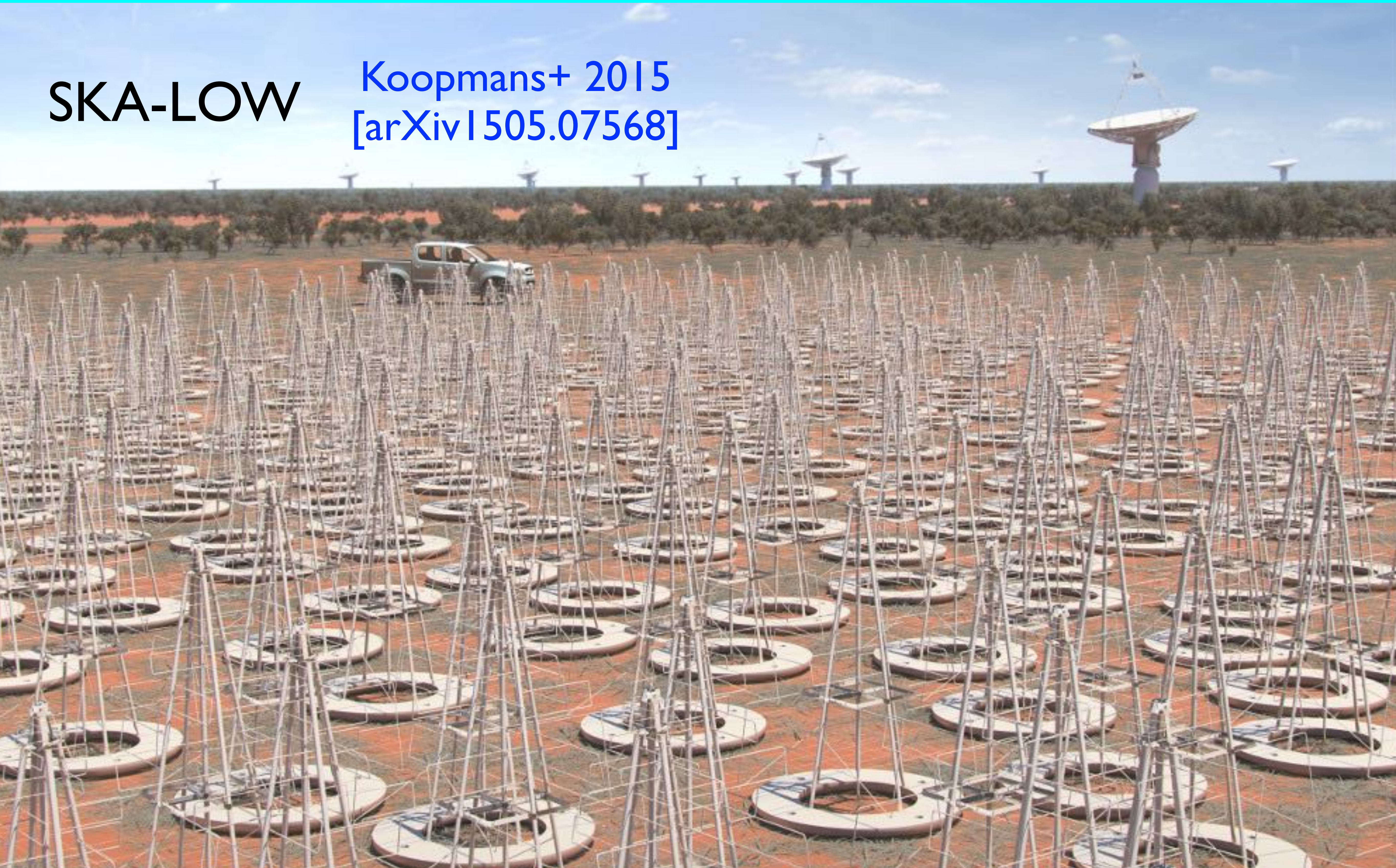
# Epoch of Reionization and Cosmic Dawn
# Cost Control Interim Response



SKA-LOW

Koopmans+ 2015
[arXiv1505.07568]

# Overview

- EoR Science Overview
- Impact of Cost control Options
- Progress towards a KSP bid

Many questions still under investigation

LOFAR

MWA

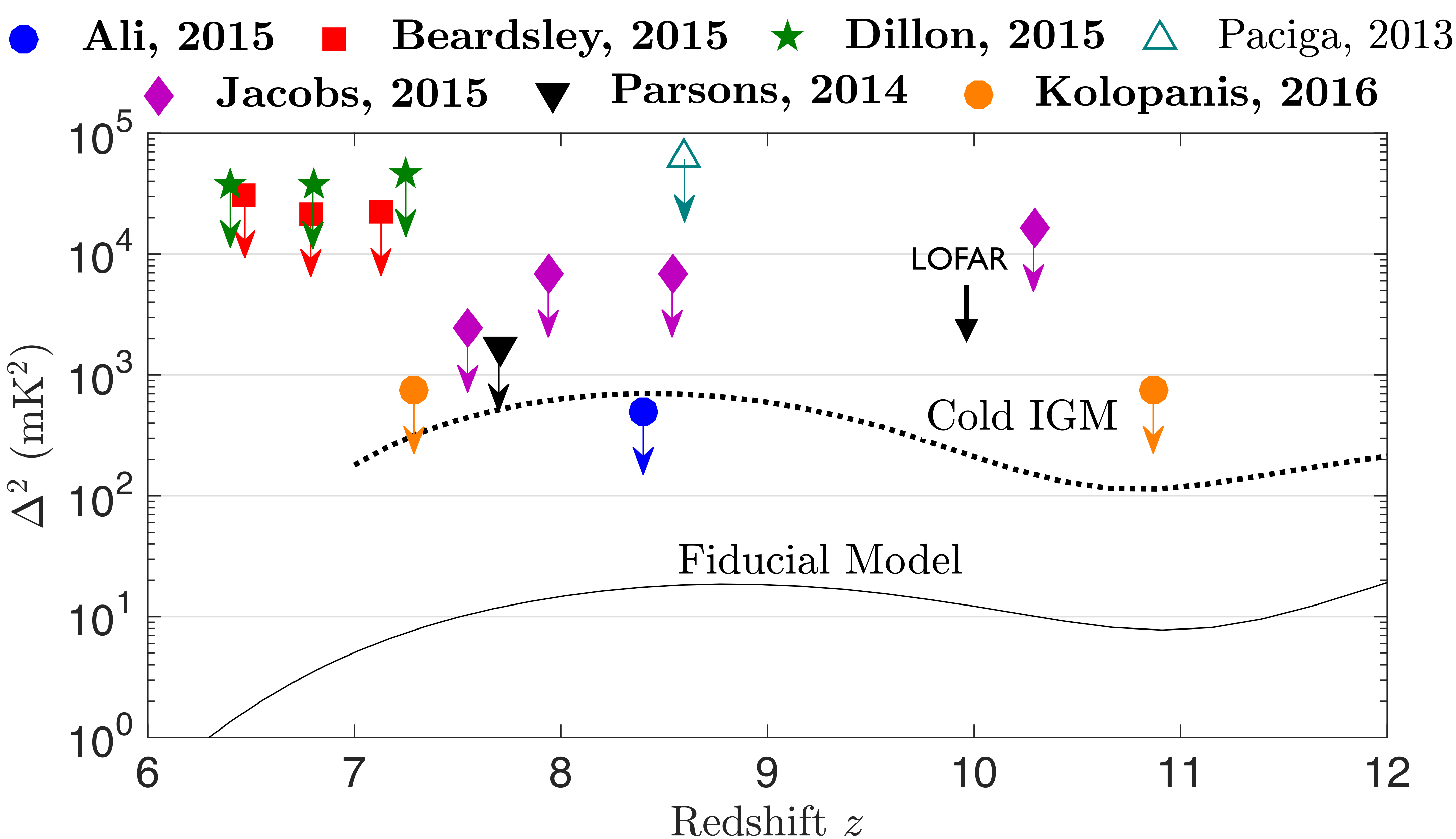
PAPER

ICIC

HERA

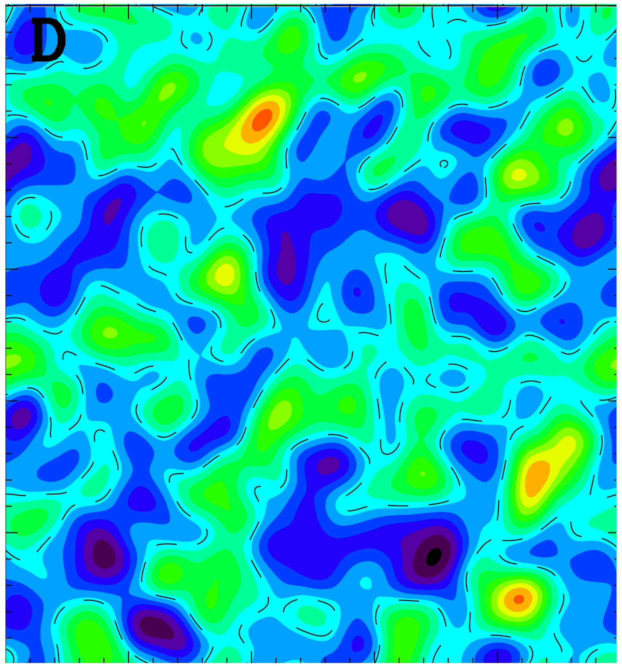
SKA

# Continuing improvement

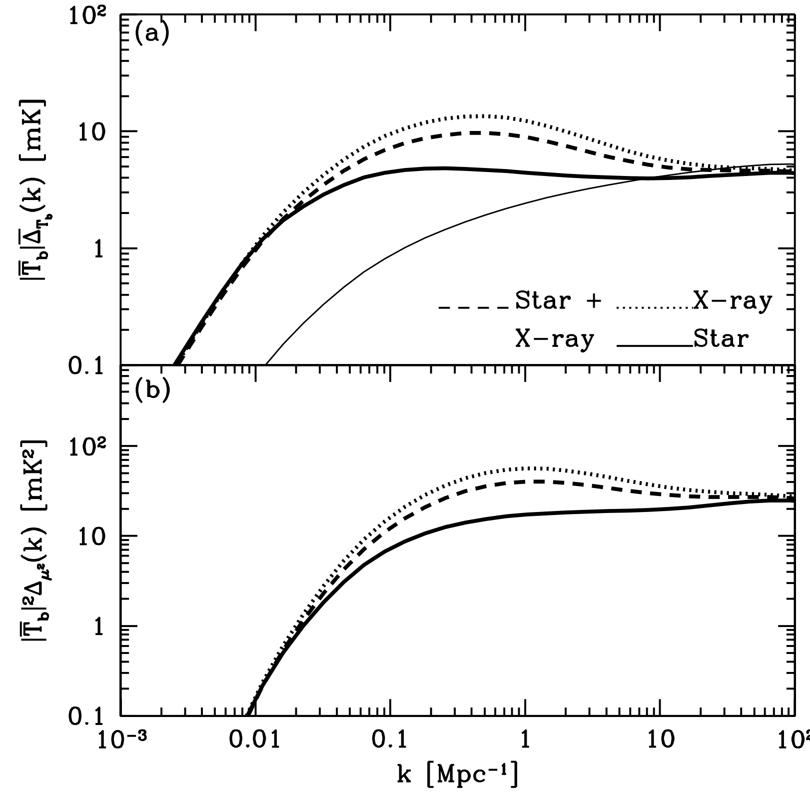



Upper limits beginning to make contact with possible models

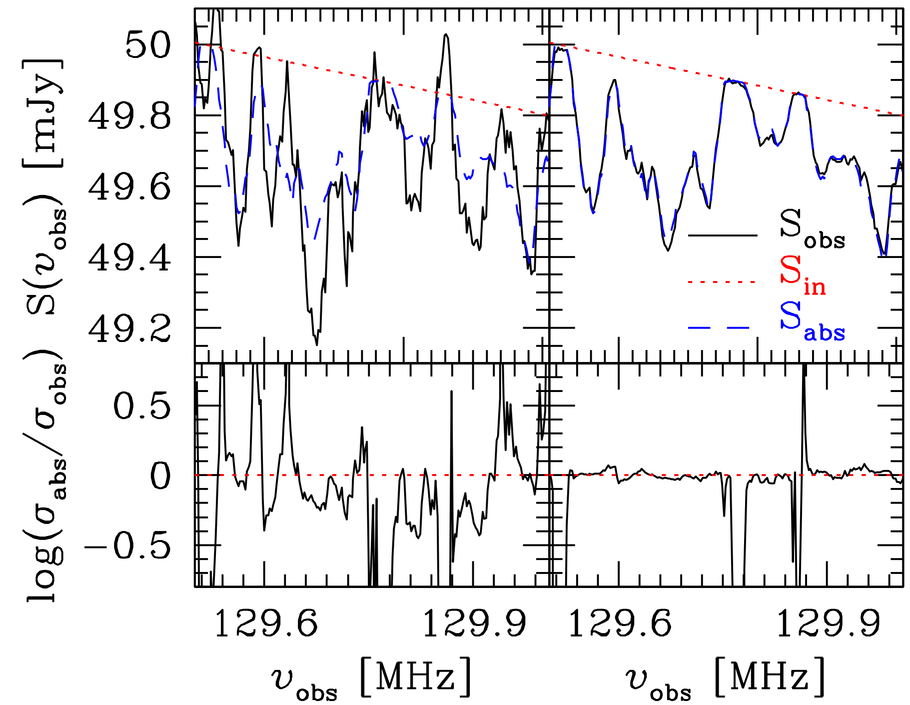
# Images, power spectra, 21cm forest



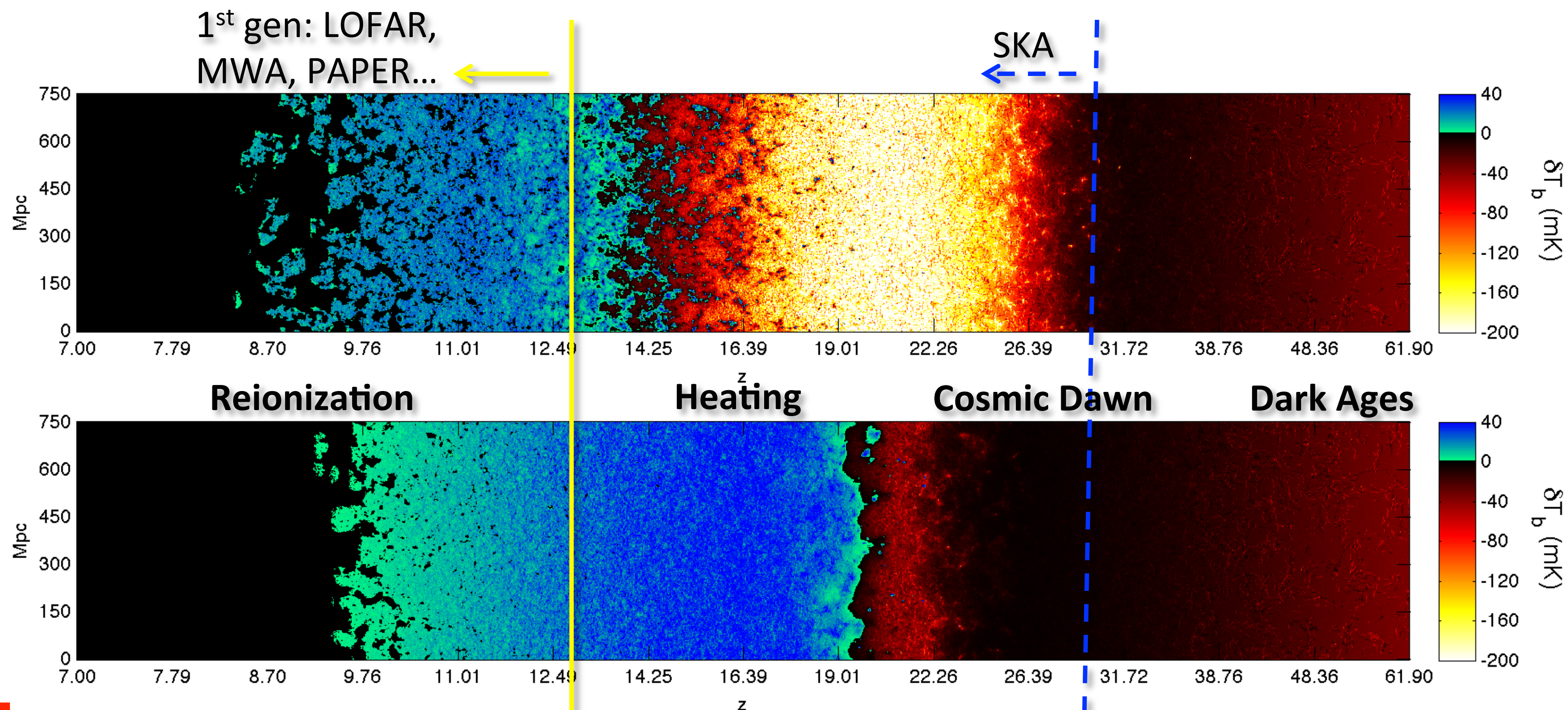

Zaroubi+2013




Pritchard & Furlanetto 2007



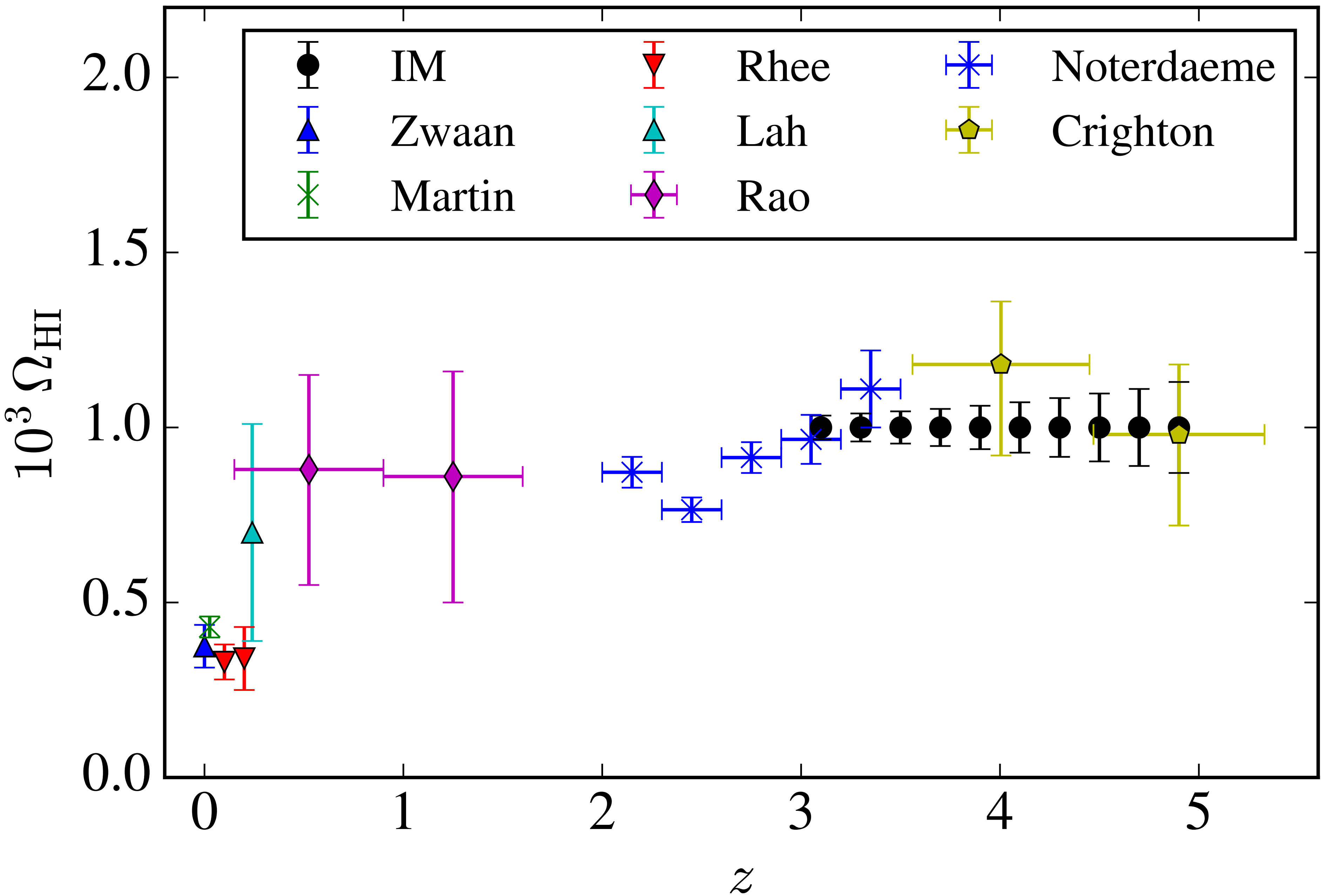

Ciardi+ 2015




Valdes+ 2012

# Intensity mapping at 3<z<5

Intensity mapping with LOW for OmegaHI & cosmology



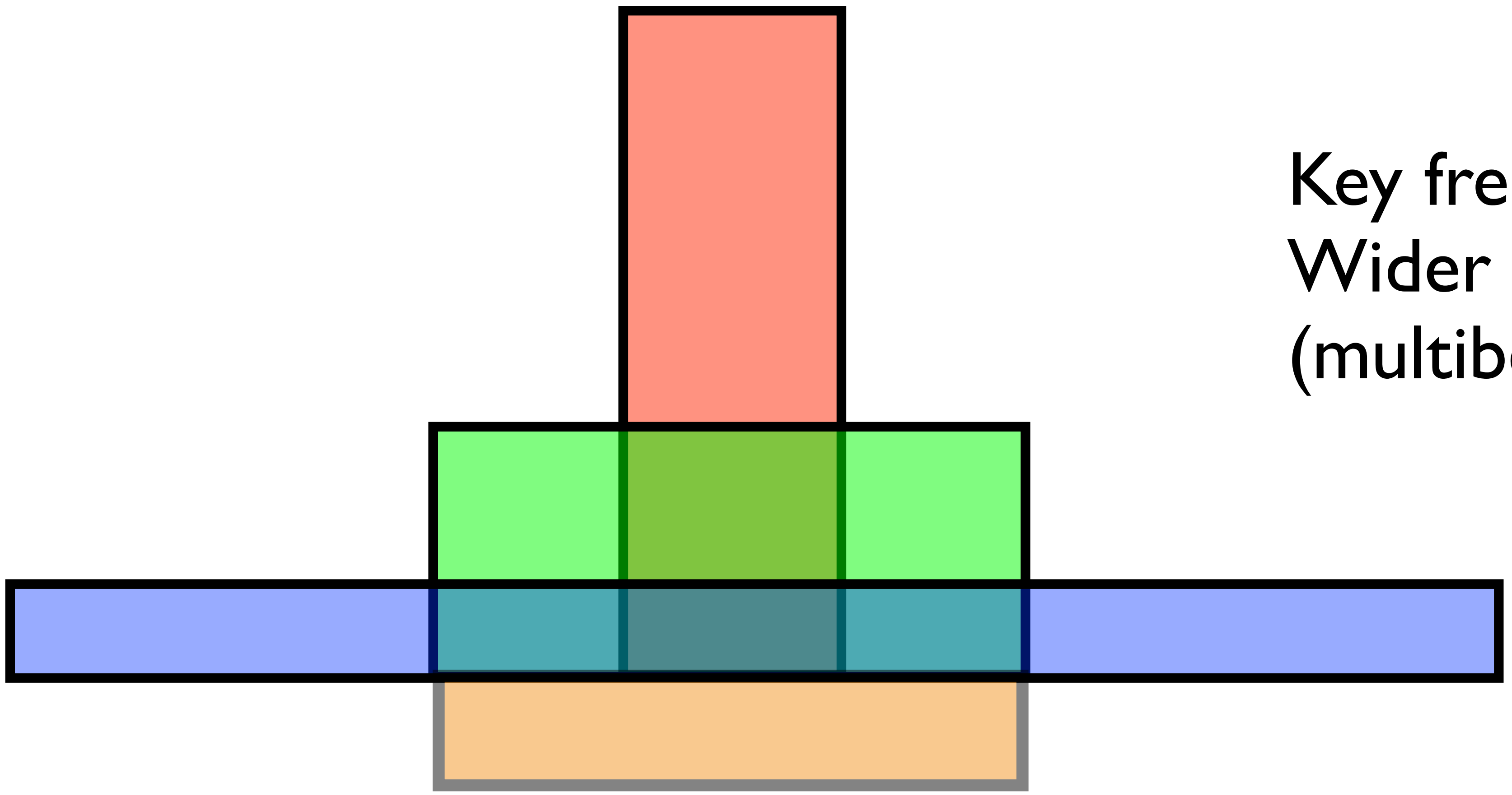

Preliminary calculation

Pourtsidou+

# SKA observing strategy

Deep: 5 x 1000hr integration => 100 deg$^2$ field
Middle: 50 x 100hr integration => 1,000 deg$^2$ field
Shallow: 500 x 10hr integration => 10,000 deg$^2$ field
IM: 50 x 100hr integration => 1,000 deg$^2$ field

Koopmans+ 2015 [arXiv1505.07568]



Key frequencies: 200-50MHz
Wider band for foregrounds
(multibeaming to reduce tint)

Shallow: LOFAR-like power spectrum sensitivity over 10000 deg$^2$.
Middle: Shallow imaging + power spectrum over 1000 deg$^2$
Deep: Power spectrum to z<28 & deep imaging over 100 deg$^2$
IM: OmegaHI & cosmology at 3<z<6 over 1000 deg$^2$

# Impact on EoR of cost control options

# No effect on EoR

EoR is essentially focussed on LOW.
MID useful for some synergy science, but not priority

Some items have no impact

| | | | | |
|---|---|---|---|---|
| 5.34 | **Maximise use of code produced during Pre-Construction** Pre-existing code is being identified for re-use in SKA. This may include code developed in precursors, code developed as part of other projects (typically open source) and code developed in pre-construction as part of prototyping. | **COMMON** | **None** | 1 |
| 5.38 | **Simplify DDBH LOW** The Digital Data Back Haul (DDBH) is the science data network connecting the antenna stations to CSP. This option explores the possibility to simplify the DBBH for LOW, going from a managed network design (exploiting network switches to forward data) to a point-to-point connection (where data communication is established by other means). | **LOW** | **None** | 1 |

Pulsar science is awesome, but doesn't impact EoR

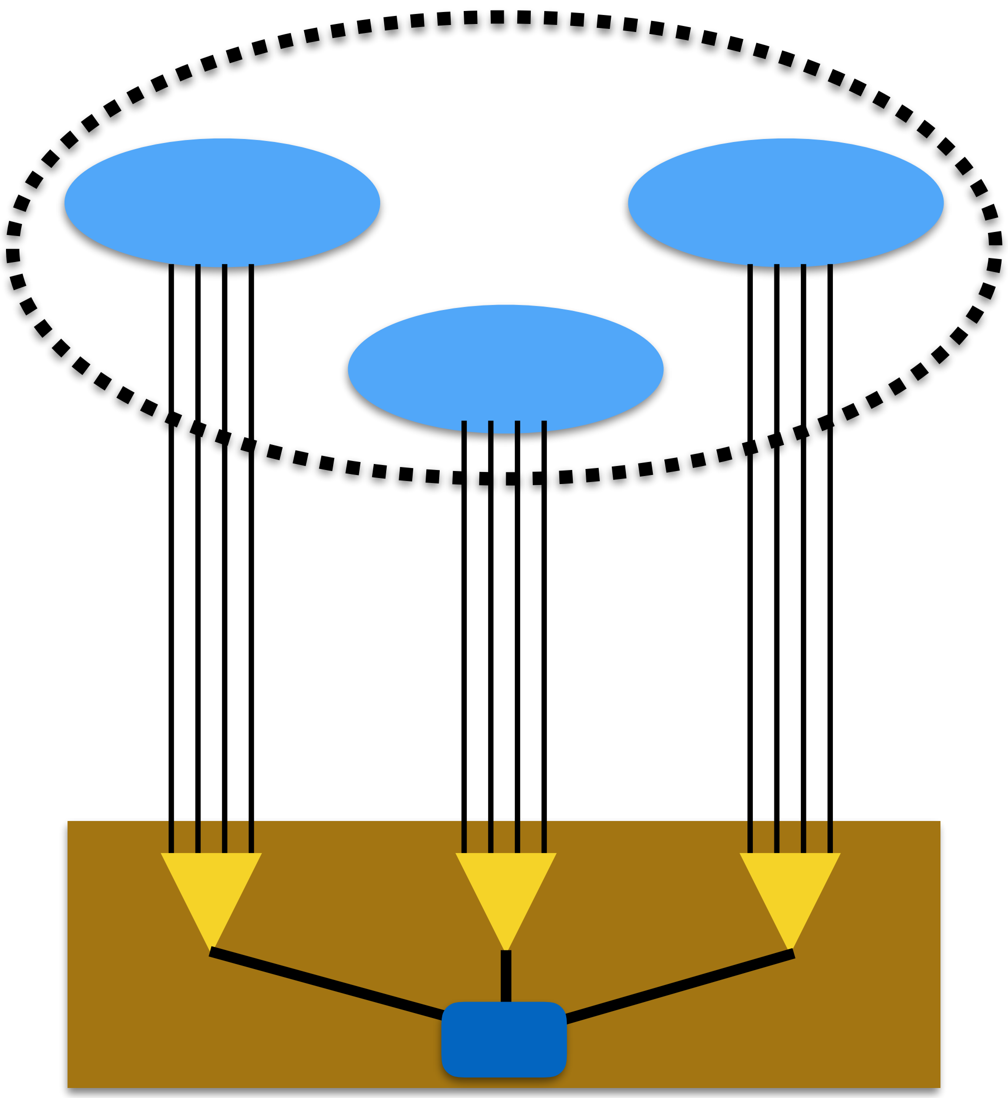
| | | | | |
|---|---|---|---|---|
| 5.25.2 | **Reduce PSS-LOW: A, 250 nodes to 167 nodes** This cost control option involves necessitating that the CSP.PSS design processes 3 tied array search beams per PSS processing node on LOW. Currently the design processes 2 tied array search beams per PSS processing node on LOW. To achieve this would require improved processing algorithms (which may not be possible) or the reduction in search parameter space (i.e. not performing a complete pulsar search so not achieving the same science). It is considered that the change from 2 beams/node to 3 beams/node will potentially be possible without needing to perform an incomplete search. If it is not possible this equates to a cut in the number of pulsar beams that can be processed across the full search parameter space by 166, i.e. from 500 to 334 beams. | **LOW** | **Likely none, or small reduction of pulsar search parameter space.** | 1 |

et cetera

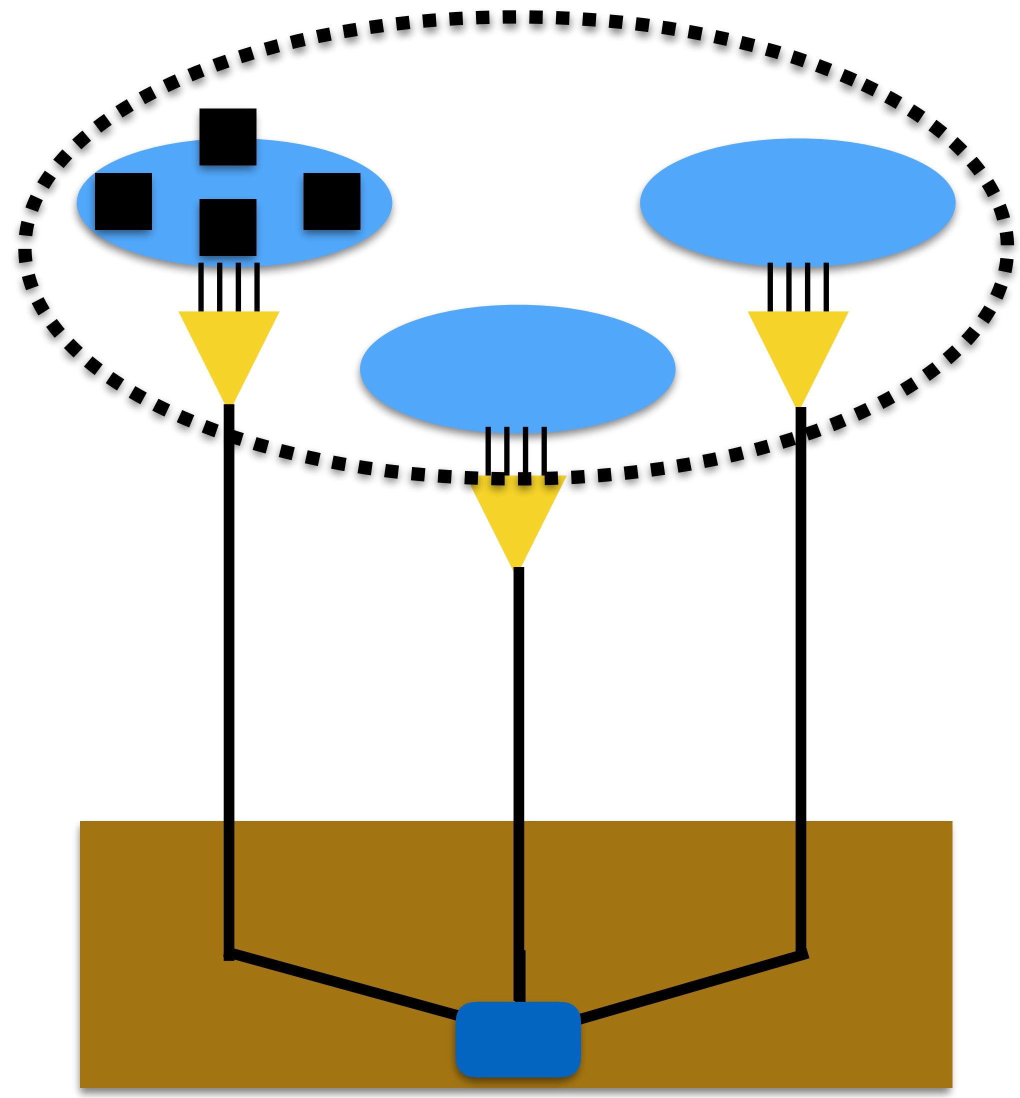
# Beam forming

| | | | | |
|---|---|---|---|---|
| 5.26 / 5.29 | **LOW RPF: Early Digital Beam Formation vs. Analogue Beam Formation** This cost control option involves a change to the beam forming architecture by placing either a digital or analogue beam-former immediately adjacent to the stations, thereby eliminating long distance analogue signal transport. The chosen option would need to satisfy the current System requirements. | LOW | None | 1 |

Details? - semi-real time beam calibration? Gains at 16 dipole level?



Station level beam forming



Tile level beam forming (analog or digital)

# Beam forming

- MWA and LOFAR both make use of tile level beam forming
- Currently not the limiting factor in their EoR observations. Not enough experience to say if it might limit at deeper integrations. DD calibration ought to absorb residual beam errors(?)
- MWA experience illustrates importance of good design e.g. bit depth for analog delay system, sensible (irregular) arrangement of dipoles within tiles, etc.
- Retaining flexibility in analysis is desirable. Gain corrections at dipole level? Unclear how much computation allows?
- Concerns on fidelity of beam modelling required. Trade off between computational effort in beam forming versus calibration.
- Hard to capture impact in current simulations. Hence worrying. Actively being pursued e.g. with OSKAR

Chapman, Mouri, others

# Reduced processed bandwidth

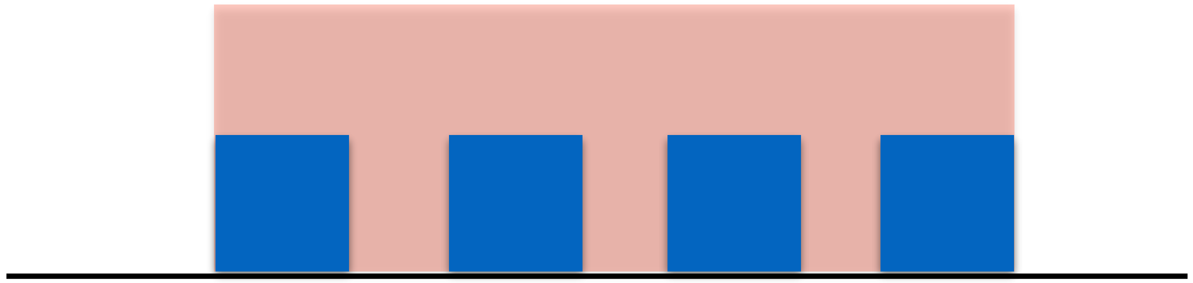
| | | | | |
|---|---|---|---|---|
| 5.31 | **Reduce CBF-LOW BW: A, 300 to 200 MHz** This option entails reducing the maximum bandwidth processed by the correlator from 300 to 200 MHz. Continuum applications may require longer integration times to achieve the same sensitivity. Multi-beamed imaging observations, that relied on 2x150 MHz beams, would instead be limited to 2x100MHz beams. | LOW | **Longer observing times for continuum applications (1.5x)** | 2 |

Desired survey: 7500hrs using 2 x 150 MHz beams

SHALLOW (10,000 sqdeg, 500 x 10 hour, 50-200 MHz), 5000 hrs/Nbeam
MIDDLE (1000 sqdeg, 50 x 100hr, 50-200 MHz), 5000 hrs/Nbeam
DEEP (100 sqdeg, 5 x 1000 hr, 50-200 MHz). 5000 hrs/Nbeam

Reduced bandwidth reduces scope for multi beaming essentially doubling total integration time to 15,000 hrs on sky

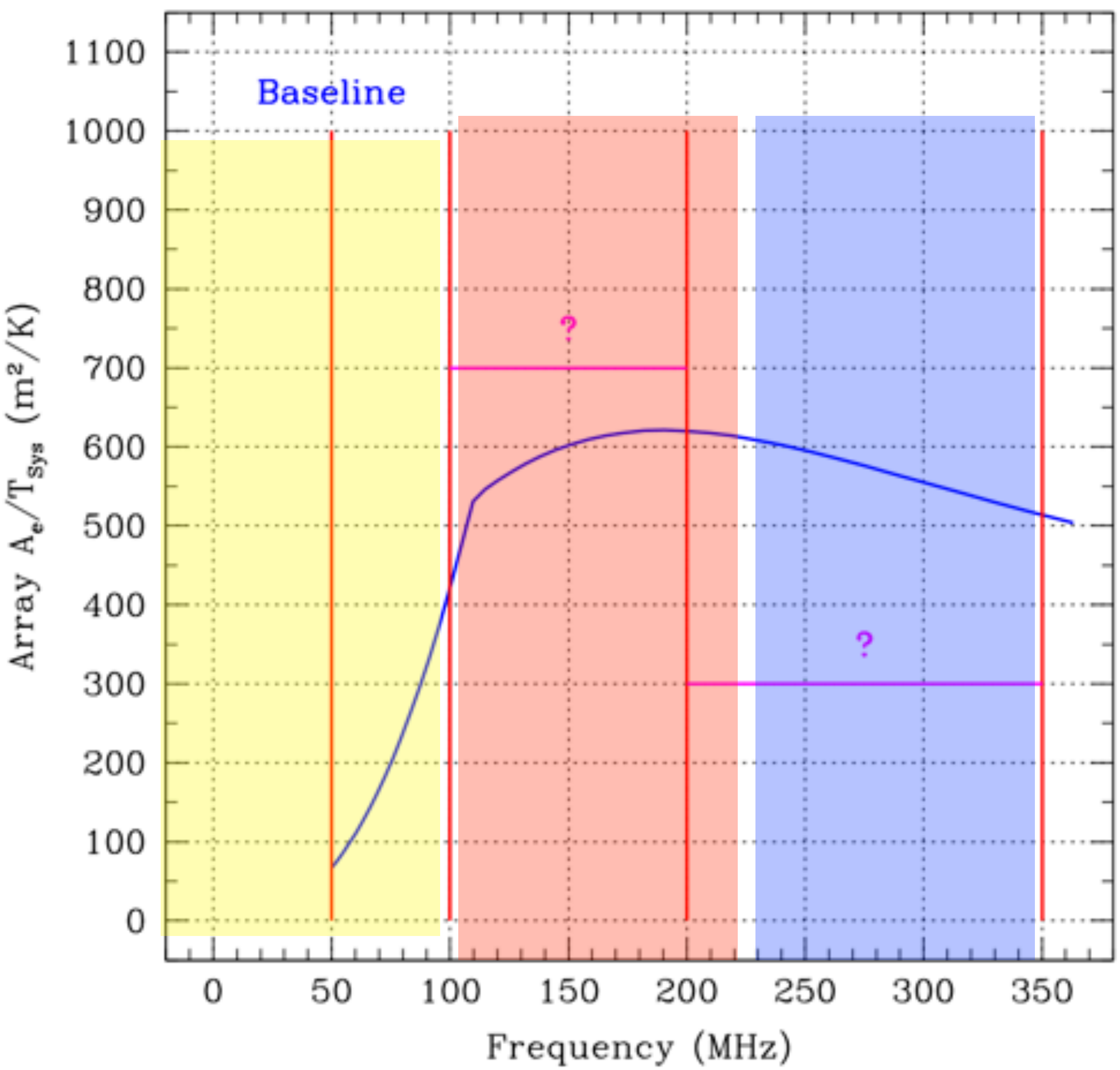
May be scope for redesigning surveys to compensate somewhat
- maybe focus on 100-200 MHz for SHALLOW to use multibeaming
- non-contiguous frequency bins e.g. 10MHz band w 5 MHz gap



Possible impact on ionospheric calibration (measuring TEC gradients for 1-10 mJy calibrator sources)

# Antennae design

| | | | | |
|---|---|---|---|---|
| 2 | **LOW Antenna: Log Periodic Design vs. Dipole Design**<br>This option refers to undertaking a choice of antenna that best matches the scientific needs and cost constraints. Since none of the current antenna designs meet the System Requirements, further work is needed to optimise the antenna solution. | LOW | **None of the current designs meet the L1 requirements** | 3 |




Some desire to preserve sensitivity to ~ 220-240 MHz (& no cliff please!)

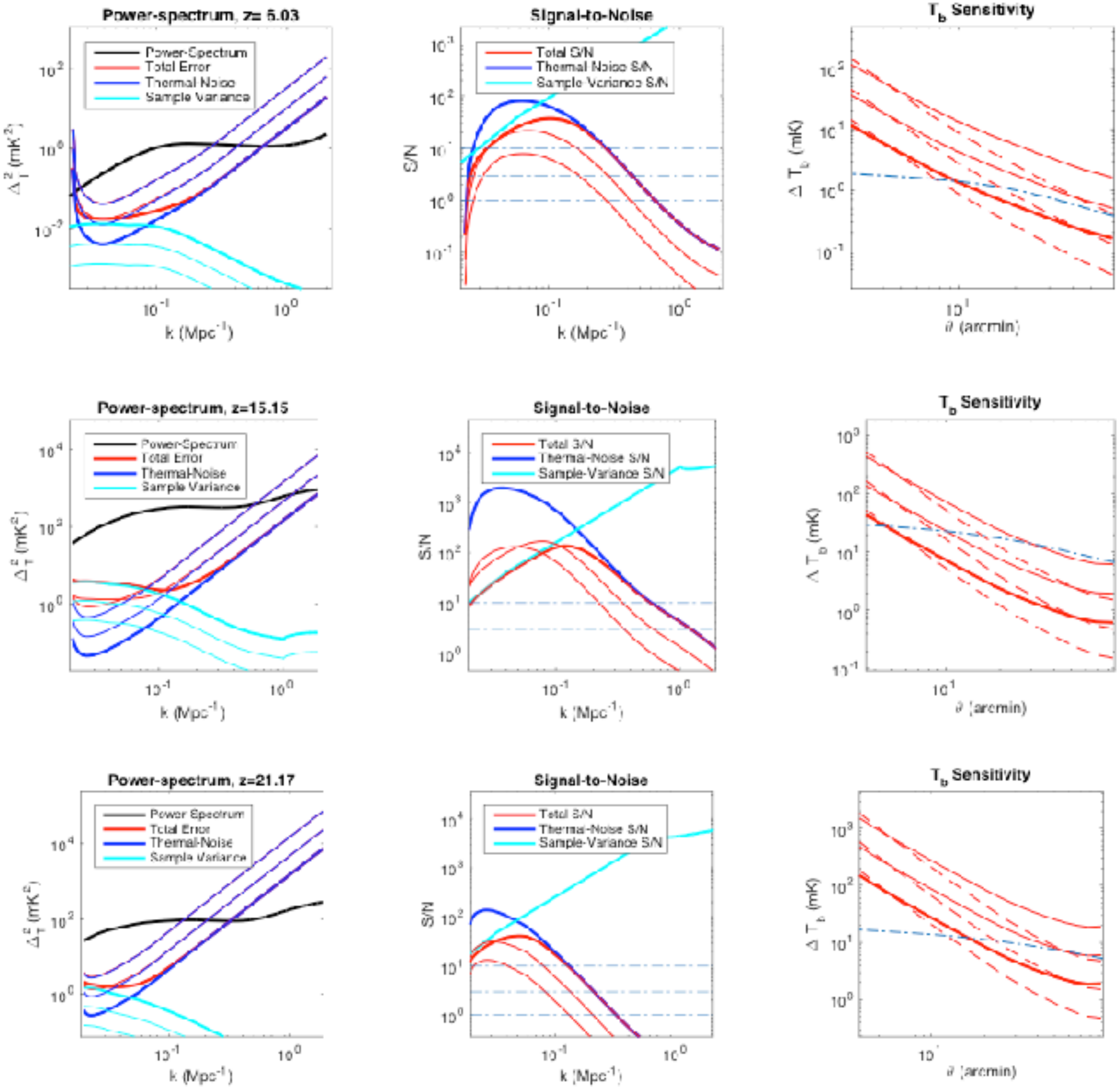
EoR target 50 - 220 MHz

Maybe something like:
Cosmic Dawn 50-100 MHz
Reionization 100 - 200 MHz
Reionisation tail
200-220 MHz (240 MHz?)

Recent observations lower redshift of reionisation

200 - 350 MHz lower priority for intensity mapping of HI in galaxies for OmegaHI/Cosmology

# Updated sensitivity curves




Koopmans

Including rebaseline & new antennae proposal (+20% @100-200MHz, …)

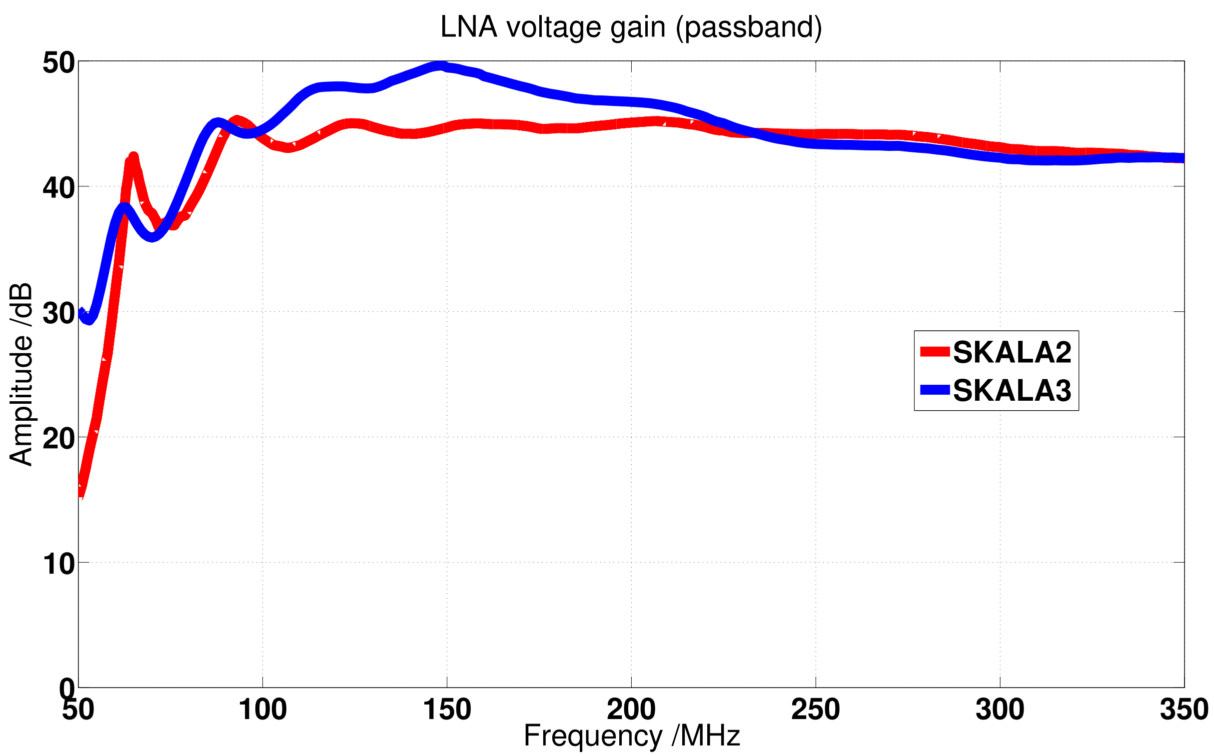
100-200 MHz:
may gain some S/N.

50 - 100 MHz:
Hard to reach z~25,
but good S/N at z=21

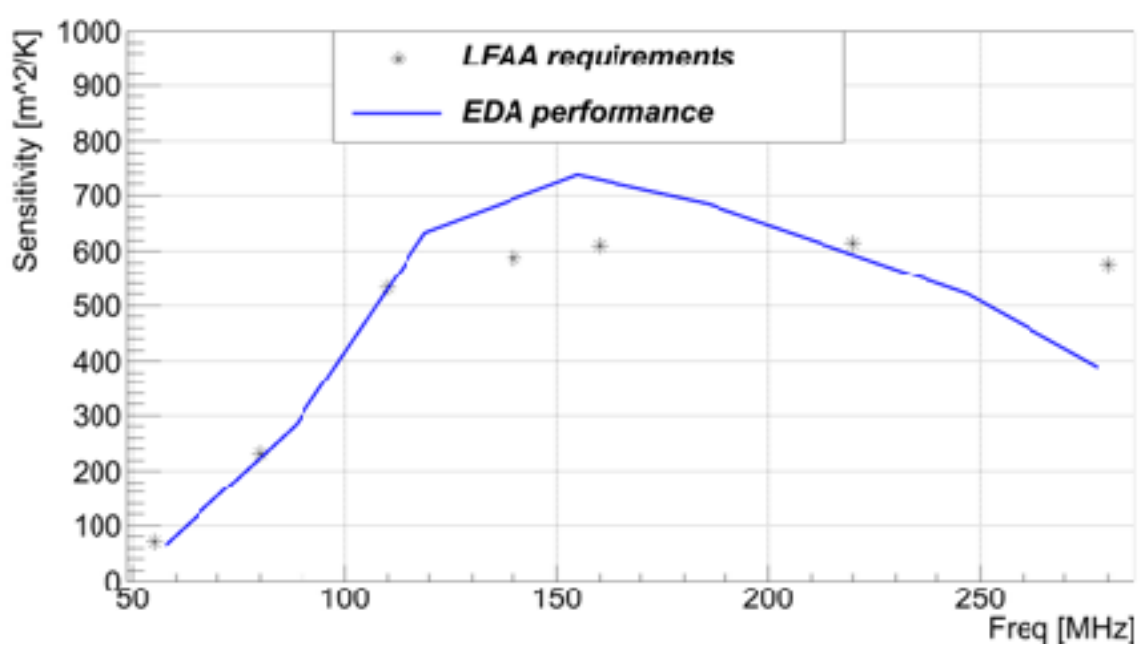
200-350 MHz:
- tail of reionisation z~5.5 (220MHz) harder to observe.

- IM at z<6 significantly affected

# Band smoothness



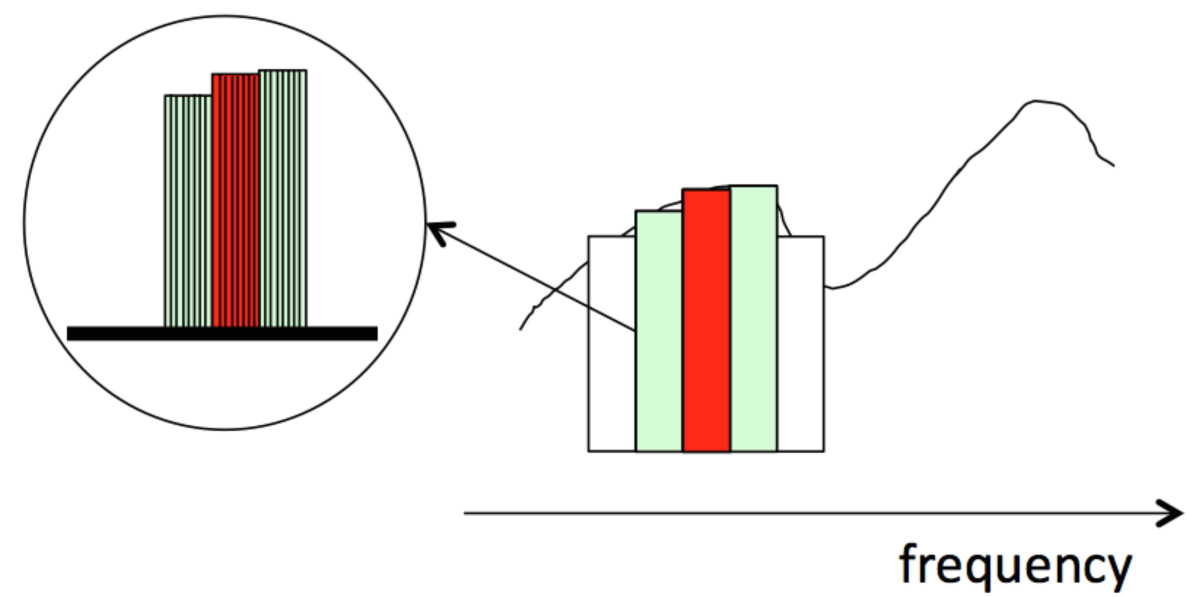

de Lera Acedo+




Wayth

Typically will calibrate bandpass on ~ 2MHz chunks. Residuals are issue

Trott & Wayth

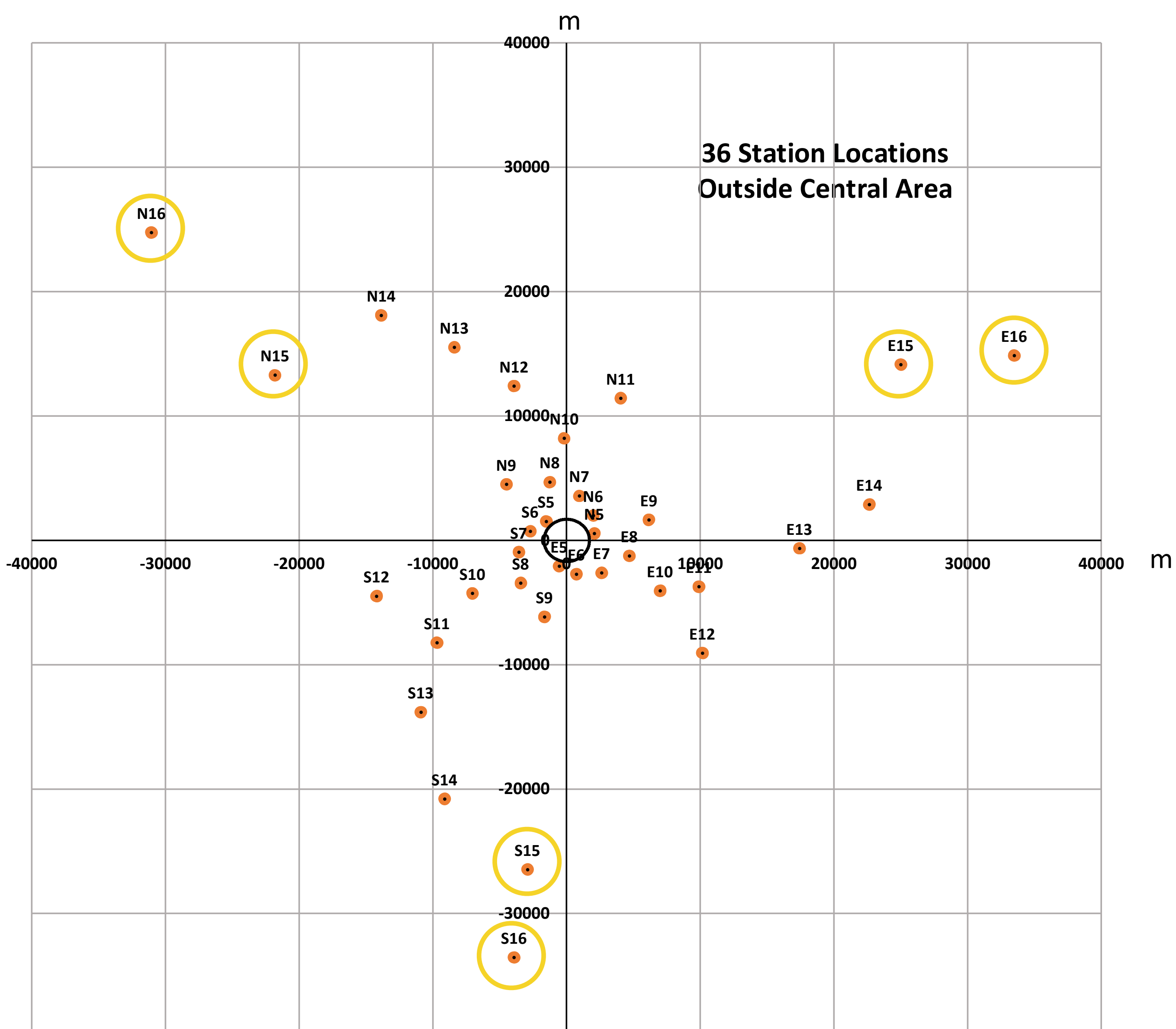



General preference for smoother frequency response in antennae. Stability an issue.

Critical to get this right!

# Reduced maximum baselines

| | | | | |
|---|---|---|---|---|
| 5.30.0 | **Reduce Bmax LOW to 50km: A, remove infra, add 18 stations to core** In this scenario, the outer three clusters of (6x3) stations are moved into the central region of 3km diameter. The reduction in maximum baseline may have adverse consequences for foreground continuum source characterisation and removal. | LOW | Science Risk to EoR: Bmax. | 3 |
| 5.30.0 | **Reduce Bmax LOW to 50km: B, remove 18 stations** Similar to the previous scenario where the outer three clusters of (6x3) stations are removed, but these are not added to the central core region. The reduction in maximum baseline may have adverse consequences for foreground continuum source characterisation and removal. | LOW | Science Risk to EoR: Bmax | 3 |
| 5.30a | **Reduce Bmax LOW to 40km: C, remove next 18 stations** This scenario involves removing the second outermost clusters of (3x6) stations as well as the supporting infrastructure. These are not added to the core region. The reduction in maximum baseline may have adverse consequences for foreground continuum source characterisation and removal. | LOW | Science Risk to EoR: Bmax | 3 |




Resolution team studying this has significant EoR input

- Ionosphere
- Calibration
- Foregrounds

Requires full simulation of calibration process to study

Ongoing work in Science Team

# Sky model for SKA

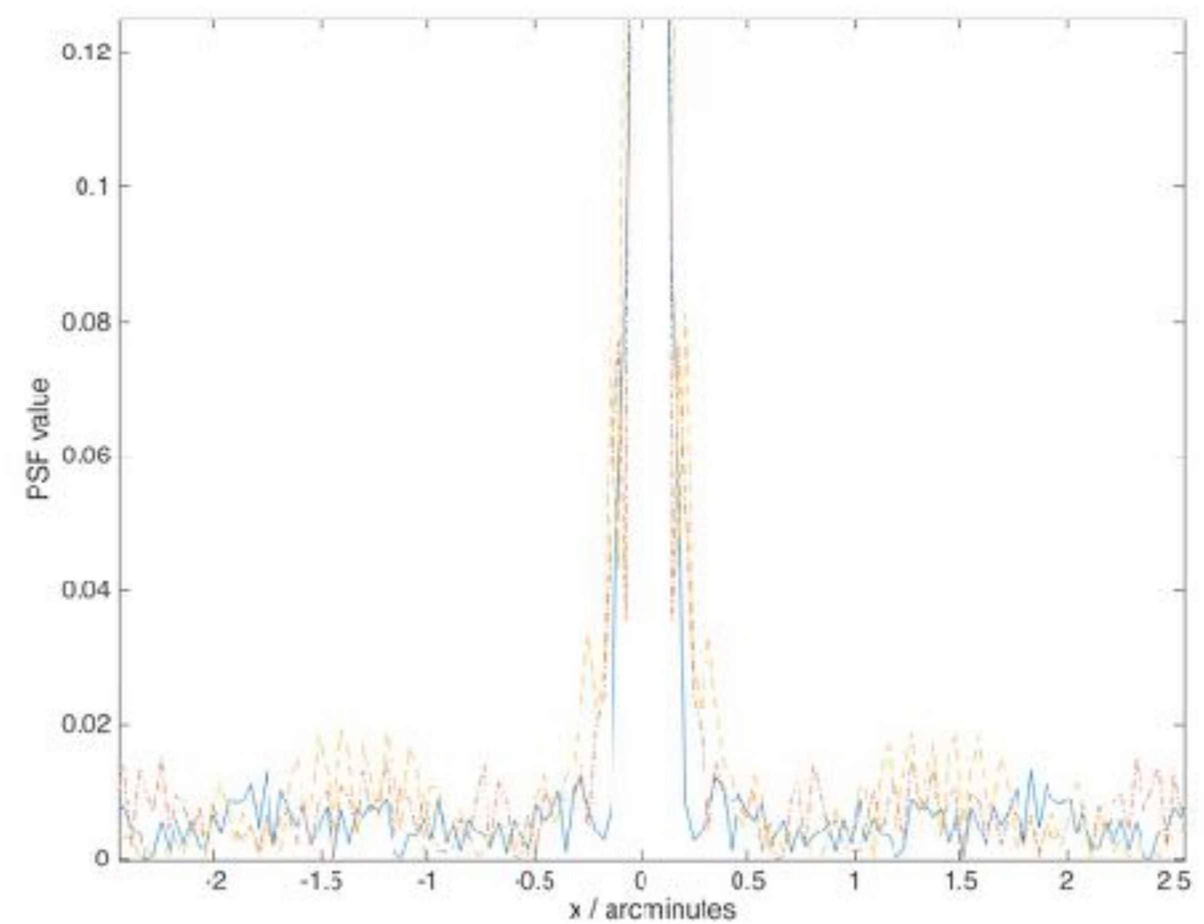
EoR imaging set by core, so here talking calibration/sky model

1) Use SKA-LOW to build sky model
2) Exploit high frequency surveys to build sky model

e.g. PUMA Line+

Working towards a full test
with OSKAR simulation of SKA,
SAGECAL calibration
& model of sky based
on FIRST, VLSSr+model faint pts

Should be cautious if interim
calculations are capturing all
the key points




Chapman

# Calibration

Direction dependent calibration

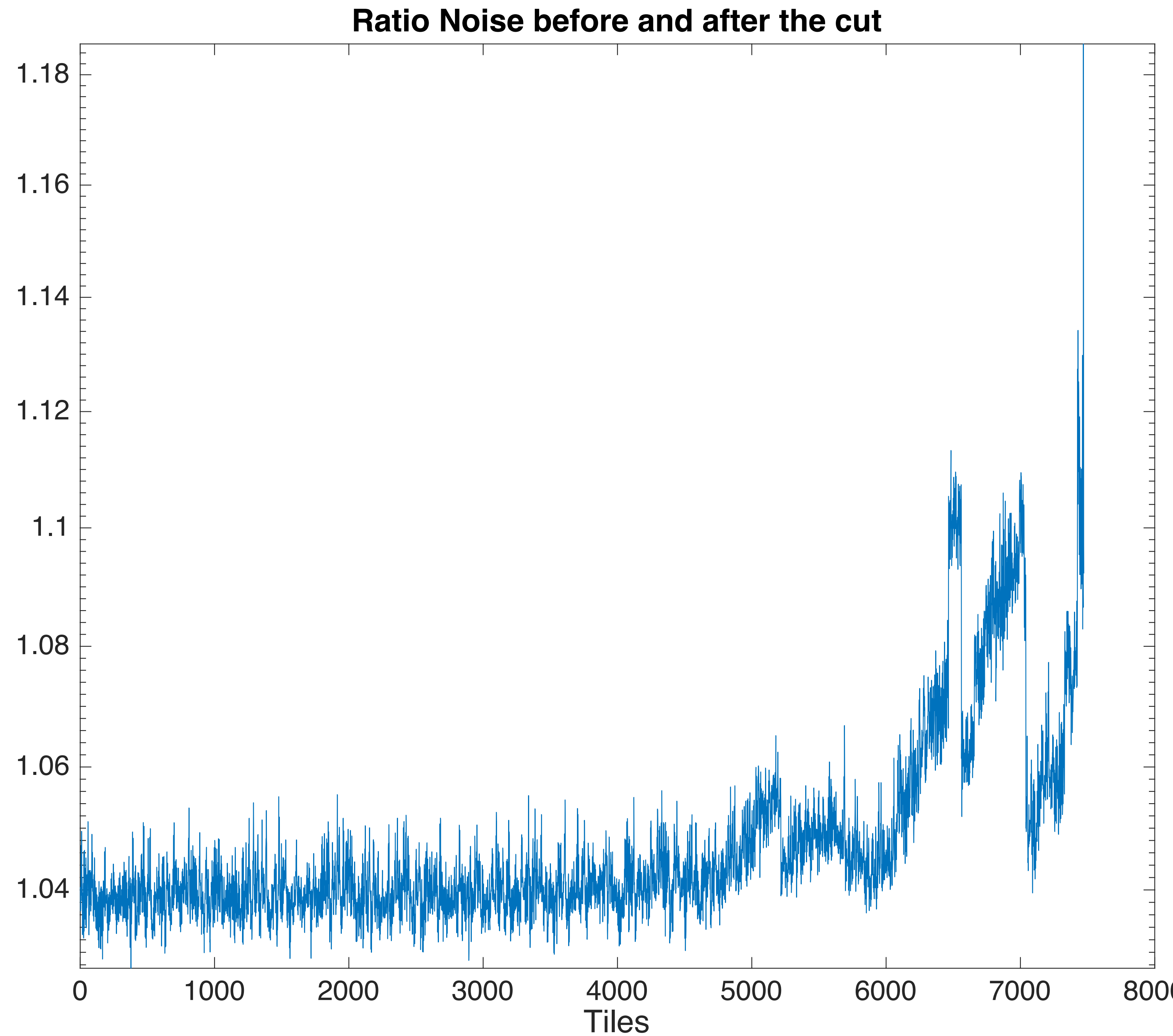
Calibrate on 4x4 tiles with 256 dipole stations

20,000+ pt source sky model

$B_{max}$ = 49 km

On EoR scales, increase in noise of ~4%. Larger on longest baselines

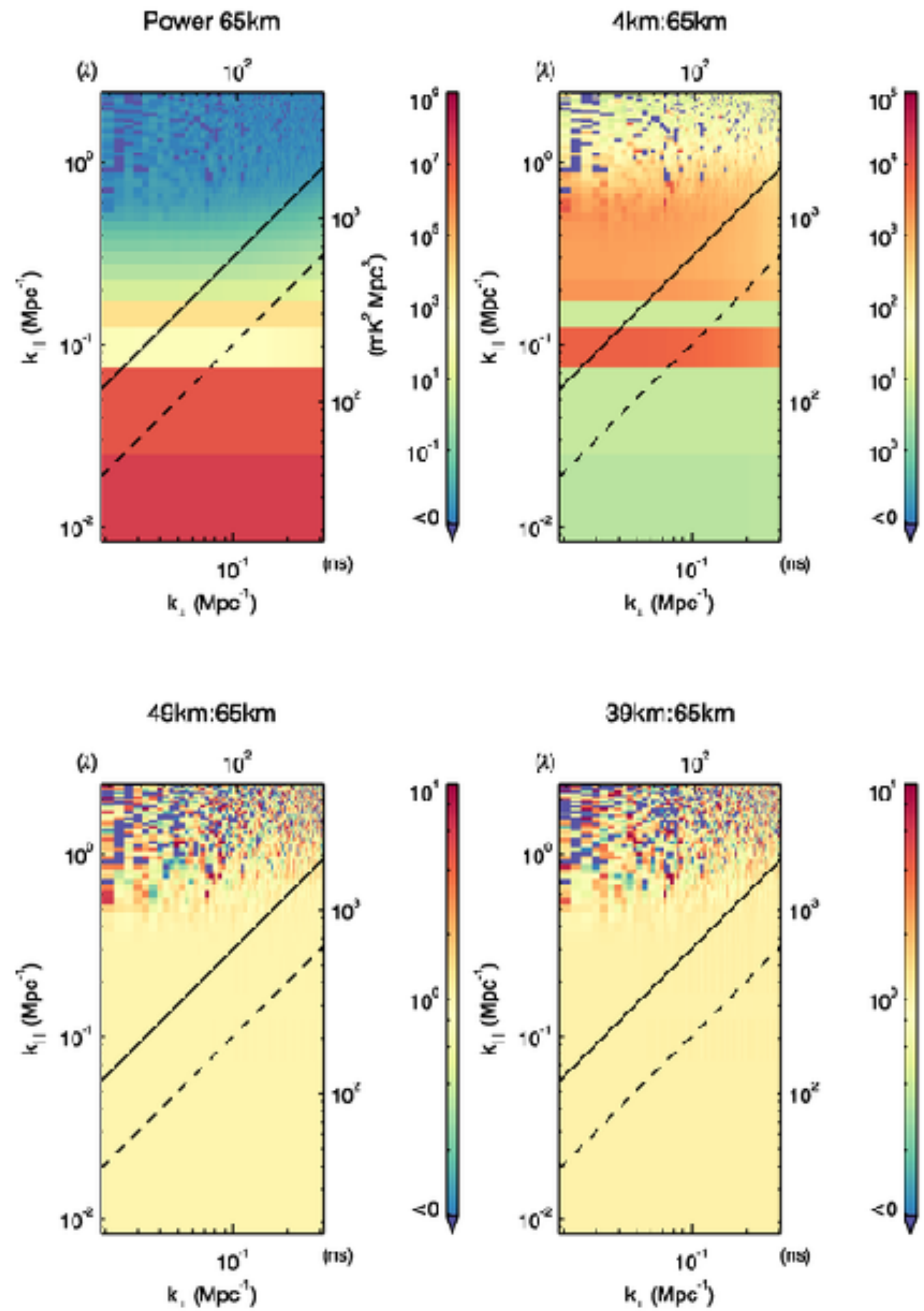
(ionosphere not included)




Mouri & Koopmans

# Impact of source structure on EoR

1) Model sources with structure
2) Build sky model with SKA
3) Propagate errors into EoR PS




Estimated impact for 1 Jy source modelled as multi-scale Gaussian with ~300 parameters
+ other simple sources

Increased errors generally acceptable for $B_{max} \gtrsim 39$ km

Trott & Wayth

# Reduced SDP HPC

| | | | | | |
|---|---|---|---|---|---|
| 8 | **SDP- HPC: Deploy 200 Pflops (rather than 260 Pflops)** It is expected that the cost of SDP processors will decrease in time due to more efficient and cheaper technology becoming available (Moore's Law gain). Therefore, re-sizing the first major purchase and assembling the full 260 Pflops SDP system in stages, rather than in a single deployment, results in a saving. If a smaller SDP is initially deployed, computationally demanding observations need to be observed for smaller fractions of time, in order for the sustained load on SDP to be compatible with the reduced system size. Those observations will still be possible, but will be accumulated more slowly. | **COMMON** | **Lower allowed duty cycle for HPC-intensive observations.** | 2 | ~80% |
| 8 | **SDP- HPC: Deploy 150 Pflops (from 200 Pflops)** It is expected that the cost of SDP processors will decrease in time due to more efficient and cheaper technology becoming available (Moore's Law gain). Therefore, re-sizing the first major purchase and assembling the full 260 Pflops SDP system in stages, rather than in a single deployment, results in a saving. If a smaller SDP is initially deployed, computationally demanding observations need to be observed for smaller fractions of time, in order for the sustained load on SDP to be compatible with the reduced system size. Those observations will still be possible, but will be accumulated more slowly. | **COMMON** | **Lower allowed duty cycle for HPC-intensive observations.** | 3 | ~60% |
| 8 | **SDP- HPC: Deploy 100 Pflops (from 150 Pflops)** It is expected that the cost of SDP processors will decrease in time due to more efficient and cheaper technology becoming available (Moore's Law gain). Therefore, re-sizing the first major purchase and assembling the full 260 Pflops SDP system in stages, rather than in a single deployment, results in a saving. If a smaller SDP is initially deployed, computationally demanding observations need to be observed for smaller fractions of time, in order for the sustained load on SDP to be compatible with the reduced system size. Those observations will still be possible, but will be accumulated more slowly. | **COMMON** | **Lower allowed duty cycle for HPC-intensive observations.** | 4 | ~40% |
| 8 | **SDP- HPC: Deploy 50 Pflops (from 100 Pflops)** It is expected that the cost of SDP processors will decrease in time due to more efficient and cheaper technology becoming available (Moore's Law gain). Therefore, re-sizing the first major purchase and assembling the full 260 Pflops SDP system in stages, rather than in a single deployment, results in a saving. If a smaller SDP is initially deployed, computationally demanding observations need to be observed for smaller fractions of time, in order for the sustained load on SDP to be compatible with the reduced system size. Those observations will still be possible, but will be accumulated more slowly. | **COMMON** | **Lower allowed duty cycle for HPC-intensive observations.** | 4 | ~20% |

- Suggestion in CC report that 80% (60%) cuts to HPC translate into 40%(30%) EoR observing efficiency. Assumptions unclear.
- Naively increases time needed to complete survey by 1.25 (1.7)?
- Critical path for EoR is likely DEEP survey - needs full 5 yrs to accumulate integration time on sky, so can't lose nights.
- Need more information to properly assess, but a concern

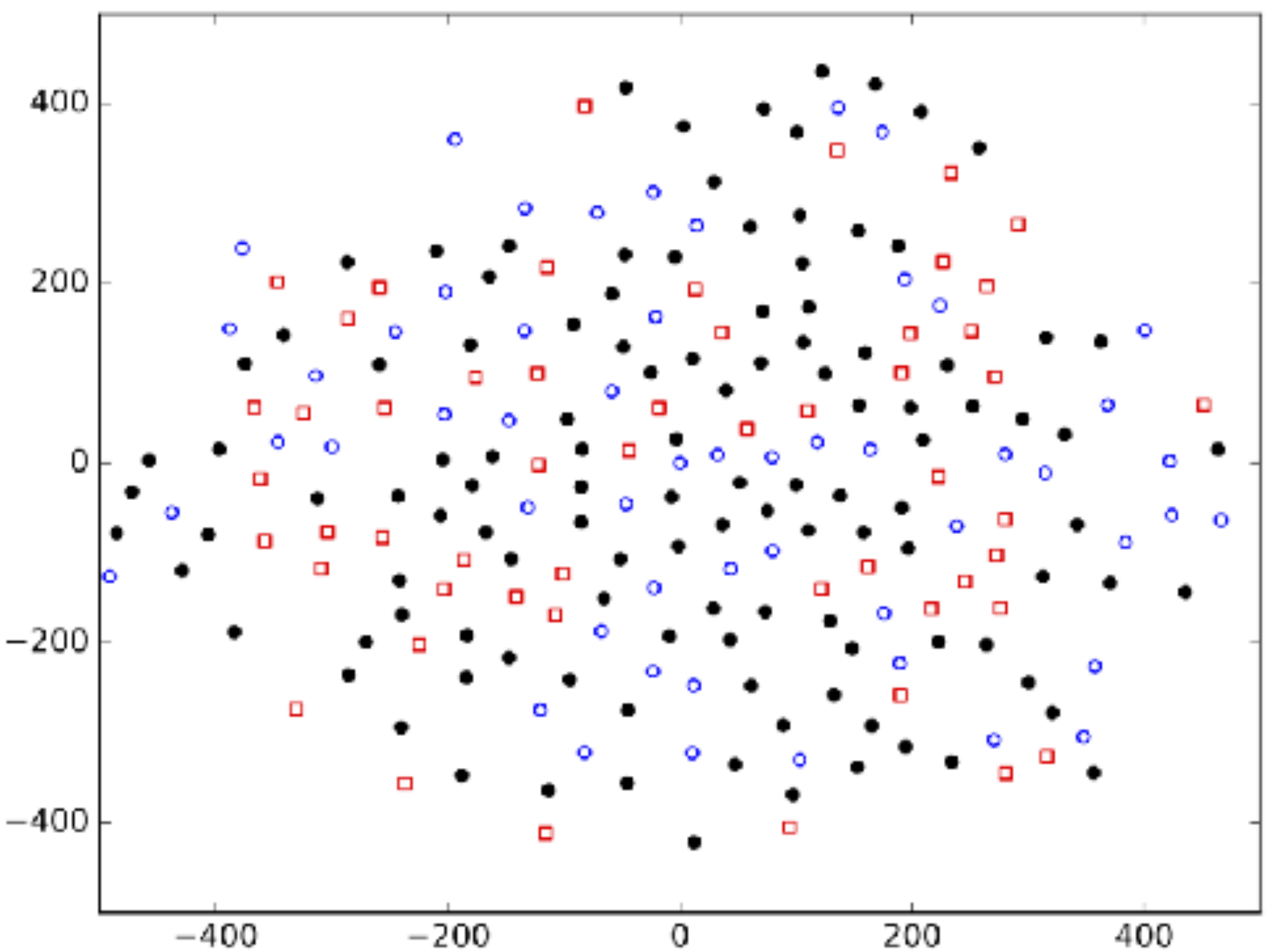
# Reduced core sensitivity

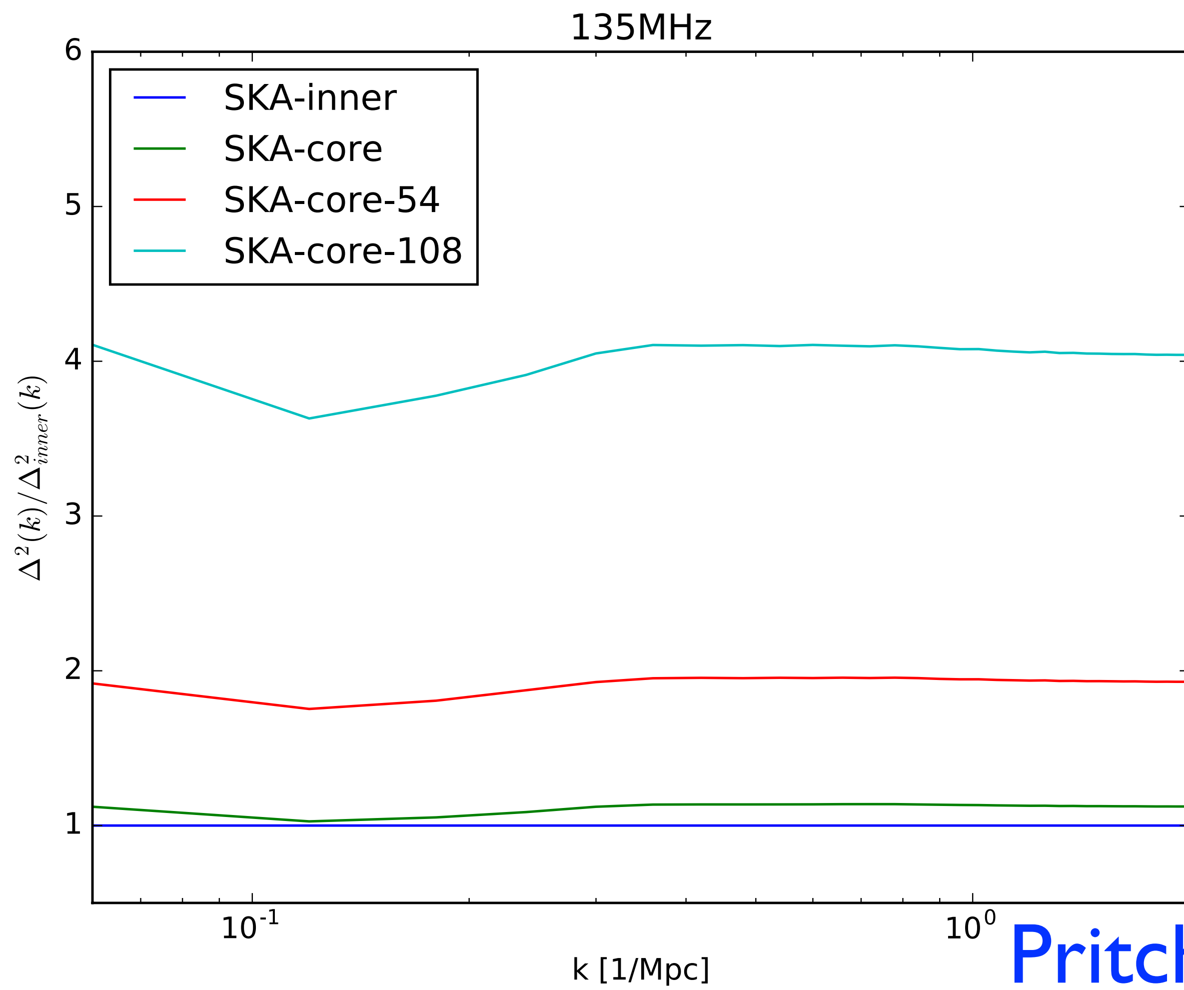
| | | | | |
|---|---|---|---|---|
| 5.30 / Deeper Savings | **Remove 54 LOW stations from core** In this measure, 54 stations are randomly removed from the inner 1.5 km radius region (where there are currently 282 stations). There is a 20% loss in core sensitivity, which would require about 40% increased integration time to compensate. | LOW | 20% Sensitivity loss in core | 4 |
| 5.30 / Deeper Savings | **Remove additional 54 LOW stations from core** In this measure, an additional 54 stations are randomly removed from the inner 1.5 km radius region (where there are currently 282 stations) for a total loss of 108 stations in the core, and 36 stations beyond 40km. There is a 40% loss in core sensitivity, which would require a doubling of integration time to compensate. | LOW | 40% Sensitivity loss in core | 4 |

Core is key for power spectrum sensitivity, which scales as $1/N_{station}^2$

Cutting 54 (108) from 225 core stations increase integration by x2 (x4)

Extremely damaging for EoR => Less sensitive than HERA




Pritchard

# Summary

Lots of work still needed to have total confidence

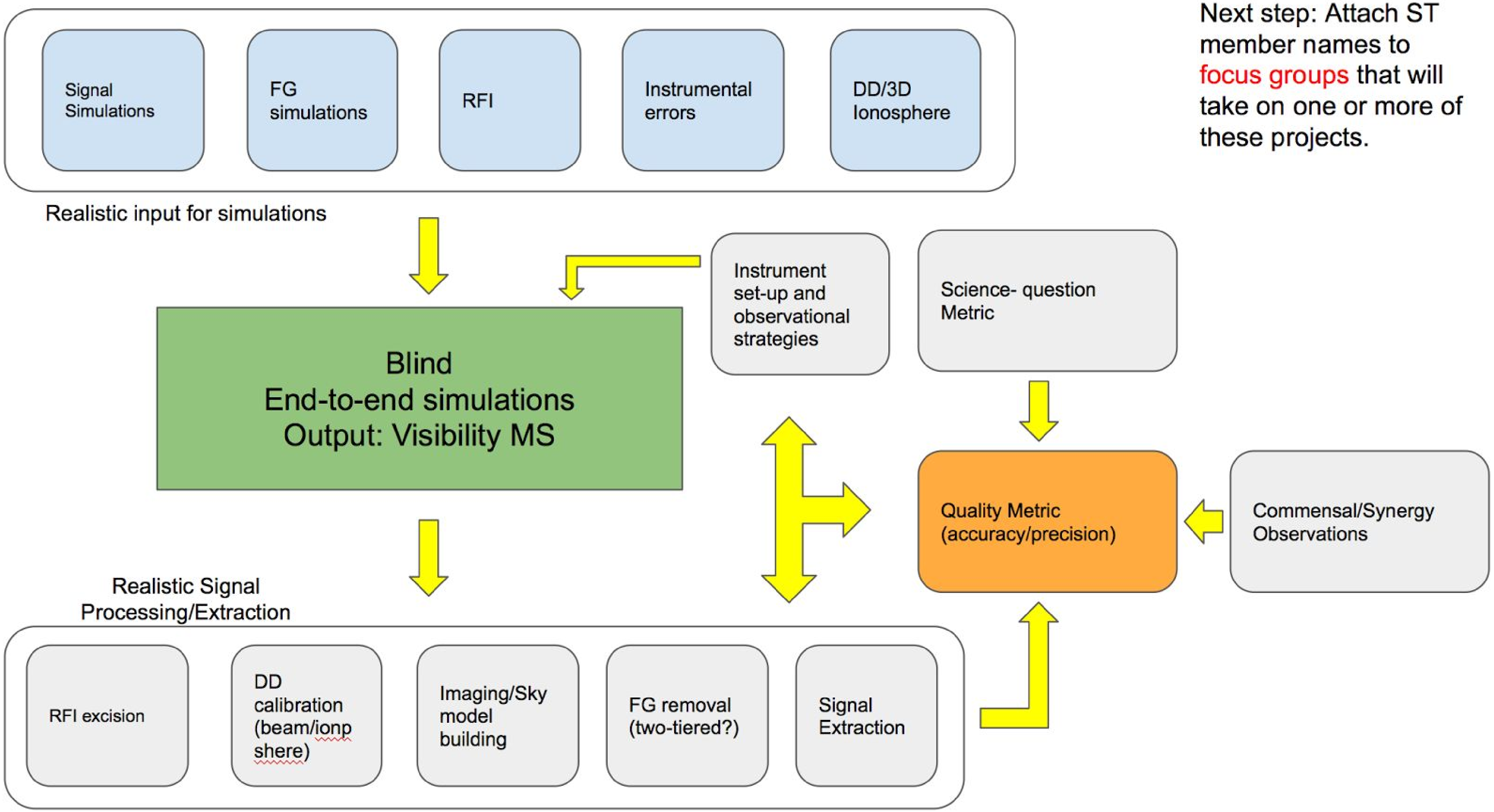
- Station level beam forming - Needs more thought.
- Reduced bandwidth - ~double integration time.
- Antennae redesign - Mostly ok for EoR. Impacts IM at $z<6$. Smoother bandpass preferred. Critical.
- 49km max baselines - Seems mostly ok. Sky model issues.
- Reduced SDP-HPC - Possible issues with data collection rate
- Reduced core sensitivity - Don't! x2 (x4) increases to integration time for EoR PS & possible calibration issues.

Proposed changes seem manageable without compromising EoR science. Loss of science due to increased integration time. Caution with cumulative impacts (x2-x3.5).

# Overview of Science Team Structure and Goals

# Path to a KSP Proposal

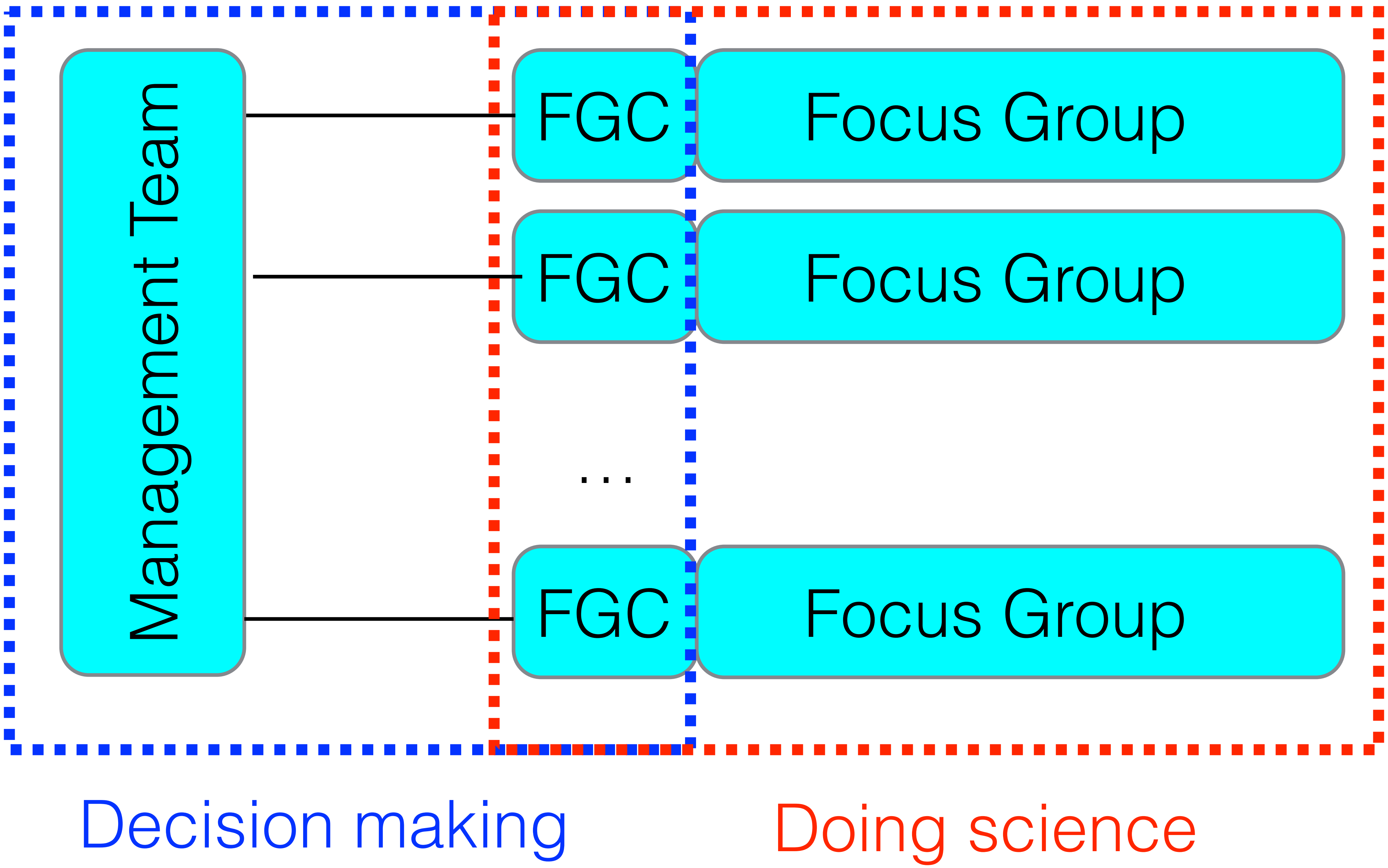
## Demonstrate we have expertise and tools for analysis of SKA data based on precursor experience




Next step: Attach ST member names to focus groups that will take on one or more of these projects.

Figure 1: Flow-diagram for End-to-End data simulations.

# Science Team Structure




Membership levels: core, associate, consulting

Imagine growing over time to become large KSP

# Focus Groups & Coordinators

A) Theory/Numerical Simulations:

1. Theory/Physics for understanding model space/subgrid physics - Rennan Barkana, Benoit Semelin
2. Full numerical simulations for calibration - Ilian Iliev, Garrelt Mellema
3. Fast simulations for analysis - Andrei Mesinger, Tirth Roy Choudhury
4. Foreground Simulations - **Mario Santos,** Vibor Jelic

B) Observational Strategies

1. Interferometric: Cath Trott, Vibor Jelic
2. Global Signal: Uday Shankar, Abhi Datta
3. 21cm Forest - Bene Ciardi, Yidong Xu

C) Data Processing

1. RFI Excision - Andre Offringa, Ben McKinley
2. Calibration/Ionosphere - Daniel Mitchell, Sarod Yatawatta
3. Imaging/Sky-model building - Bart Pindor, Andre Offringa
4. Foreground Fitting/Removal - Emma Chapman**, Anna Bonaldi**
5. New Algorithmic Development - **Kris Zarb-Adami,** Leon Koopmans
6. Computational and Other Resources - Mike Jones, Pandey V.

D) Signal Extraction and Error Analysis - Cath Trott, Saleem Zaroubi

E) Signal Analysis and Interpretation - **Jonathan Pritchard,** Andrei Mesinger

F) Synergy (SKA + Other instruments) - Erik Zackrisson, Pratika Dayal, Andrea Ferrara

G) End-to-End (Data) Simulations - **Leon Koopmans, Gianni Bernardi**

# Management Team

**Leon Koopmans** (NL, Chair ST), Gianni Bernardi (SA),
**Garrelt Mellema** (SE), Andrei Mesinger (IT)
**Jonathan Pritchard** (UK, Chair SWG), Cath Trott (AU)
Abhi Datta (IN)

## Mix of SKA member representation and expertise

- The Board contains one representative from each SKA member-state which is substantially represented in the EoR/CD Science Team. This number can increase, per decision by the Board, in the future depending on scientific needs and commitments.
- Board members can also be added by (a) new SKA member-states joining, (b) ST members bringing in substantial resources to the benefit of the ST, (c) balancing expertise.
- The founding-Board term finishes after the outcome of the KSP bid is made public. Thereafter, terms are 1-3 years[1], renewable and Board members per member-state are decided by their respective ST members.
- The chair/vice-chair positions are elected by the Board members. Terms are 1-3 years[2] with successors selected from the existing Board members (process TBD by the Board; e.g. election per member-state, rotation, etc.)
- Board members should be those with senior/faculty positions to ensure long-term stability.

# Progress towards forming a Science Team

Stockholm Aug 2015 - SKA KSP Meeting
Groningen Oct 2015 - ST Kick off meeting
Goa Nov 2016 - SKA Science Meeting
Pisa Mar 2017 - ST Science Meeting
Zagreb Sept 2017 - next ST Science Meeting

Current discussions focussed on:

- mock pipeline data challenges
- deciding key priorities for different FGs
- SKA design

https://sites.google.com/site/skacdeorscienceteam/home

# Summary

Lots of work still needed to have total confidence

- Station level beam forming - Needs more thought.
- Reduced bandwidth - ~double integration time.
- Antennae redesign - Mostly ok for EoR. Impacts IM at z<6. Smoother bandpass preferred. Critical.
- 49km max baselines - Seems mostly ok. Sky model issues.
- Reduced SDP-HPC - Possible issues with data collection rate
- Reduced core sensitivity - Don't! x2 (x4) increases to integration time for EoR PS & possible calibration issues.

Proposed changes seem manageable without compromising EoR science. Loss of science due to increased integration time. Caution with cumulative impacts (x2-x3.5).

# SKA-LOW

http://astronomers.skatelescope.org/documents/