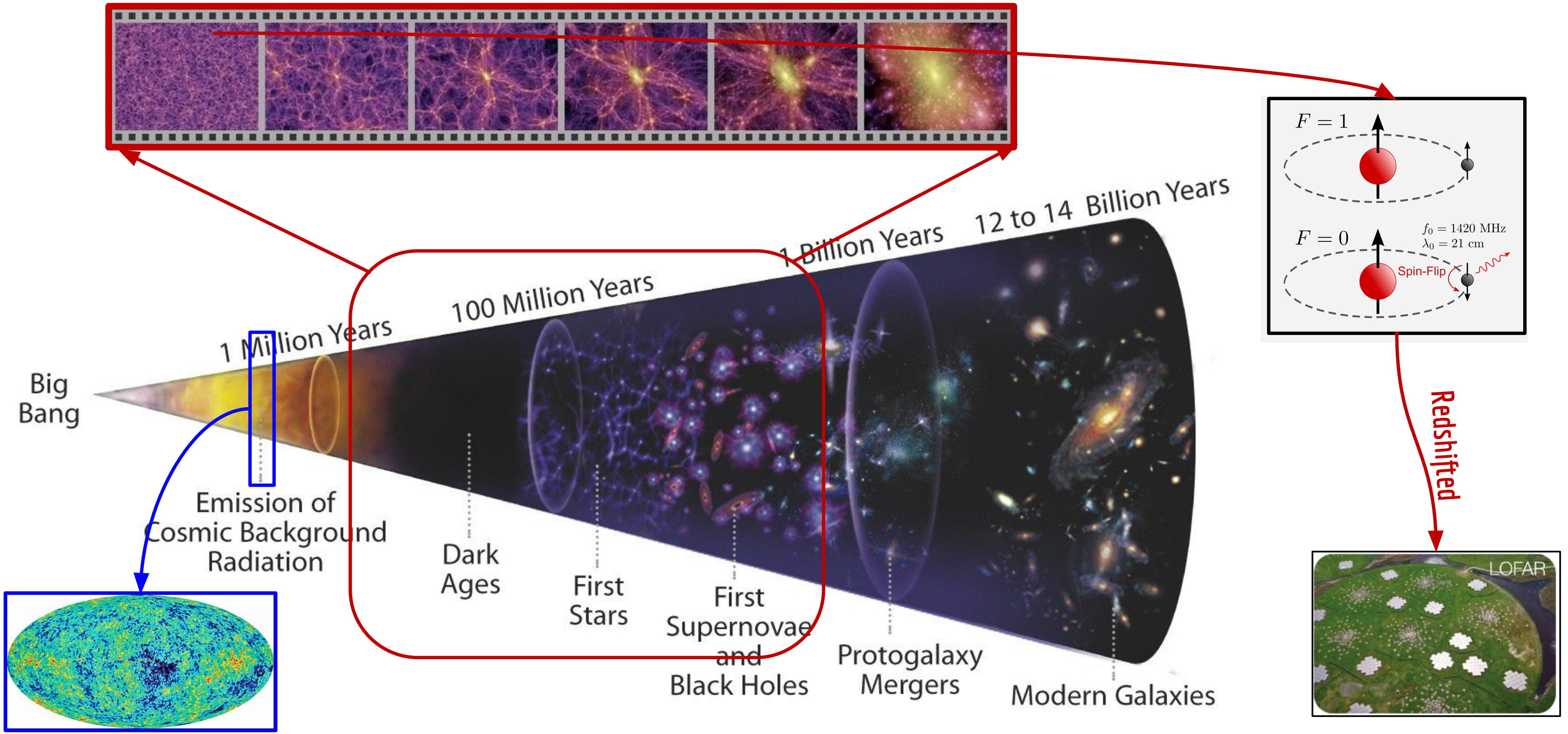
university of groningen

LOFAR



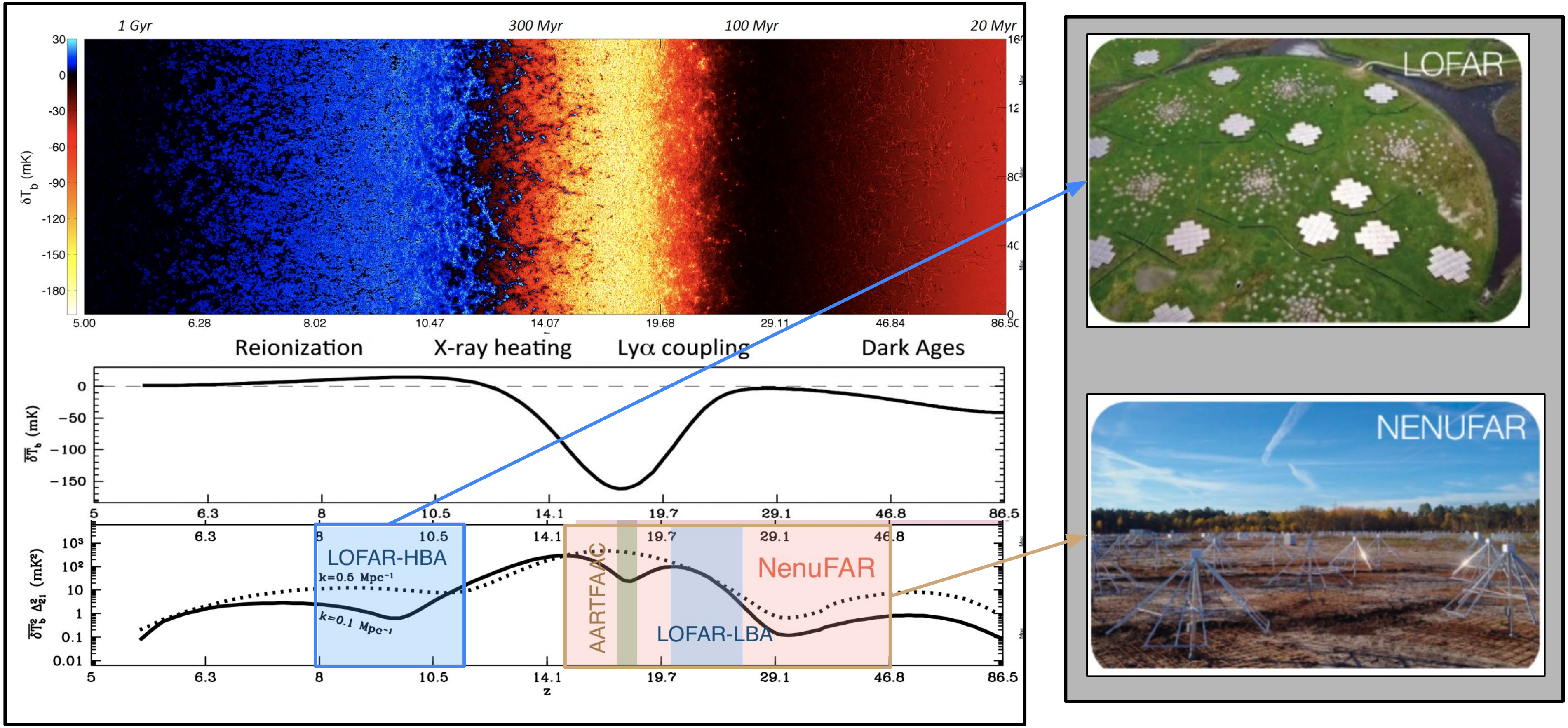
LOFAR EoR KSP team

Introduction



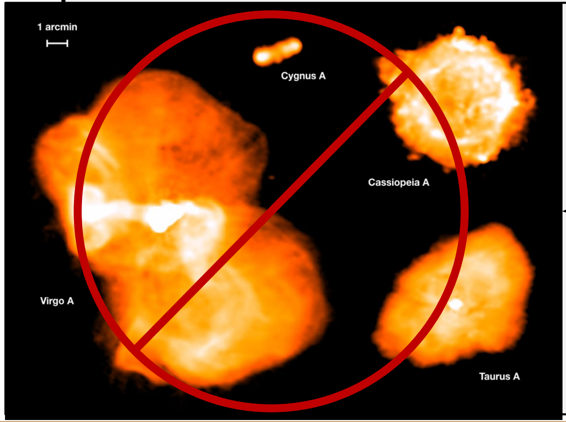
CoDEX –
ERC-Advanced program

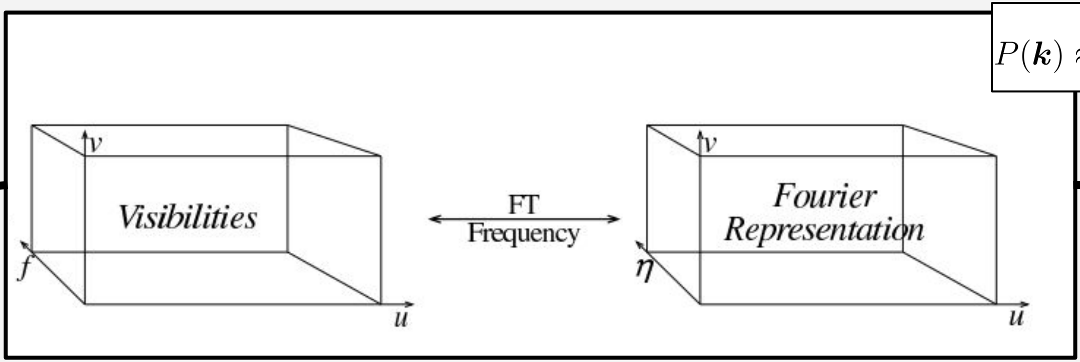
$$\delta T_b \approx 28 x_{\rm HI}(1+\delta)\left(1-\frac{T_\gamma}{T_s}\right)\left(\frac{\Omega_b h^2}{0.023}\right)\left(\frac{0.15}{\Omega_m h^2}\frac{1+z}{10}\right)^{1/2}\left[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\right] \rm mK$$



The power spectrum statistic

- We lack enough sensitivity to observe the spatial 21 cm fluctuations directly.
- Power spectrum is the 2-point correlation function - measure of excess signal (above random) on all spatial scales.
- Computed by the angle averaged sum of the Fourier transform of the normalized 21 cm brightness temperature.
- Produces a distribution of modes characterizing the amount of structural information in the signal.





$$P(\boldsymbol{k}) \approx \frac{\langle|\tilde{T}(\boldsymbol{k})|^2\rangle}{V}$$

Power Spectra

21-cm signal:

- Uncorrelated - MHz
- Largely isotropic

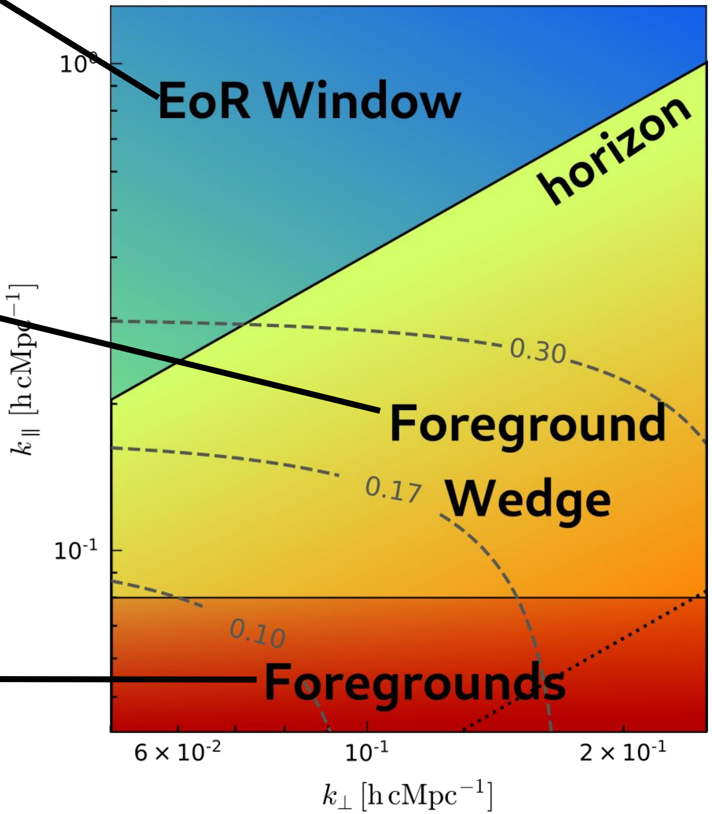
Foregrounds Wedge:

- Due to a chromatic instrument
- Dependent on :
  - Coordinates of the Instrument and phase center
  - Time of the observation
  - Beam/UV coverage
  - Off-axis emission upto the horizon

Foregrounds:

- Much brighter than the signal
- Mainly synchrotron and free-free emission
- Smooth in frequency

Spatial vs line-of-sight power-spectra

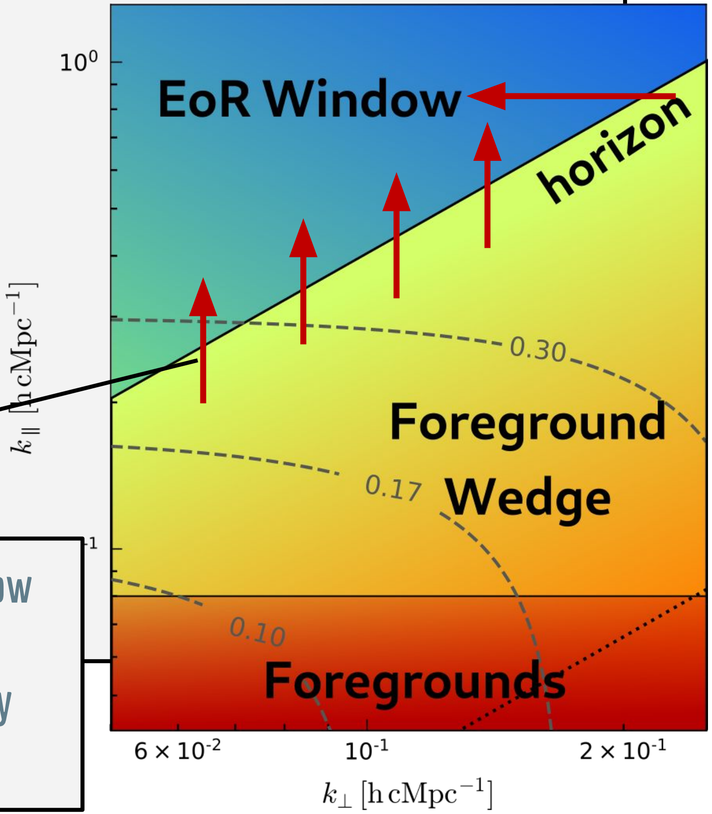


Main Challenge: Foregrounds

Foregrounds are much brighter than the signal

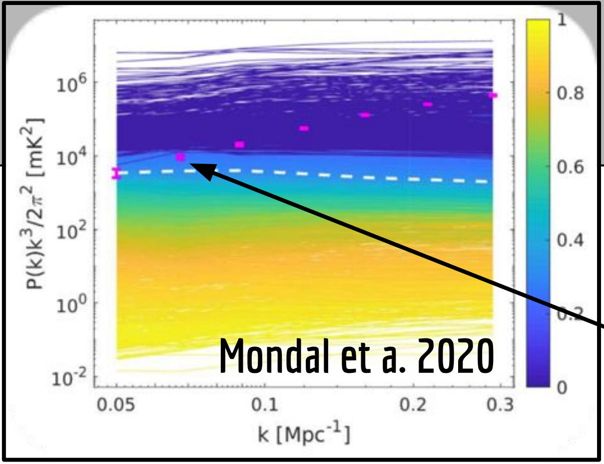
Several factors and systematics exacerbate the challenge

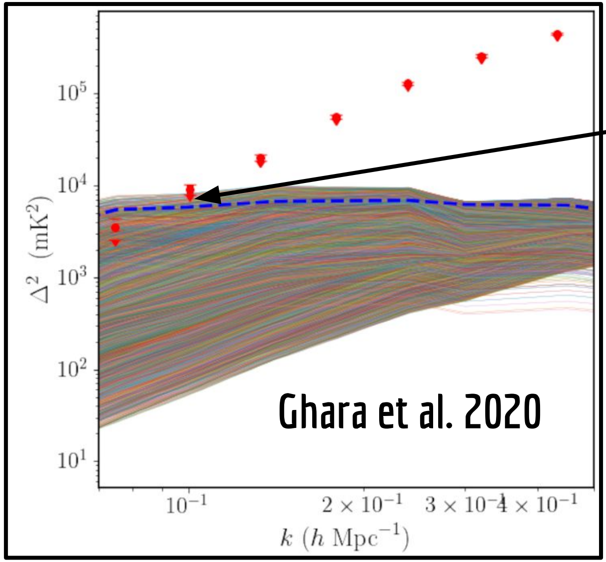
- Chromatic instrument (beam/uv-coverage)
- Incomplete sky models
- Calibration errors
- Ionosphere
- Local and transient RFI
- Polarization leakage

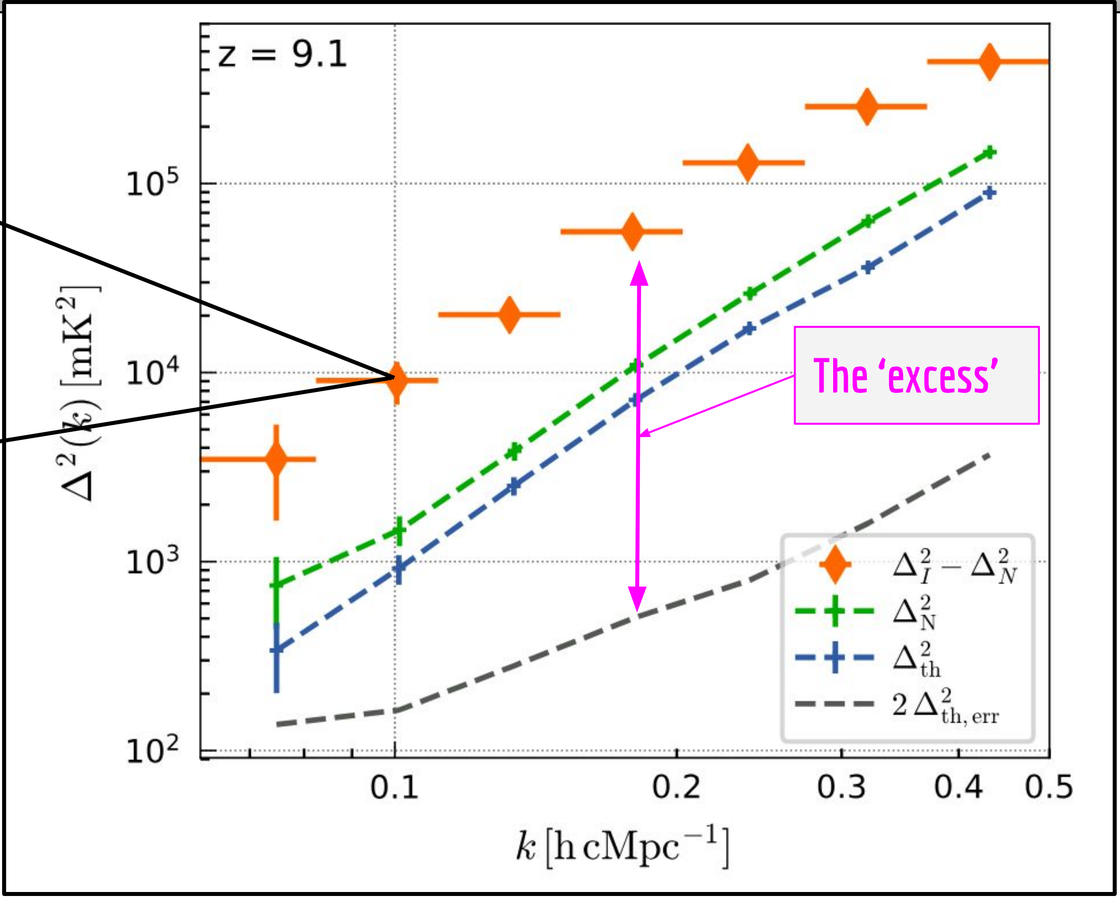


- Systematics cause foregrounds to leak further up the EoR window or across angular scales
- Results in 'excess variance' which currently hinders detection by any instrument, including HERA, MWA...

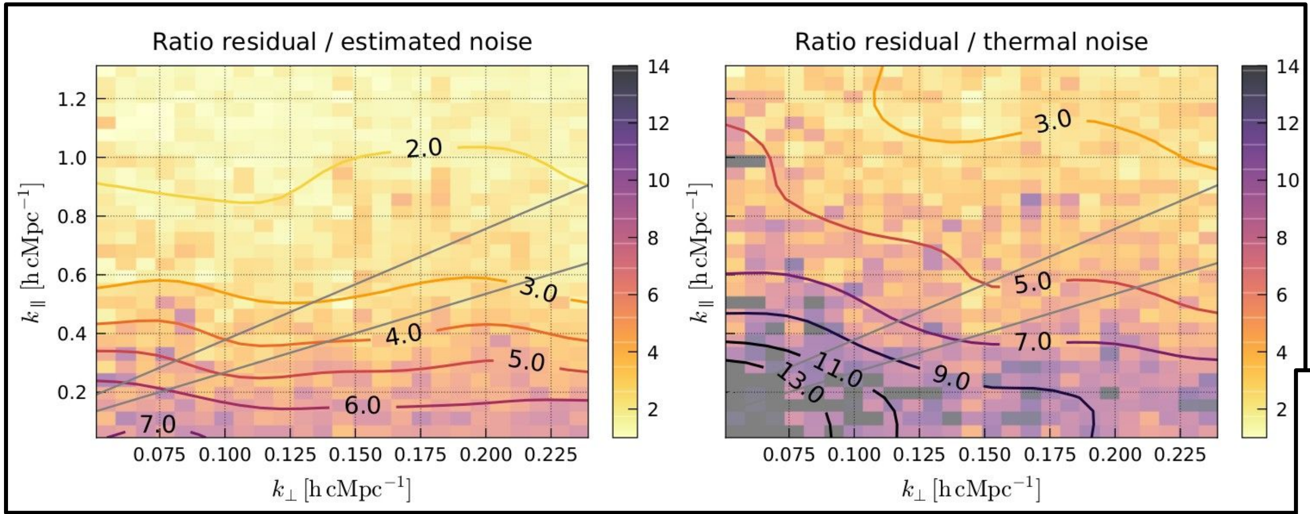
Limits on the 1D power spectra



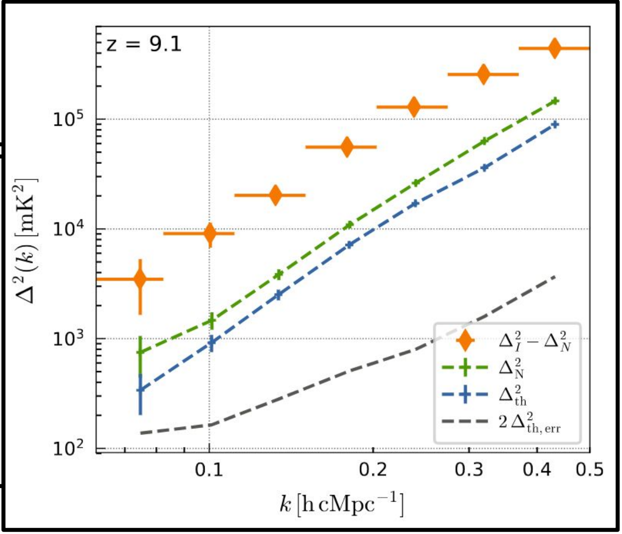




NCP field upper limits (2020)



[Mertens et al. 2020]



- ➔ Our 2020 upper limit is an improvement by a factor ~10 over previous upper limit.
- ➔ However, residual power still much higher than thermal noise.
- ➔ Residual excess power spreads over a region far larger than the foregrounds wedge

NCP Pipeline

3 redshift bins

DI calibration:

Regularization strategies,
New 3C61.1 model

DD calibration:

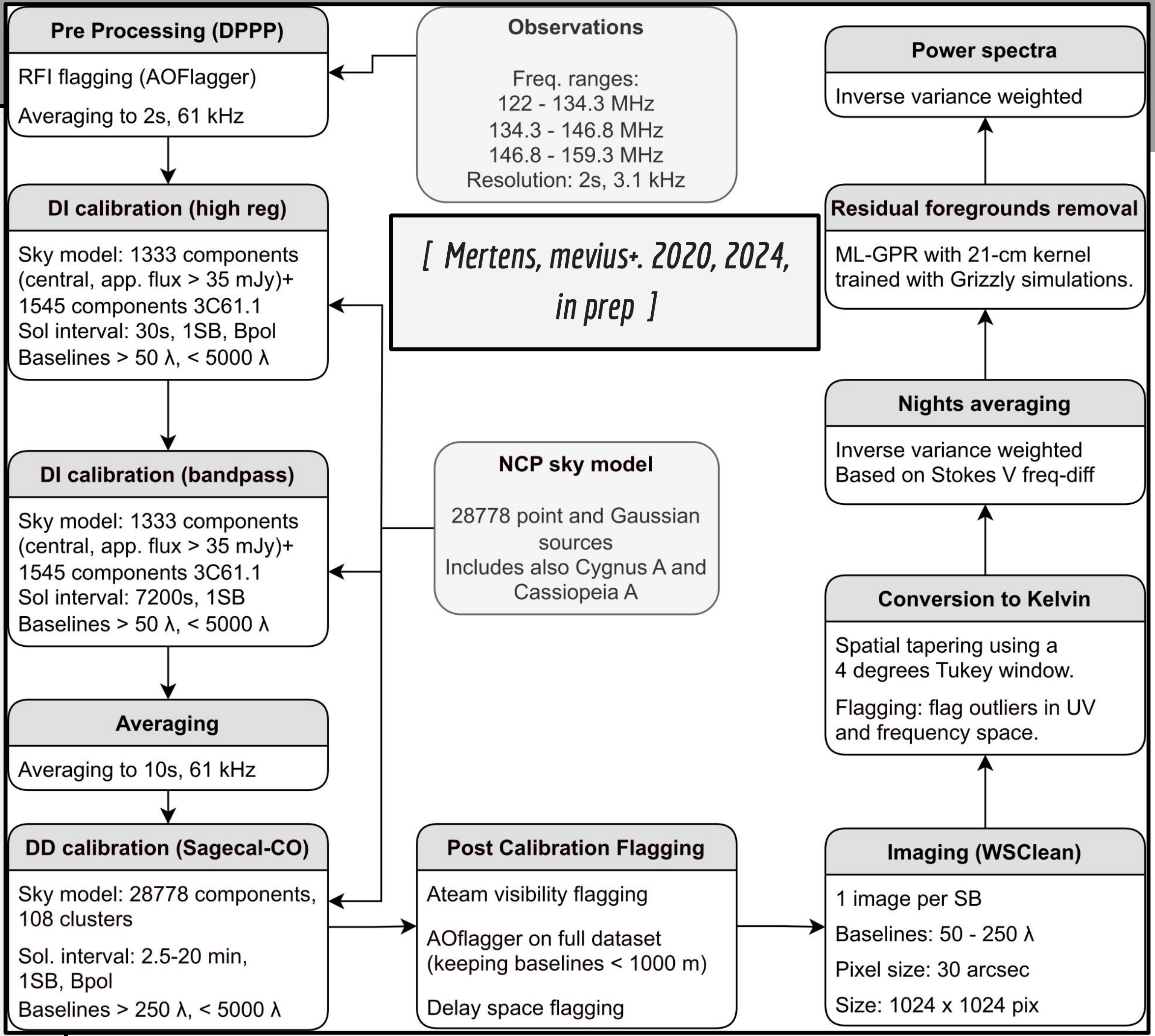
Apparent model, less directions
Simpler A-team model

Post processing flagging:

RFI excision in:

1. uv space
2. delay space

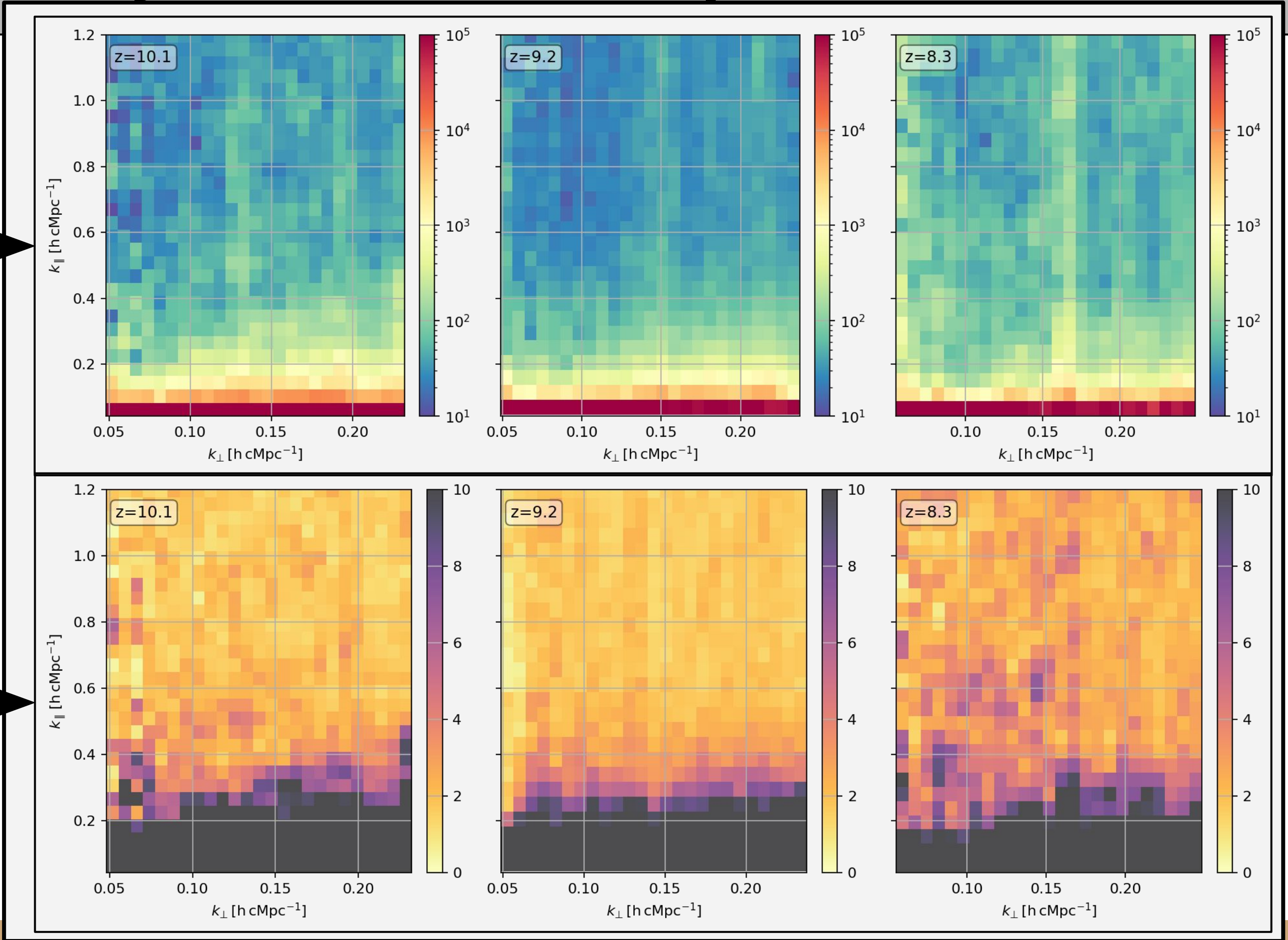
GPR → ML-GPR



New results (2024) - Residuals before GPR

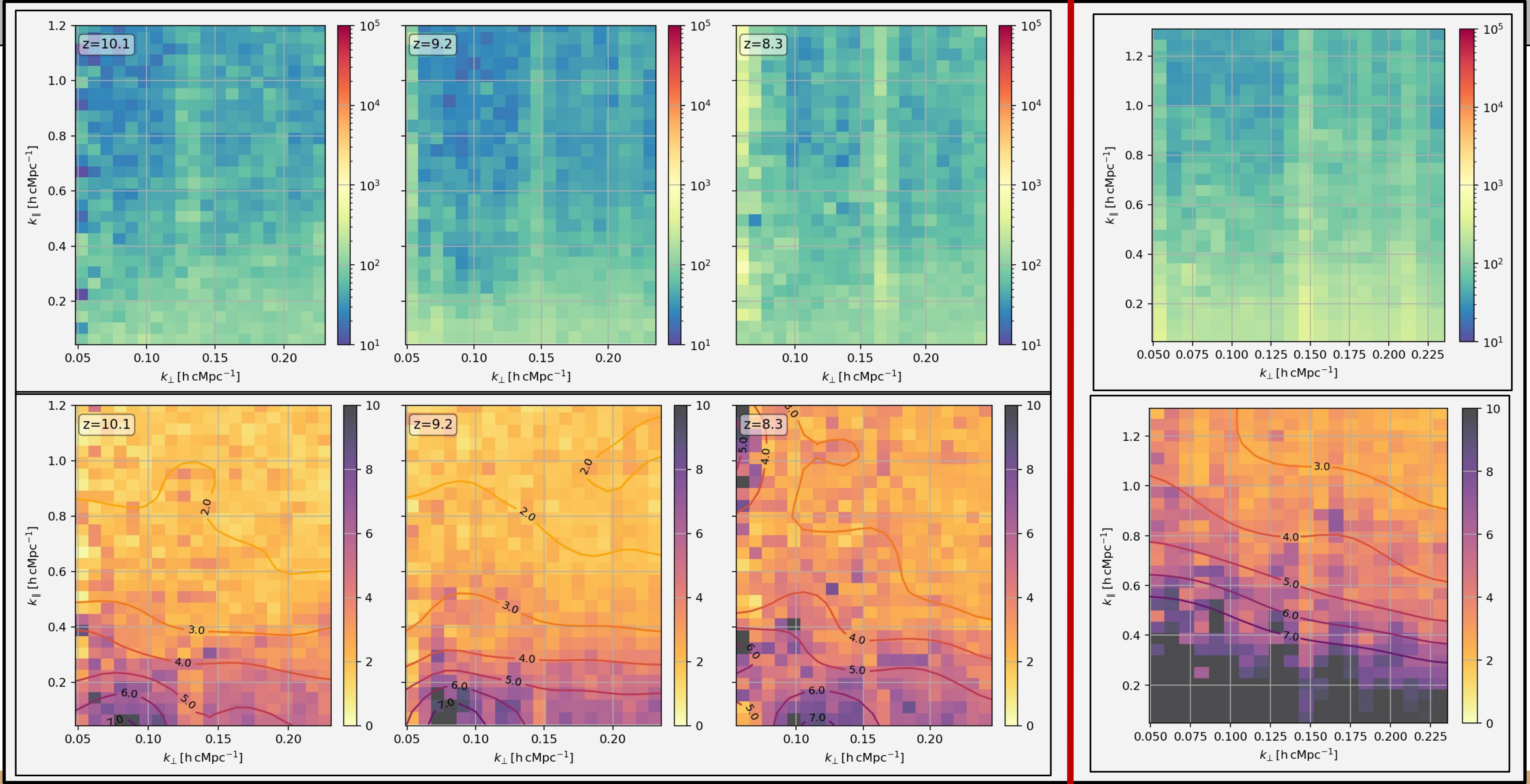
Stokes I Power Spectrum

Ratio with thermal noise spectrum

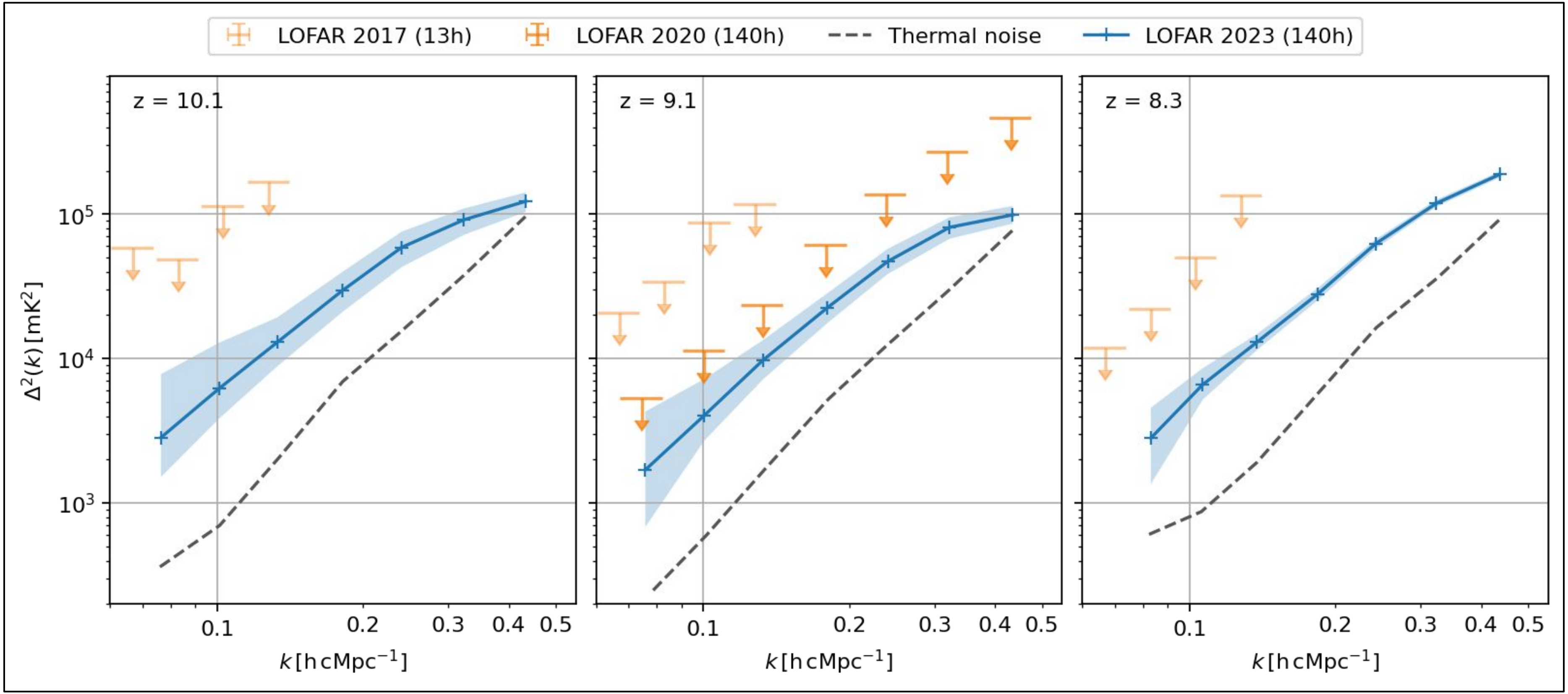


Residuals after GPR vs Thermal noise

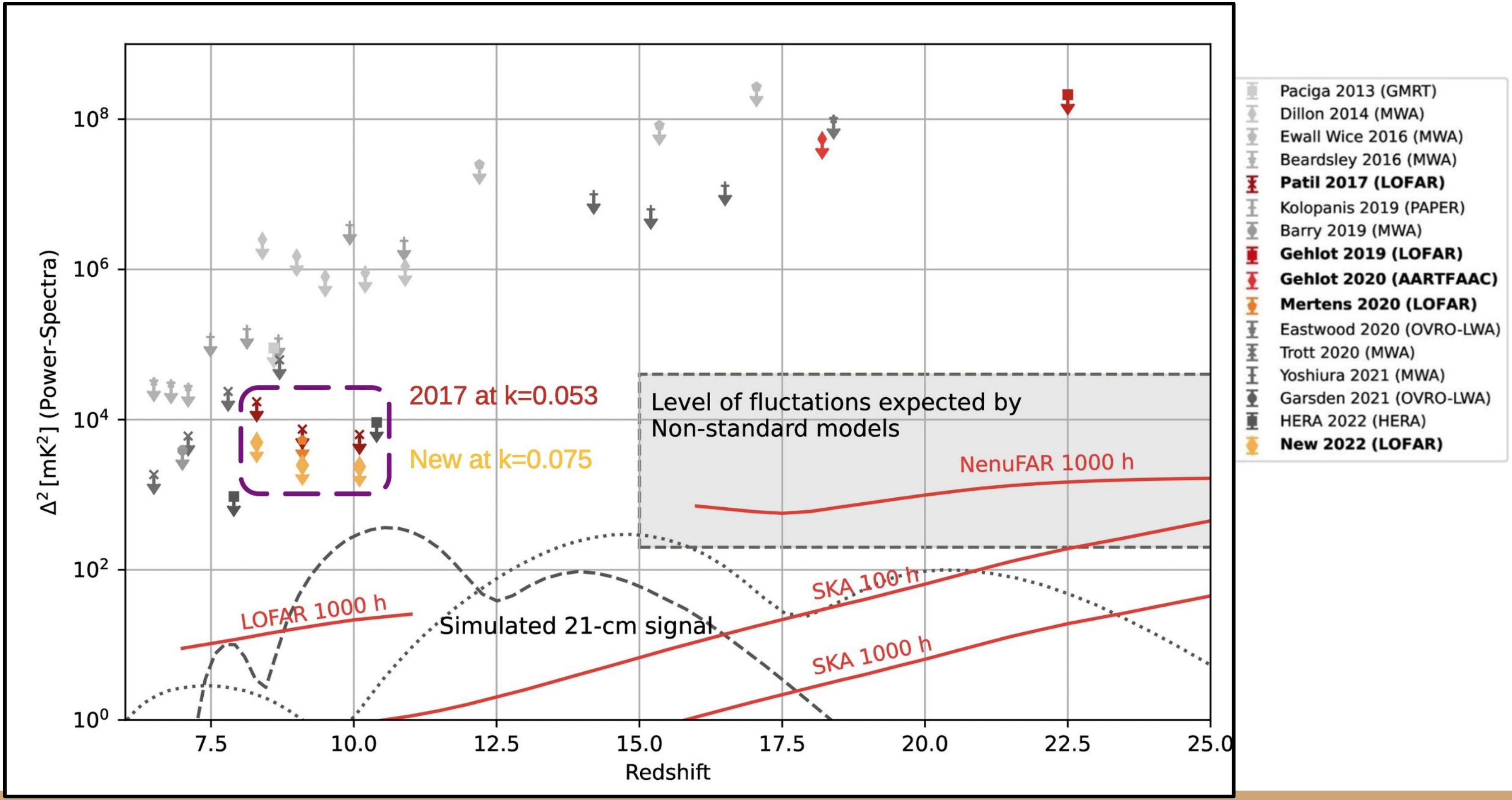
2024 (z=8.3, 9.2, 10.1) | 2020 (z=9.1)



Latest NCP Field Upper Limits!



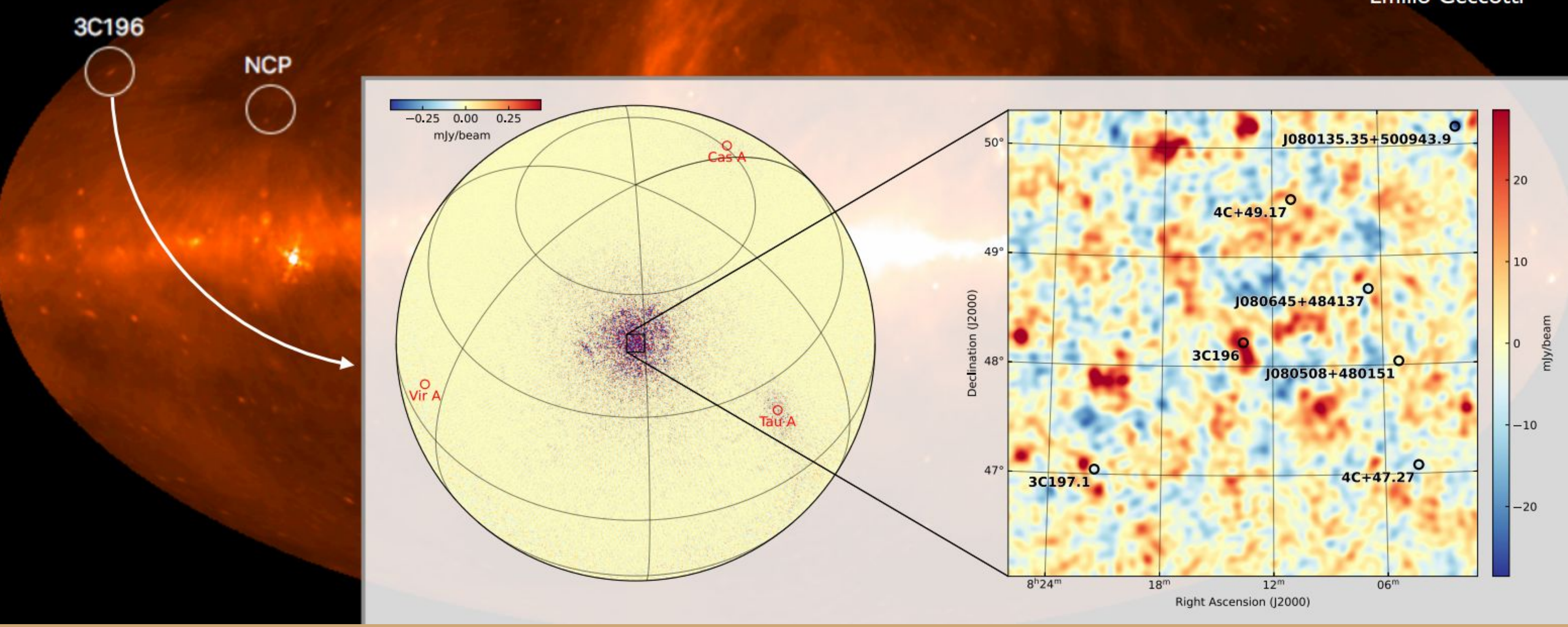
Where do we stand now?



NEW RESULTS FROM 3C196 FIELD AT z~9.2

- Sensitivity higher than NCP (closer to zenith)
- Strong source in the centre that makes DI calibration easier
- Cas A & Cyg A more than 66° far from the centre

Emilio Ceccotti



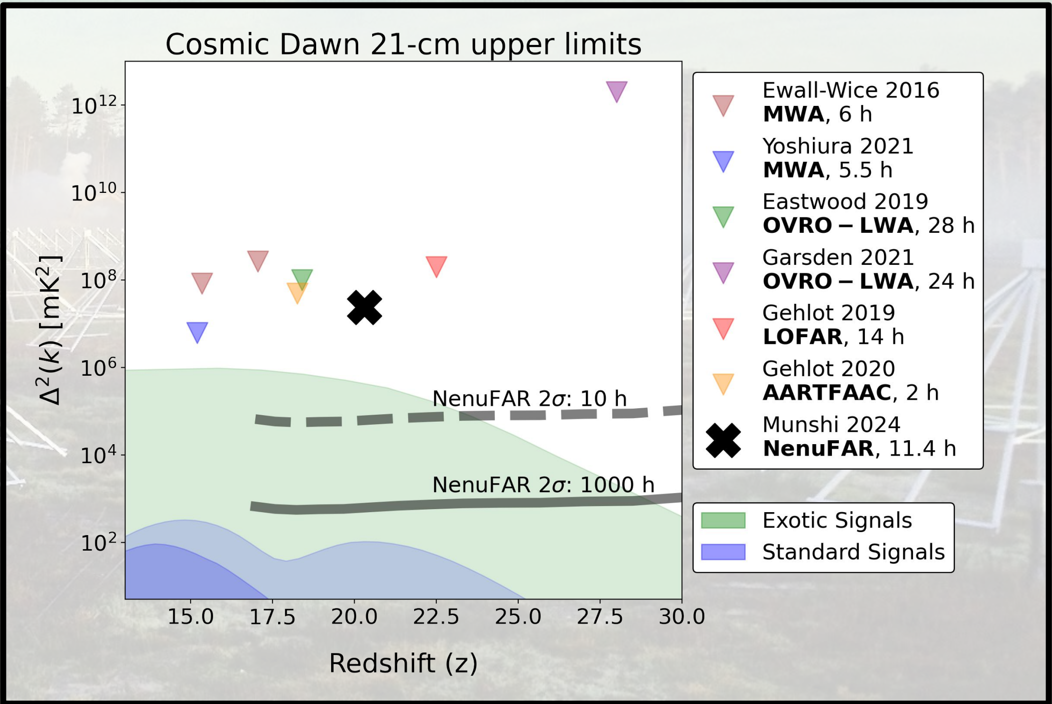
First ever NenuFAR upper limits !

S. Munshi

Status:

- **75** core and **4** remote stations active, out of 96 core and 8 remote stations
- **1250+ hours** of observation of the NCP
- **First upper limits** from one night
- **Strong excess** due to
  - Local RFI sources
  - Off-axis sky sources

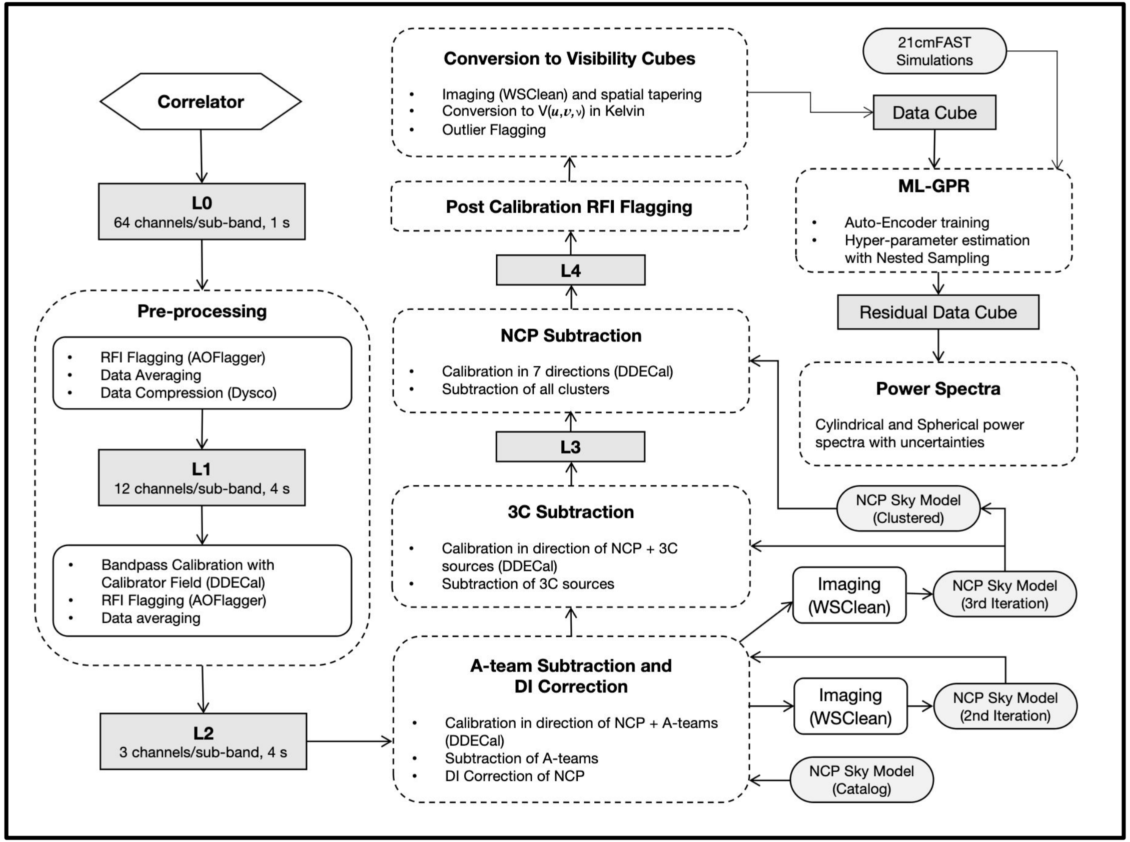
Munshi+ 2024

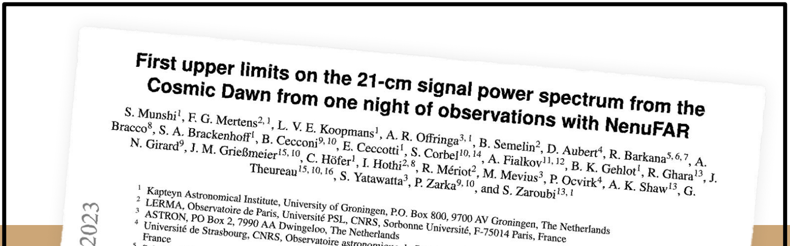


Methods based on NCP tools

Munshi+ 2024

- Using tools and methods developed for the LOFAR NCP field
- Shows our methods are robust and easily transferable to the future including SKA





First upper limits on the 21-cm signal power spectrum from the Cosmic Dawn from one night of observations with NenuFAR

S. Munshi[1], F. G. Mertens[2,1], L. V. E. Koopmans[1], A. R. Offringa[3,1], B. Semelin[2], D. Aubert[4], R. Barkana[5,6,7], A. Bracco[8], S. A. Brackenhoff[1], B. Cecconi[9,10], E. Ceccotti[1], S. Corbel[10,14], A. Fialkov[11,12], B. K. Gehlot[1], R. Ghara[13], J. N. Girard[9], J. M. Grießmeier[15,10], C. Höfer[1], I. Hothi[2,8], R. Mériot[2], M. Mevius[3], P. Ocvirk[4], A. K. Shaw[13], G. Theureau[15,10,16], S. Yatawatta[3], P. Zarka[9,10], and S. Zaroubi[13,1]

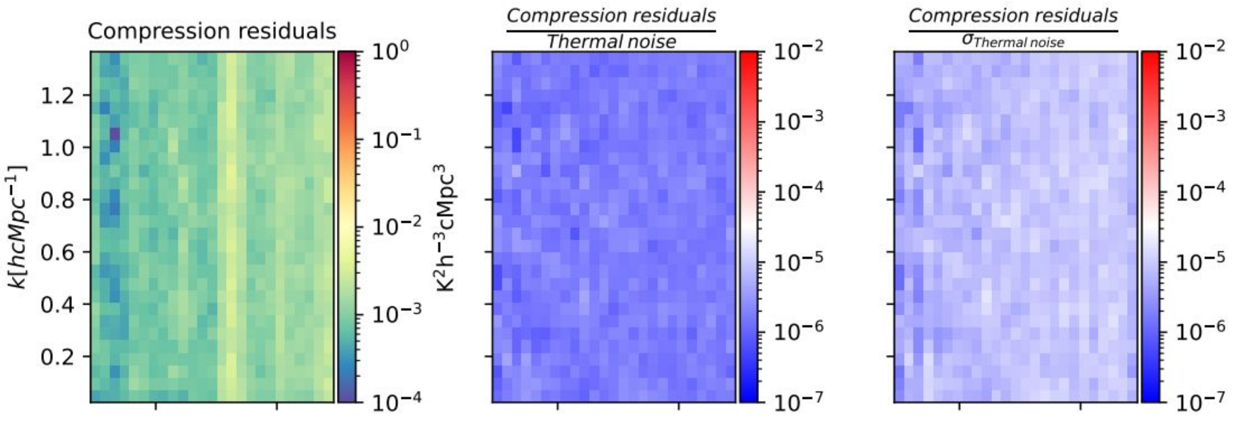
[1] Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands
[2] LERMA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, F-75014 Paris, France
[3] ASTRON, PO Box 2, 7990 AA Dwingeloo, The Netherlands
[4] Université de Strasbourg, CNRS, Observatoire astronomique ...
France

2023

Impact of Dysco data compression in the power spectrum

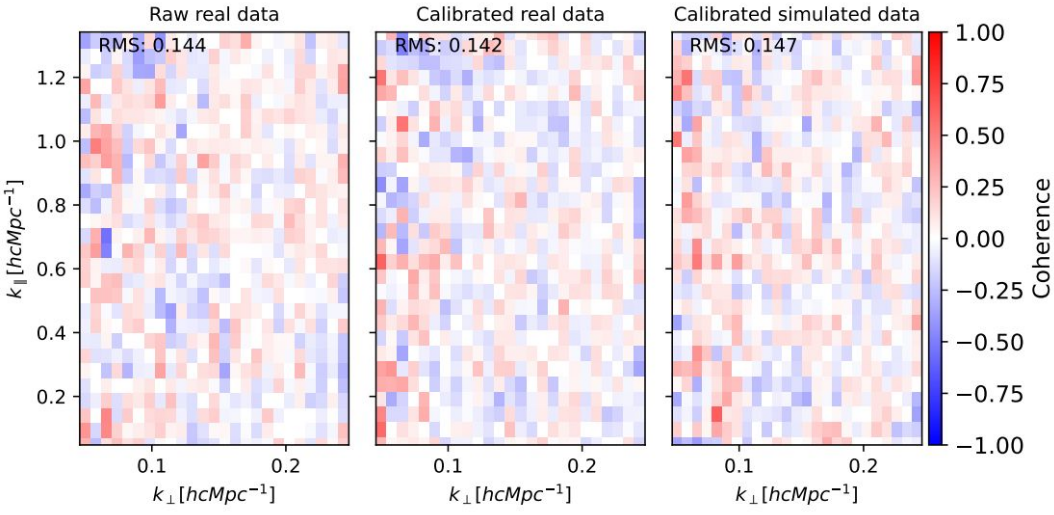
- Lossy compression of visibilities using Dysco (Offringa 2016) reduces data volumes, saving on crucial storage and compute costs
- Compression noise power spectrum is orders of magnitude lower than thermal noise spectrum

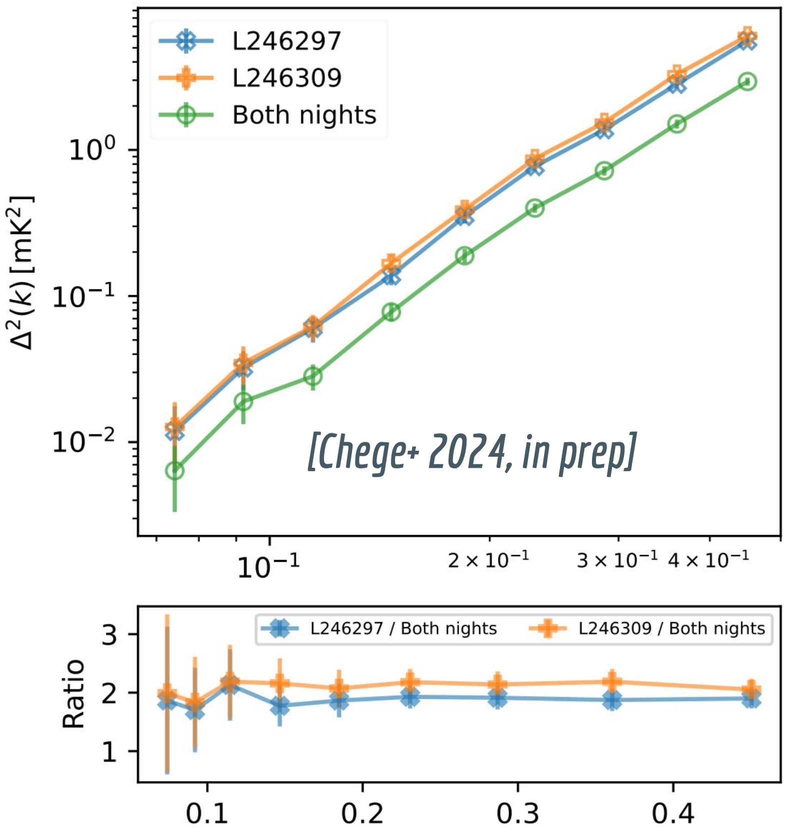
[Chege+ 2024, in prep]



Dysco compression noise coherence

- Compression noise PS is incoherent
- Consistent with random noise
- Averaging down 2 nights (same duration and LTS)
- A factor of 2 lower noise





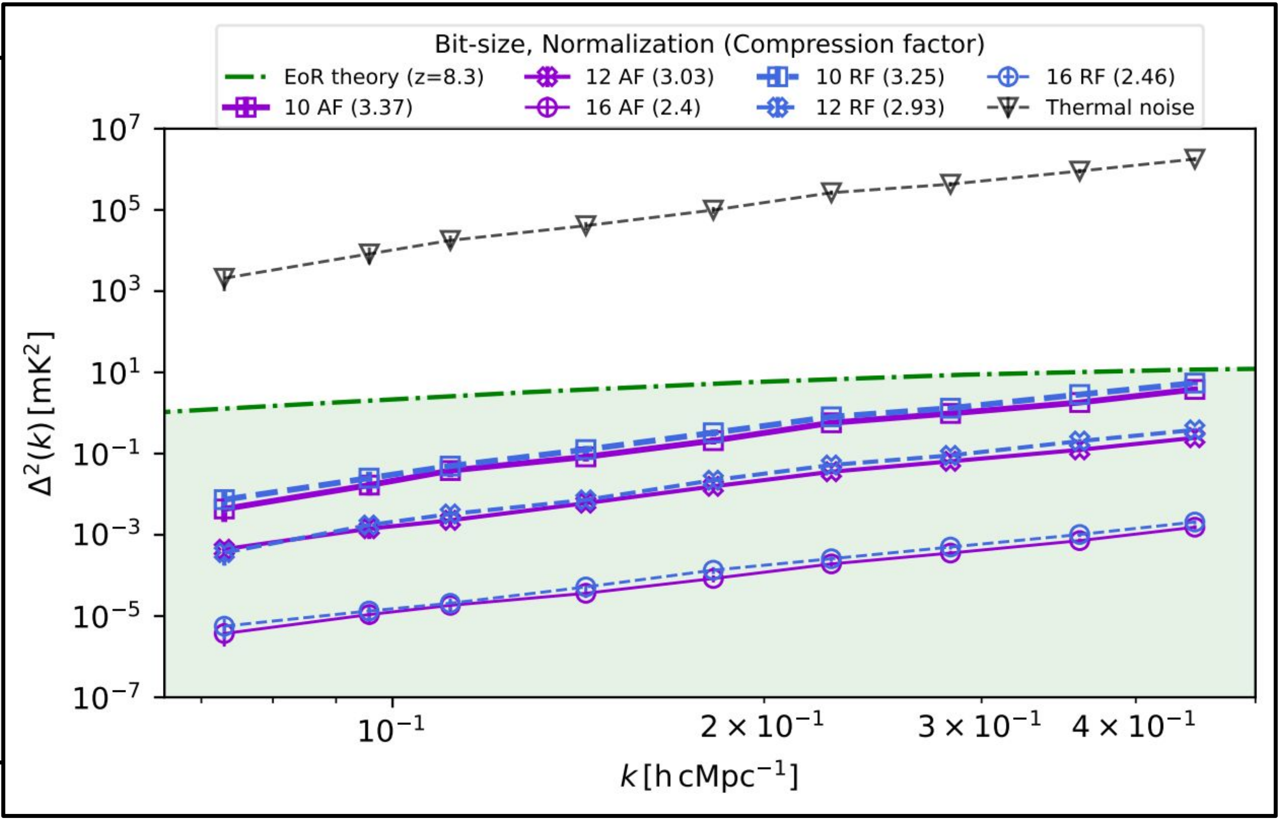
Optimal Dysco compression settings for EoR?

Compression noise scales with higher bit-rate

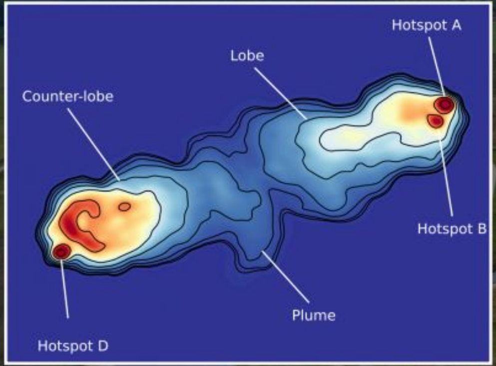
Compression factors of 3-4 achievable

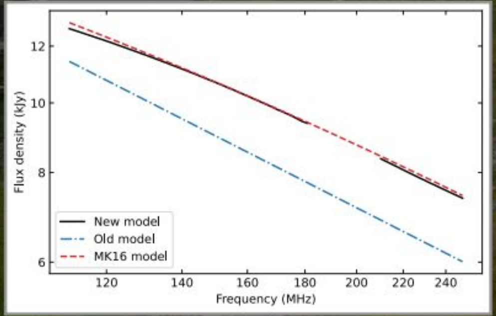
Still way below the EoR signal for even 6hrs of data

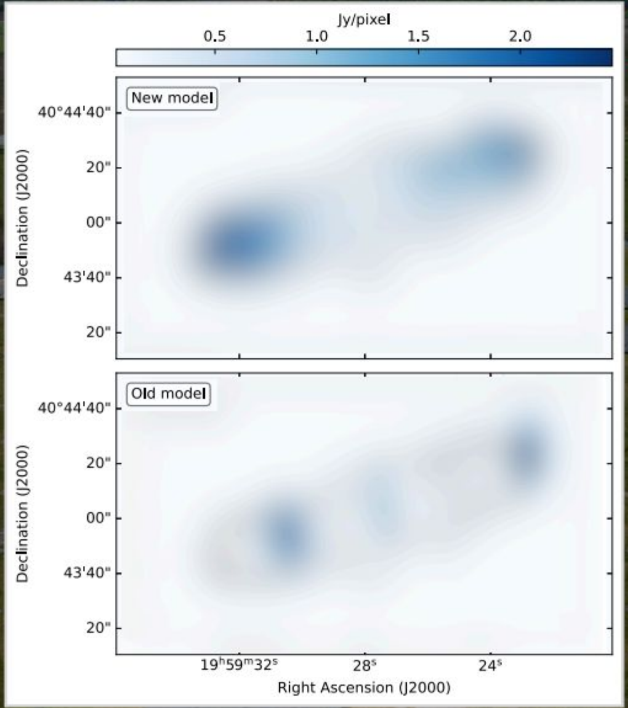
[Chege+ 2024, in prep]



IMPROVED MODEL OF CYGNUS A





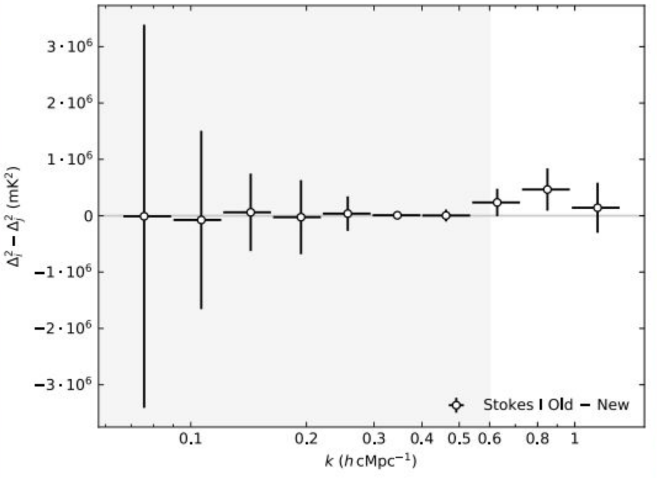


Ceccotti et al., in prep

Impact on power spectrum

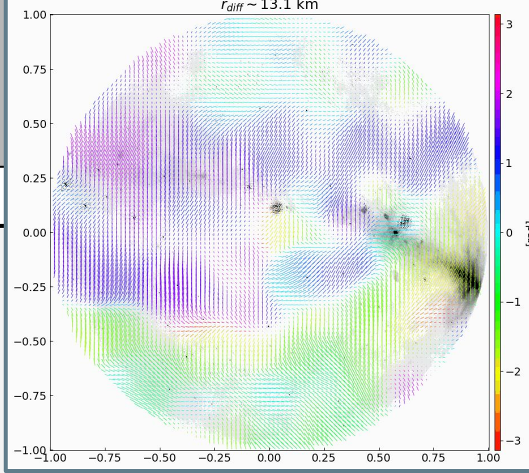
Emilio Ceccotti

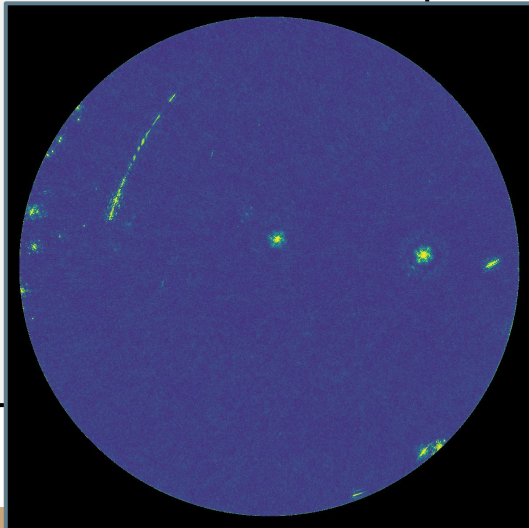
OBSERVATIONS



What is not causing the excess variance

- DI calibration looks pretty solid (Mertens et al+ in prep)
- The ionosphere is not the culprit (Brackenhoff+ in prep)
- Transient RFI is not a show-stopper (Gehlot+ 2020)
- Data compression is safe (Chege+ in prep)
- Improved sky models (always welcome) but not the main issue (Ceccotti+ in prep)
- GPR can decompose the residual visibilities into its separate constituents (Mertens+ 2023)





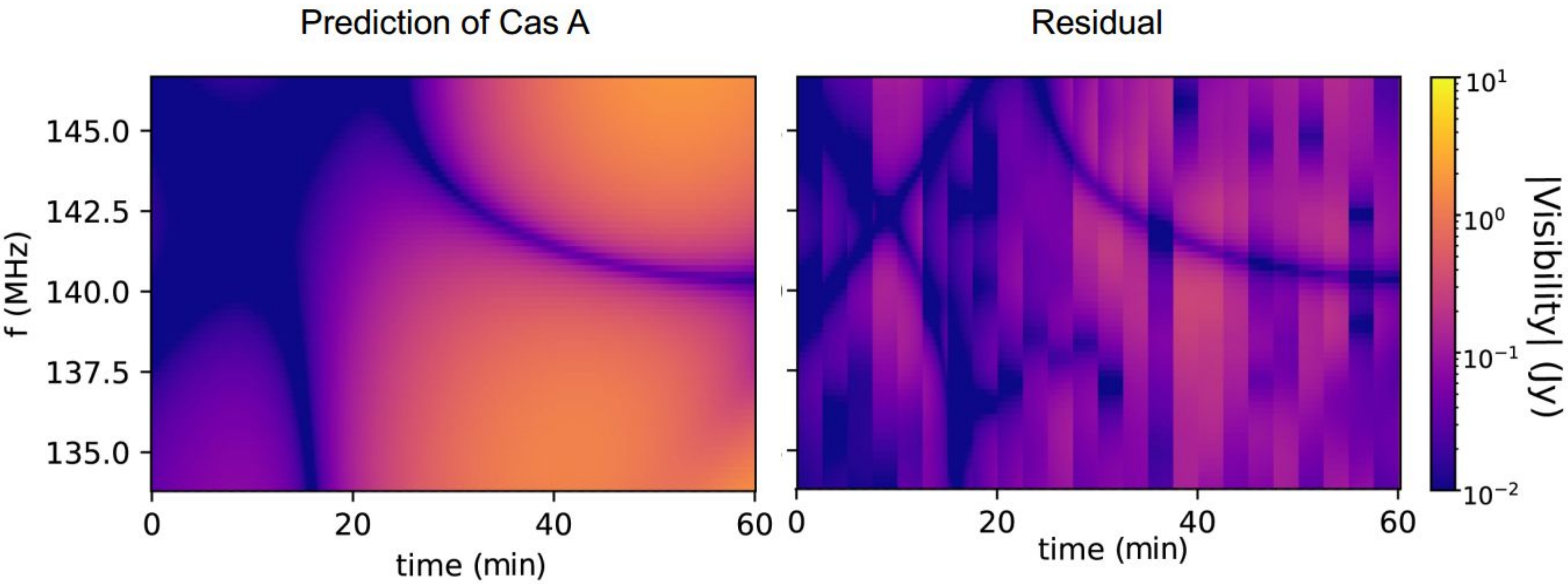
What could be causing the excess variance

- **Beam models** especially when **bright distant sources** interact with the nulls
- Degrees of freedom in **DD calibration** - optimise the directions solved for (See Stefanie's talk next).
- **Local and faint RFI** - Better modelling and filtering is required

Beyond Just Beam Errors

Stefanie Brackenhoff

- A-team sources limit our sensitivity and flagging for them leads to data loss
- How we need to change our calibration strategies due to far-field sources



university of groningen
faculty of science and engineering
Kapteyn Astronomical Institute

End-to-End automated processing pipelines

Features:

- ❑ End-to-end interfacing of the current EoR and CD pipelines
- ❑ Enables tools versions, parameters and results bookkeeping
- ❑ Computation resource usage specification & reports
- ❑ Modularized and easy to modify parameters
- ❑ Configuration and modules separation
- ❑ Standardized diagnostic metrics

On DAWN and to be ported to CODEX compute clusters

LOFAR EoR Analysis Pipeline ⟶ NextLEAP

NenuFAR Cosmic Dawn Pipeline ⟶ NenuFlow

```
N E X T F L O W  ~  version 22.10.4
Launching `/home/users/chege/theleap/leap/main2.nf` [grave_poitras] DSL2 - rev
[55/b3a6d7] process > VET_PARAMS:CheckingParams              [100%] 1 of 1 ✔
[20/f68e01] process > VET_PARAMS:SavingParams                [100%] 1 of 1 ✔
[55/d856fb] process > SET_NODES:MpiRunNodesList              [100%] 1 of 1 ✔
[c2/ad0b99] process > SET_NODES:PsshNodesList                [100%] 1 of 1 ✔
[e8/64977c] process > SAGECAL_MPI_DI:GetModels               [100%] 1 of 1 ✔
[3a/9b98f5] process > SAGECAL_MPI_DI:AddMsColumn             [100%] 1 of 1 ✔
[a4/1ee4ef] process > SAGECAL_MPI_DI:PreDIFlag               [100%] 1 of 1 ✔
[cb/389845] process > SAGECAL_MPI_DI:SagecalMPI              [100%] 1 of 1 ✔
[af/3deb21] process > SAGECAL_MPI_DI:makeEffNr               [100%] 1 of 1 ✔
[a7/9f17d0] process > SAGECAL_MPI_DI:convertSageSolutions    [100%] 1 of 1 ✔
[81/99ba2f] process > SAGECAL_MPI_DI:convertSageZSol         [100%] 1 of 1 ✔
[54/f0409c] process > SAGECAL_MPI_DI:plotDISageSols          [100%] 1 of 1 ✔
[10/abd1a0] process > SAGECAL_BANDPASS:AddMsColumn           [100%] 1 of 1 ✔
[50/43031f] process > SAGECAL_BANDPASS:SagecalStandalone     [100%] 1 of 1 ✔
[4a/85c178] process > SAGECAL_BANDPASS:ImageWSClean          [100%] 1 of 1 ✔
[ce/66ae15] process > SAGECAL_BANDPASS:makeEffNr             [100%] 1 of 1 ✔
[10/34b441] process > SAGECAL_BANDPASS:convertSageSolutions  [100%] 1 of 1 ✔
[1b/eb9d9d] process > SAGECAL_BANDPASS:plotDISageSols        [100%] 1 of 1 ✔
[7a/e96bfd] process > GEN_003_VIS:AverageVisTo003            [100%] 1 of 1 ✔
[00/837f34] process > SAGECAL_MPI_DD:GetModels               [100%] 1 of 1 ✔
[73/4ea265] process > SAGECAL_MPI_DD:AddMsColumn             [100%] 1 of 1 ✔
[ff/47c8ce] process > SAGECAL_MPI_DD:PreDDFlag               [100%] 1 of 1 ✔
[07/b8528f] process > SAGECAL_MPI_DD:SagecalMPI              [100%] 1 of 1 ✔
[3c/d00794] process > SAGECAL_MPI_DD:ImageWSClean            [100%] 1 of 1 ✔
[8f/e5772c] process > ANALYSE_GAINS:makeEffNr                [100%] 1 of 1 ✔
[43/1d8d6f] process > ANALYSE_GAINS:convertSageSolutions     [100%] 1 of 1 ✔
[8d/e5dd5e] process > ANALYSE_GAINS:convertSageZSol          [100%] 1 of 1 ✔
[66/fbd6fc] process > ANALYSE_GAINS:plotDDSageSols           [100%] 1 of 1 ✔
[57/3545ba] process > POWER_SPECTRUM:addRevision             [100%] 1 of 1 ✔
[9a/63ad3f] process > POWER_SPECTRUM:runPSPIPE               [100%] 1 of 1 ✔

Completed at: 06-May-2023 00:39:50
Duration    : 1h 57m 47s
CPU hours   : 11.2
Succeeded   : 30
```

nextflow

The next big leap

- Automated end-to-end data processing workflow (NextLEAP)
- Processing and analysing 100+ nights of NCP data at multiple-redshifts
- New GPU cluster (CODeX) at RuG/CIT-HPC
- Analysing more nights on 3C196 field
- Incorporating new calibration and GPR methods.

Next on Nenufar

Quantifying the full-sky foreground wedge

[Munshi et al. in prep]

- Foregrounds can extend well beyond previously thought
- Contributes to excess

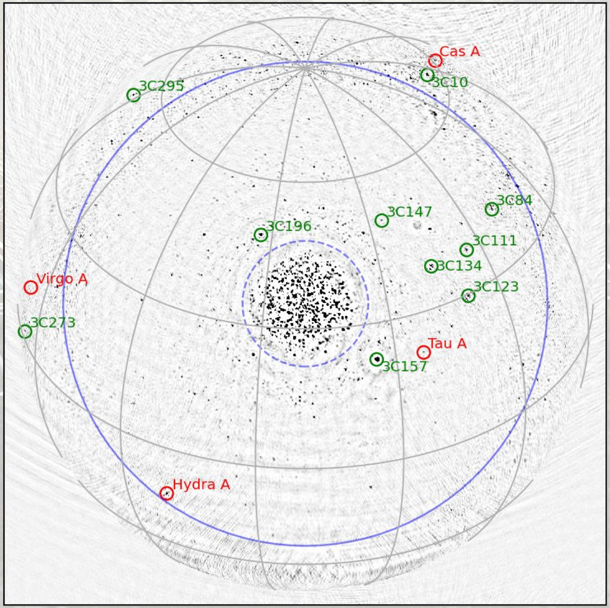
Filtering out local RFI

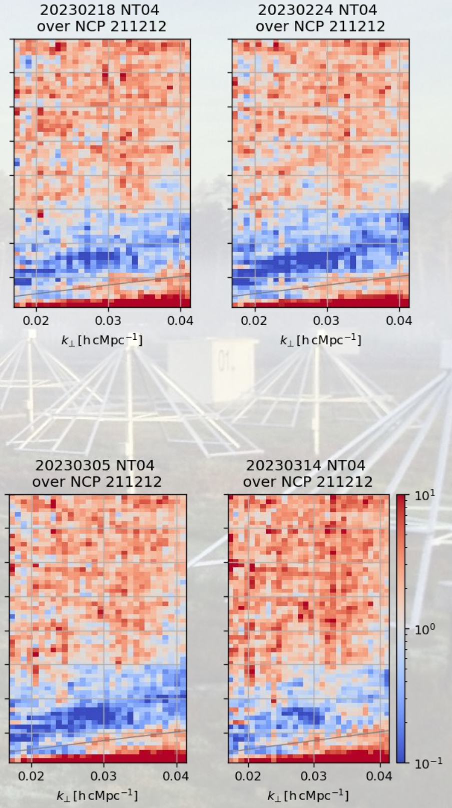
[ongoing]

- Characterize and filter out from the data

Observation of a new 'dark' field *[ongoing]*

- 5 fields surveyed - 1 chosen
- Preliminary results - promising!
- More than 250 hours observed





- HI is the only tracer (21-cm line) that allows us to study many astrophysical processes during the Cosmic Dawn & EoR over wide range of angular scales.
- No detection yet, but increasingly stronger statistical (power-spectra) upper limits
- We are scaling up significantly! (5000+ hours, 3+ sky fields, new compute cluster)
- Dysco compression will help mitigate resulting compute resource expenses without introducing additional errors
- We are ready to make biggest analysis steps to date for LOFAR & NenuFAR
- Made possible by refined processing strategies, great LOFAR tools and now, new efficient processing workflows

NextLEAP | NenuFlow

Thank You!