Forward-Modelling Beam Error in 21cm Parameter Estimation Experiments

Bella A. Nasirudin
Steven Murray
David Prelogovic
Andrei Mesinger
Gianni Bernardi

SCUOLA NORMALE SUPERIORE

European Research Council

Introduction

Detection of 21cm signal from the Epoch of Reionization (EoR) is one of the key science goals of current and upcoming radio interferometers

Main **challenges**:

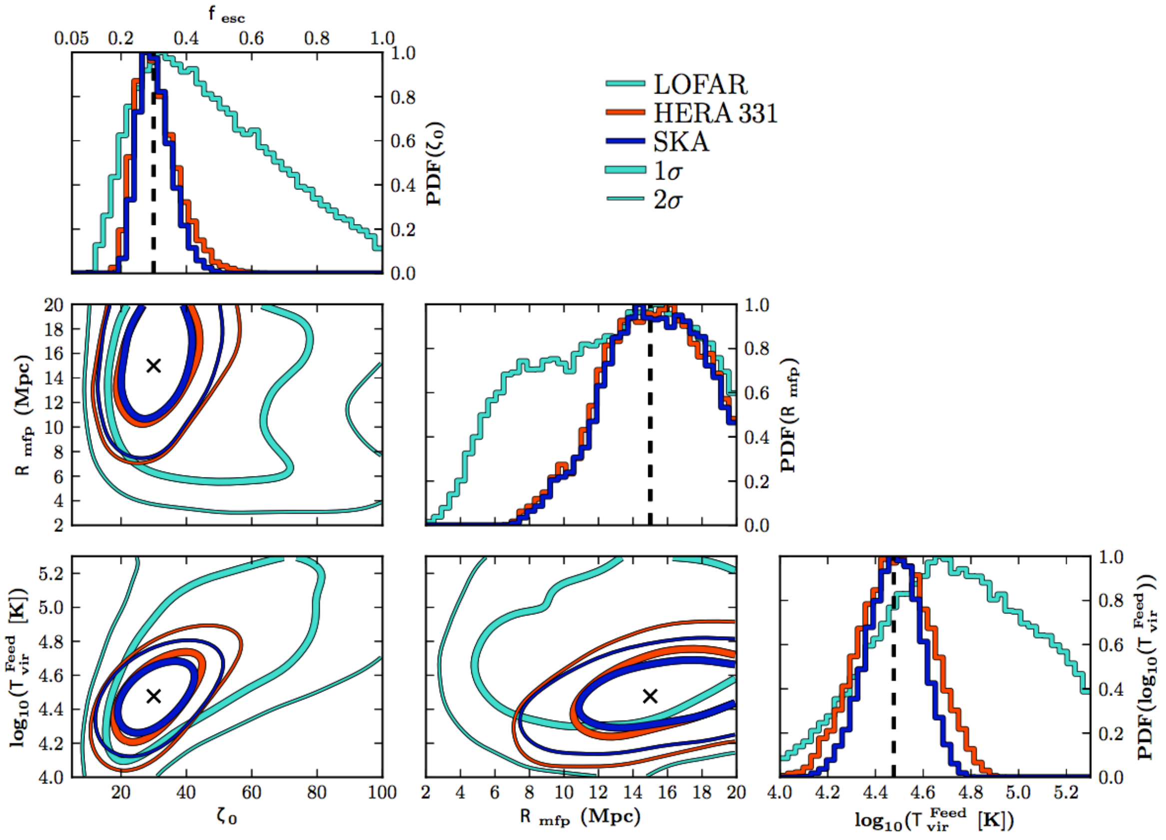
- **Foreground** that are ~ 5 orders brighter than signal
- **Instrumental** systematics e.g. frequency-dependent beam, widefield effects, cable reflections
- Other systematics: RFI, ionosphere, etc

Introduction

Ultimate goal is to understand EoR by **constraining** reionization parameters: 21cmFAST + 21cmMC

In traditional 21cm parameter estimation experiments:

- Astrophysics + cosmology are **forward** modeled
- Instrument + foregrounds are **backward** modeled



Greig & Mesinger (2015)

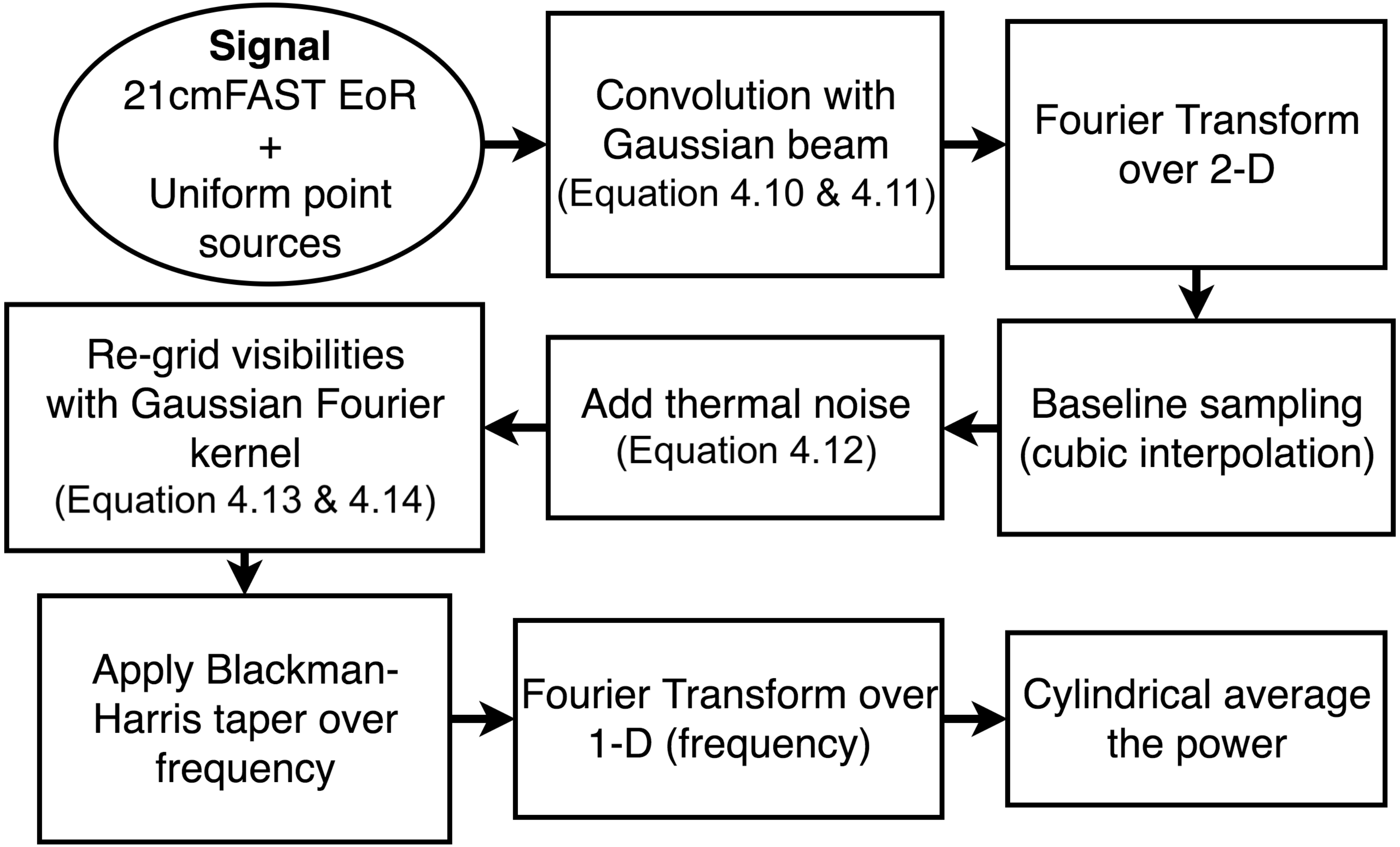
Introduction

Previous work on propagating, i.e. forward modelling, **extra-galactic foregrounds** and **instrumental effects** into 21cmMC in Nasirudin et. al (2020)

Publicly-available, plug-in framework, py21cmmc_fg, which include:

- Extra-galactic (residual) foregrounds (10mJy)
- Baseline sampling (MWA Phase II)
- MWA-based Gaussian beam
- 1000 hour of SKA-like noise

Introduction



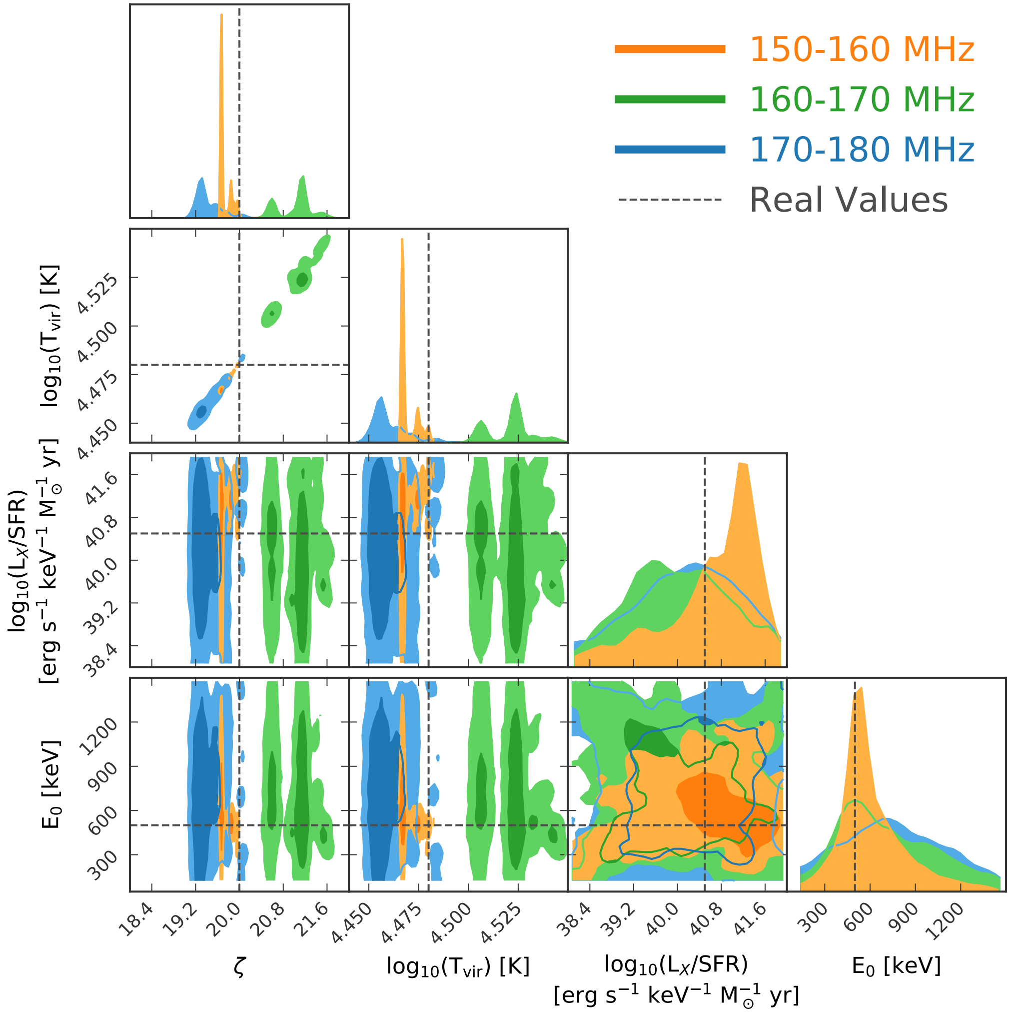
Nasirudin et. al (2020)

Introduction

The mock data is fully forward-modelled but the Bayesian framework is not because **foreground power** is added to the **21cm power**

Results show that **unaccounted** cross-power arising from not properly forward modelling can cause bias in the constrain

Gazagnes et. al (2021) verified bias results due to foreground residual



Nasirudin et. al (2020)

Goal

Develop a flexible forward modelling framework by updating py21cmmc_fg:

Best estimates + uncertainty

- Astrophysics + cosmology
- Galactic + extragalactic foregrounds
- **Realistic Beam**
- Instrumental effects

Propagate forward

Steps

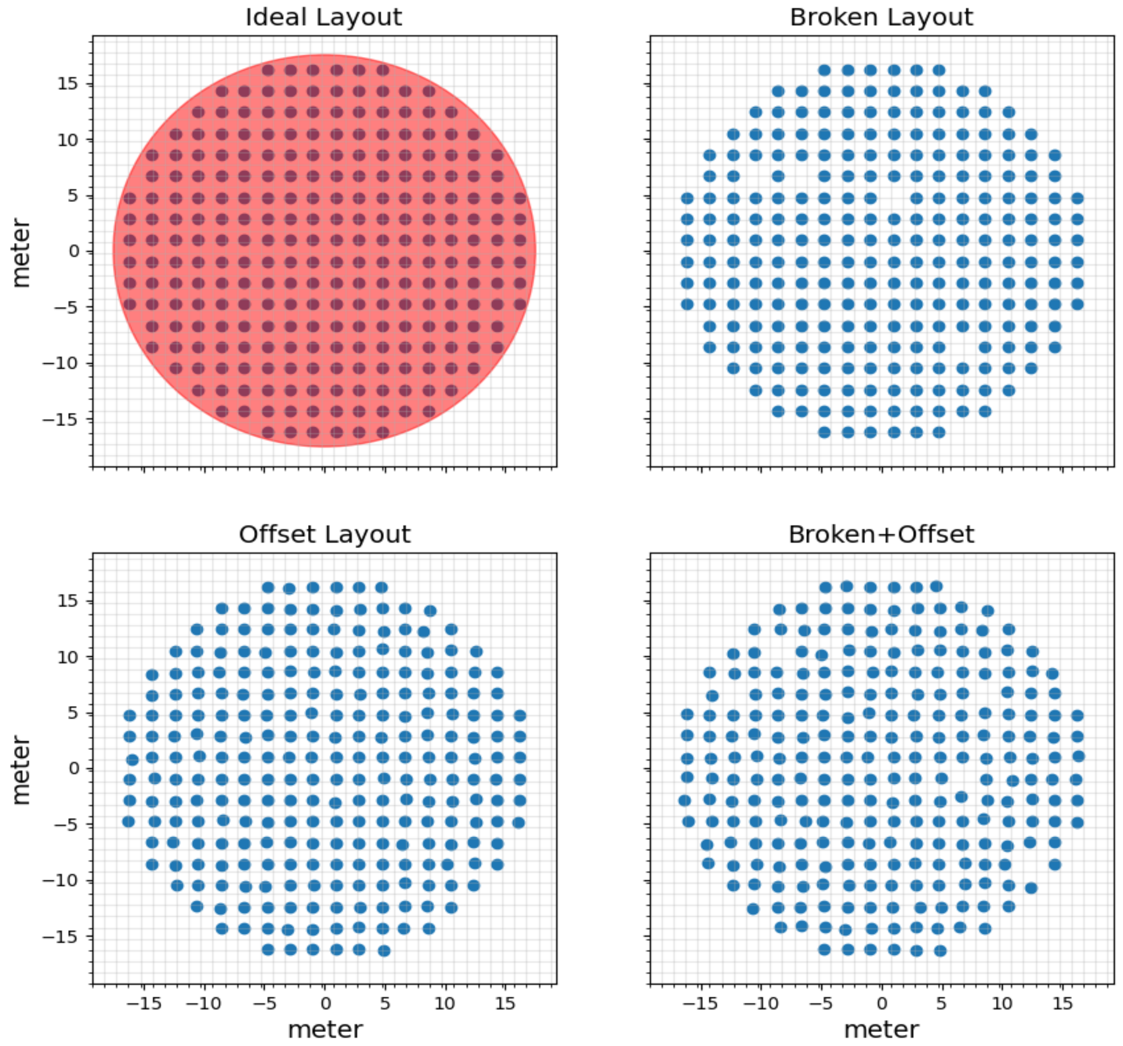
1. Make a **database** of beam realizations and empirically sample errors
2. Choose a perturbation **basis set** acting on our “best guess” for the beam
3. **Fit basis set** in 2 to the empirical database created in step 1
4. Use the distribution of best fit values (conservatively widened) as a **prior**
5. **Covary** astro + cosmo + beam parameters following the prior distribution in 4

Step 1

Make a database of beam realizations and empirically sample errors

Antenna Layouts

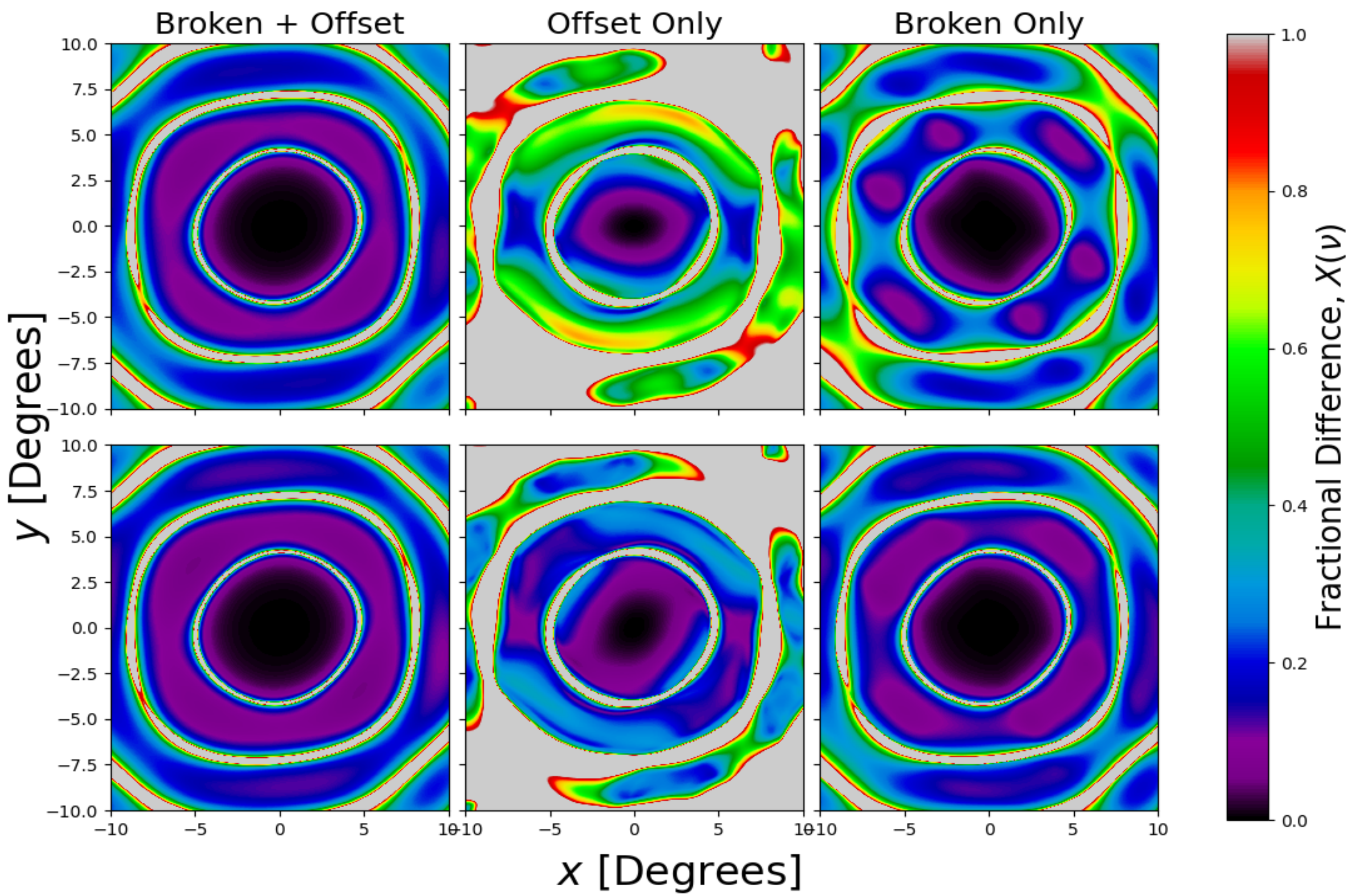
- Regularly-spaced antennas in 35-m diameter as ideal
- Broken (i.e. offline) antennas based on Joseph et. al (2020) (5%)
- Offset antennas (Normal distribution $\sigma \sim 10$ cm)
- Broken and offset antennas



Nasirudin et. al (in prep.)

Beam Realizations

- Use OSKAR: Oxford SKA Radio Telescope simulator
- Stokes I polarization
- 150, 170 and 190 MHz



Nasirudin et. al (in prep.)

Step 2

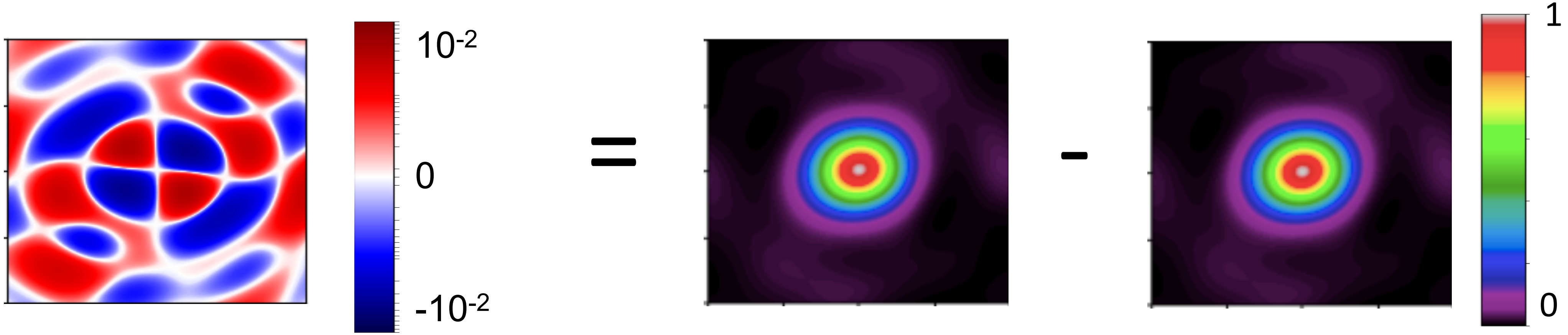
Choose a perturbation basis set acting on our best guess for the beam

Kernel Principal Component Analysis (KPCA)

Reduce the dimensionality of dataset via non-linear transformations to a higher dimension

Use SPax, a GPU and CPU-optimized PCA and KPCA code developed by David Prelogovic available at https://github.com/dprelogo/SPax

Only model the residual between perturbed and ideal beam, $\Delta B = B_{true} - B_{ideal}$



Train KPCA on 7,000 realizations of **broken + offset** data and reconstruct on remaining 3,000 realizations

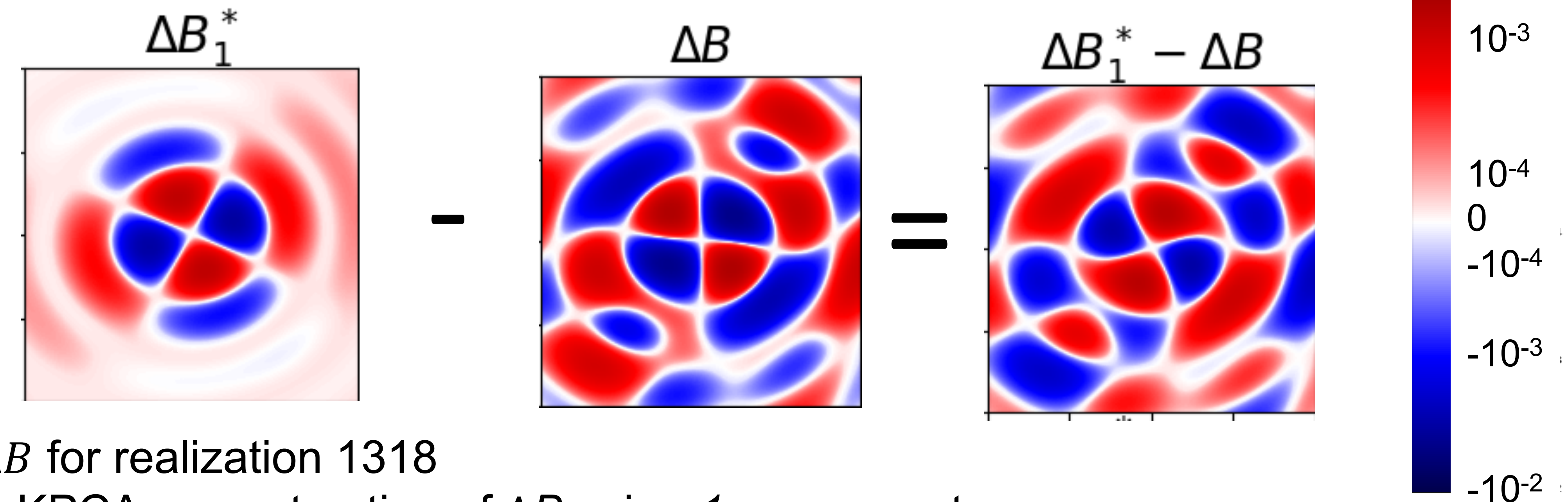
Step 3

Fit basis set in Step 2 to the empirical database created in Step 1

Results of KPCA: Number of Components

Consider one example realization (1318) with 7 offline dipoles and 23 offset dipoles

Let's look at only 1 component



ΔB : ΔB for realization 1318

ΔB^*_1 : KPCA reconstruction of ΔB using *1* components

$\Delta B^*_1 - \Delta B$: the difference between the KPCA reconstruction using 1 components and ΔB

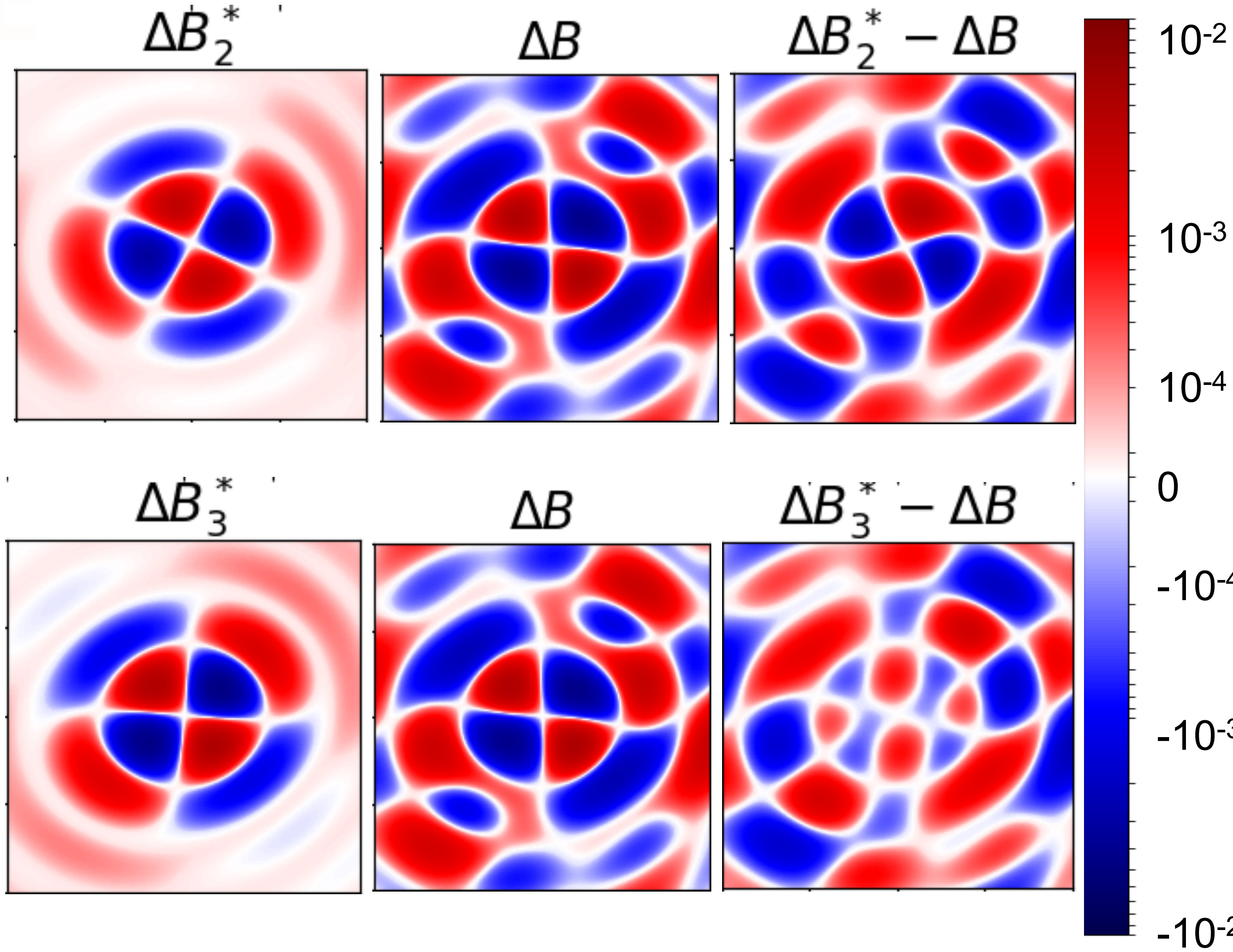
Results of KPCA : Number of Components

Now let's look at 2 more principal components

ΔB for realization 1318

ΔB^*_N : KPCA reconstruction of ΔB using N components

$\Delta B^*_N - \Delta B$: the difference between the KPCA reconstruction using N components and ΔB



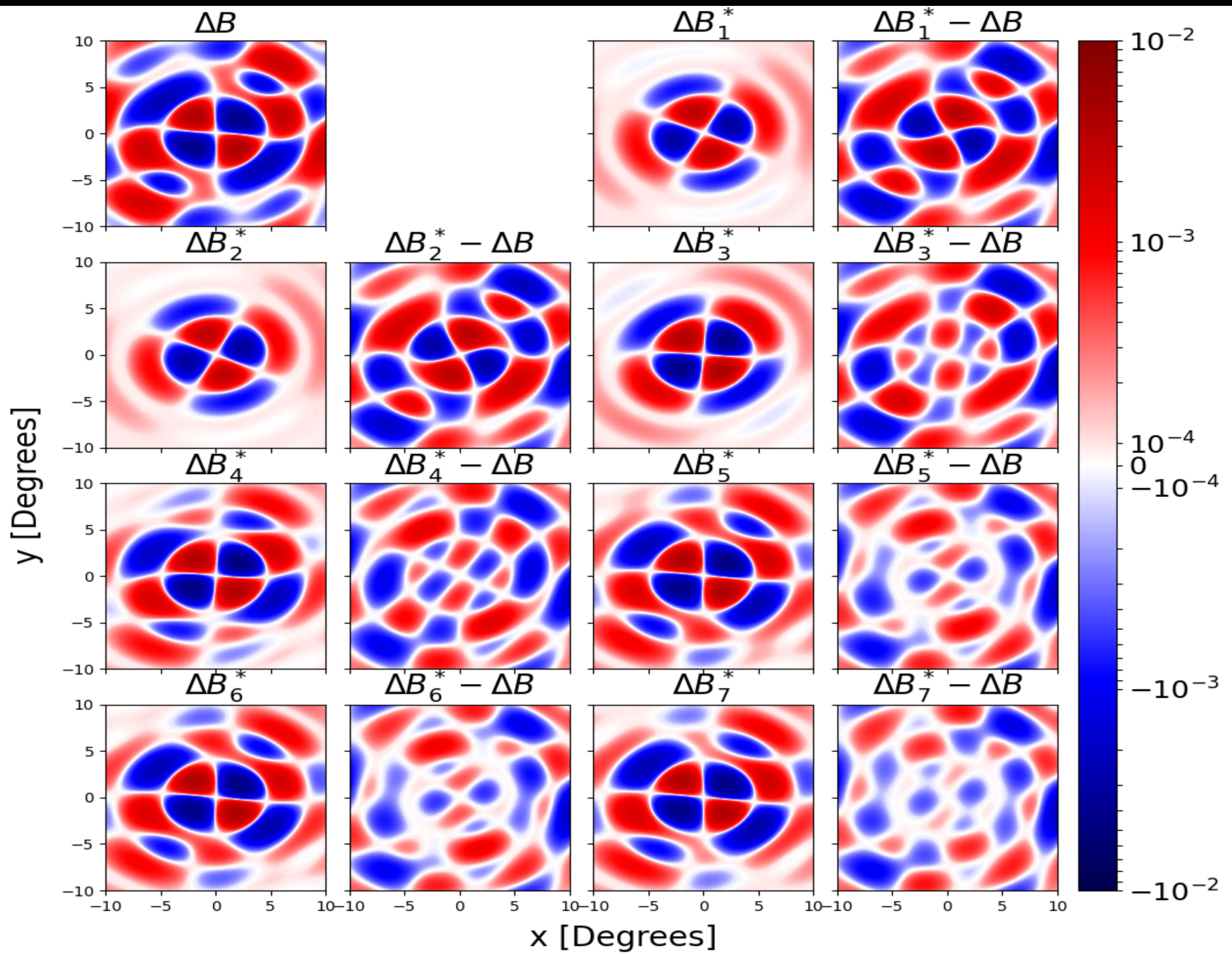
Results of KPCA : Number of Components

Now let's look at 7 principal components

ΔB for realization 1318 **(left-uppermost panel)**

ΔB^*_N : KPCA reconstruction of ΔB using N components

$\Delta B^*_N - \Delta B$: the difference between the KPCA reconstruction using N components and ΔB

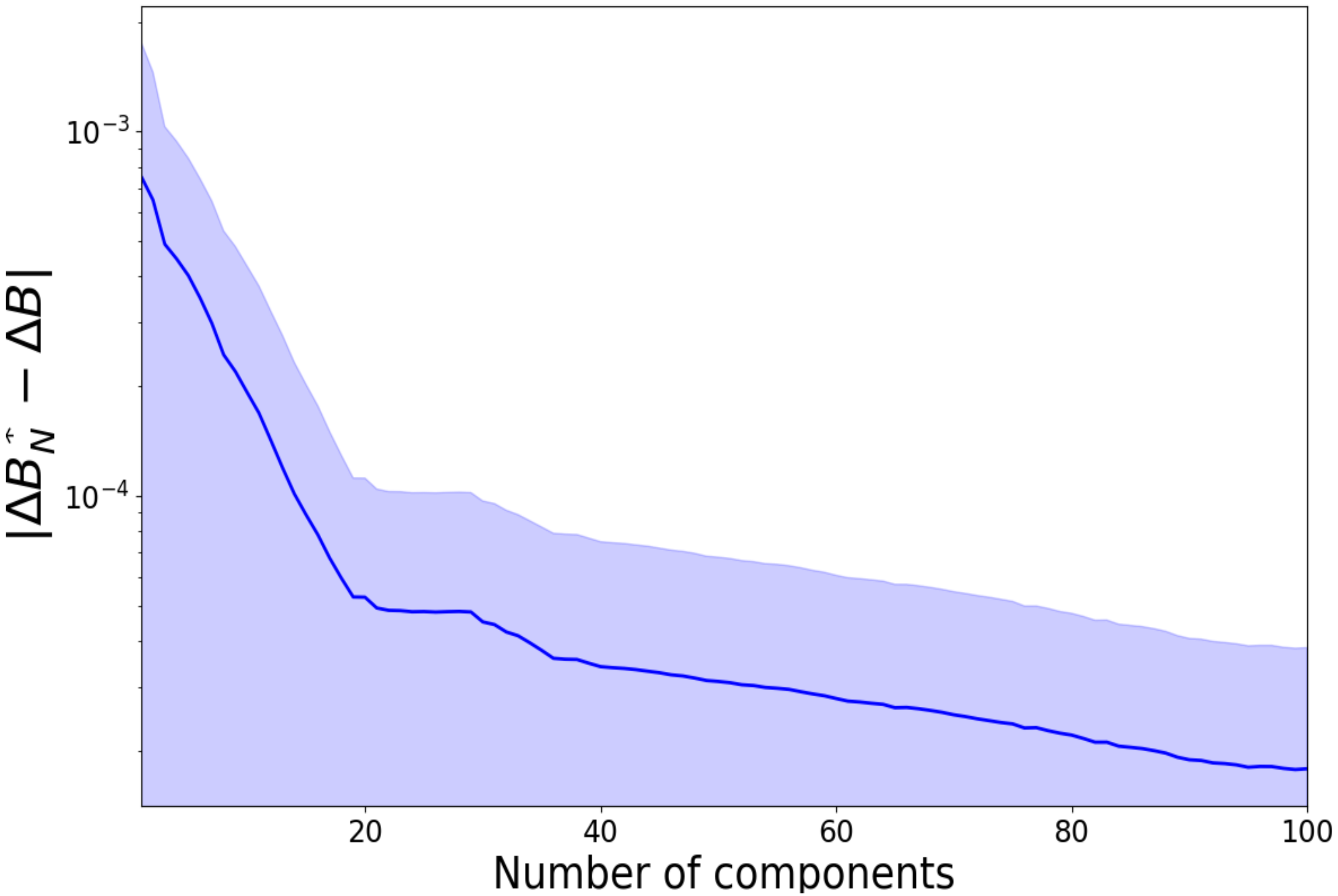


Nasirudin et. al (in prep.)

Results of KPCA : Number of Components

How many components to include?

We look at the **mean square error** between the reconstructed beam and true error to decide

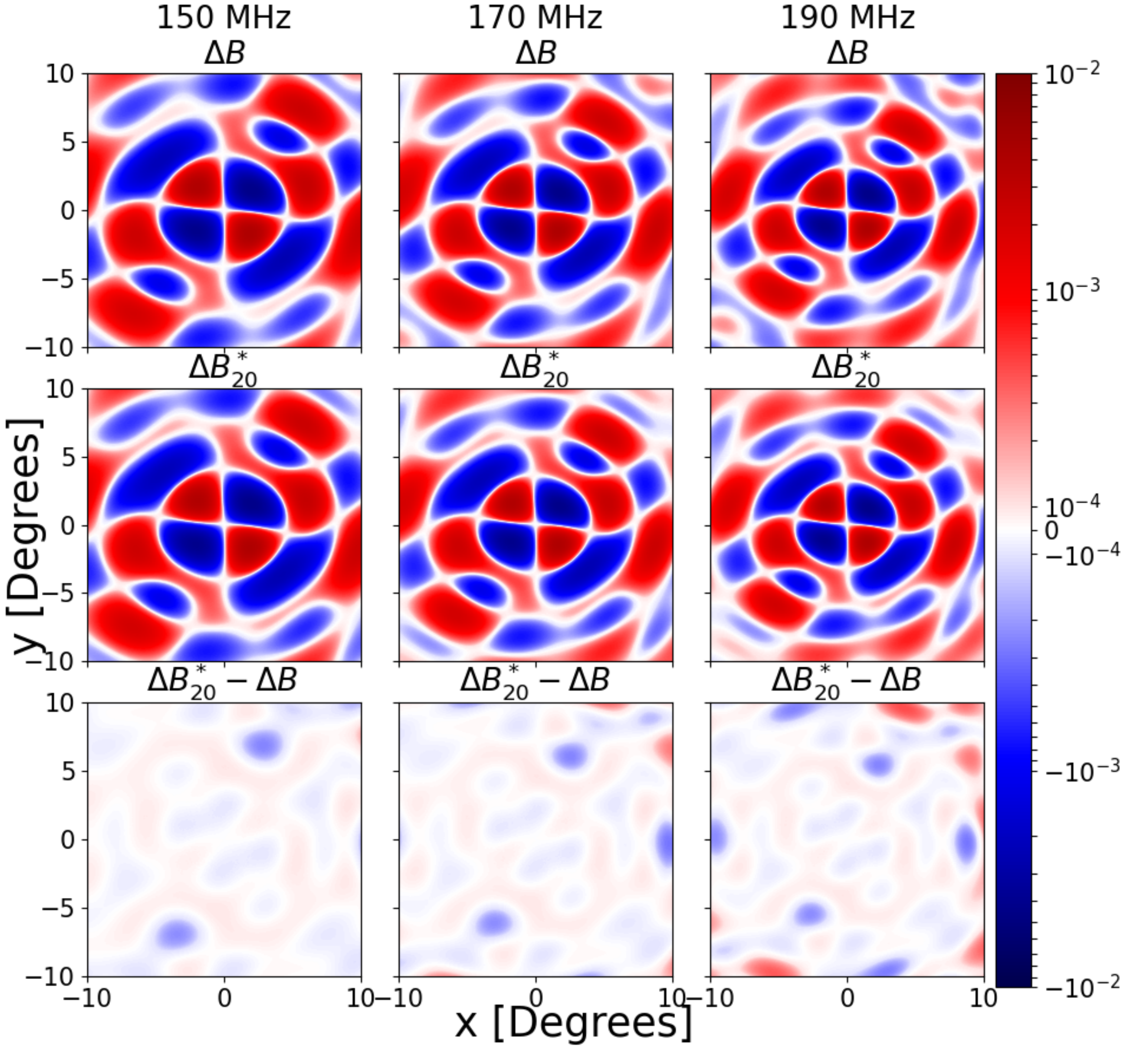


Nasirudin et. al (in prep.)

Results of KPCA : 20 Components

This is the final error reconstruction for the same realization using 20 components for all three frequencies

KPCA can characterize the beam error quite well



Nasirudin et. al (in prep.)

Step 4

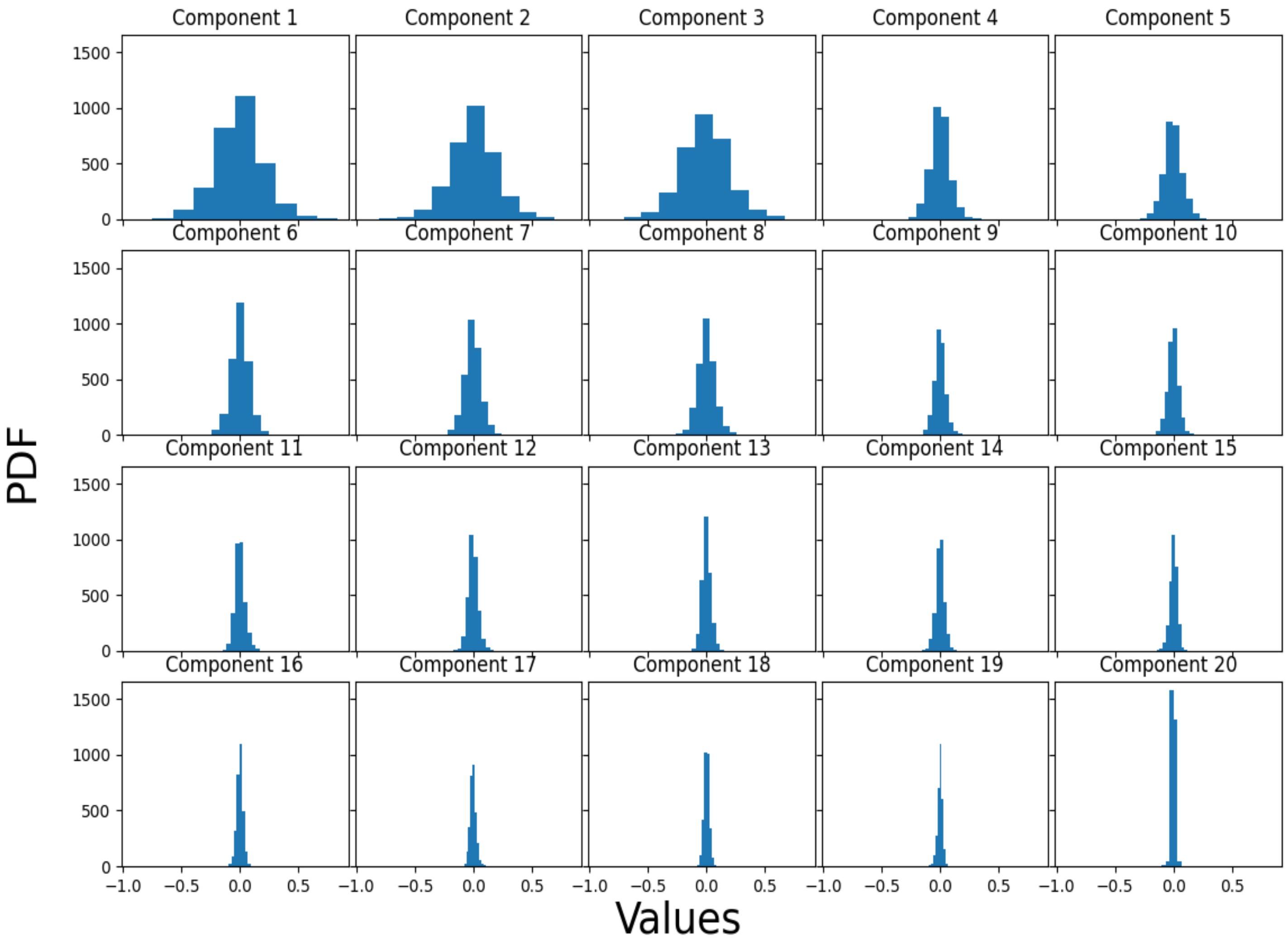
Use the distribution of best fit values as a prior

Distribution of Components

Include 20 principal components based on previous plot

PDFs from 3,000 realizations follow the Gaussian distribution

Need to conservatively widen the prior

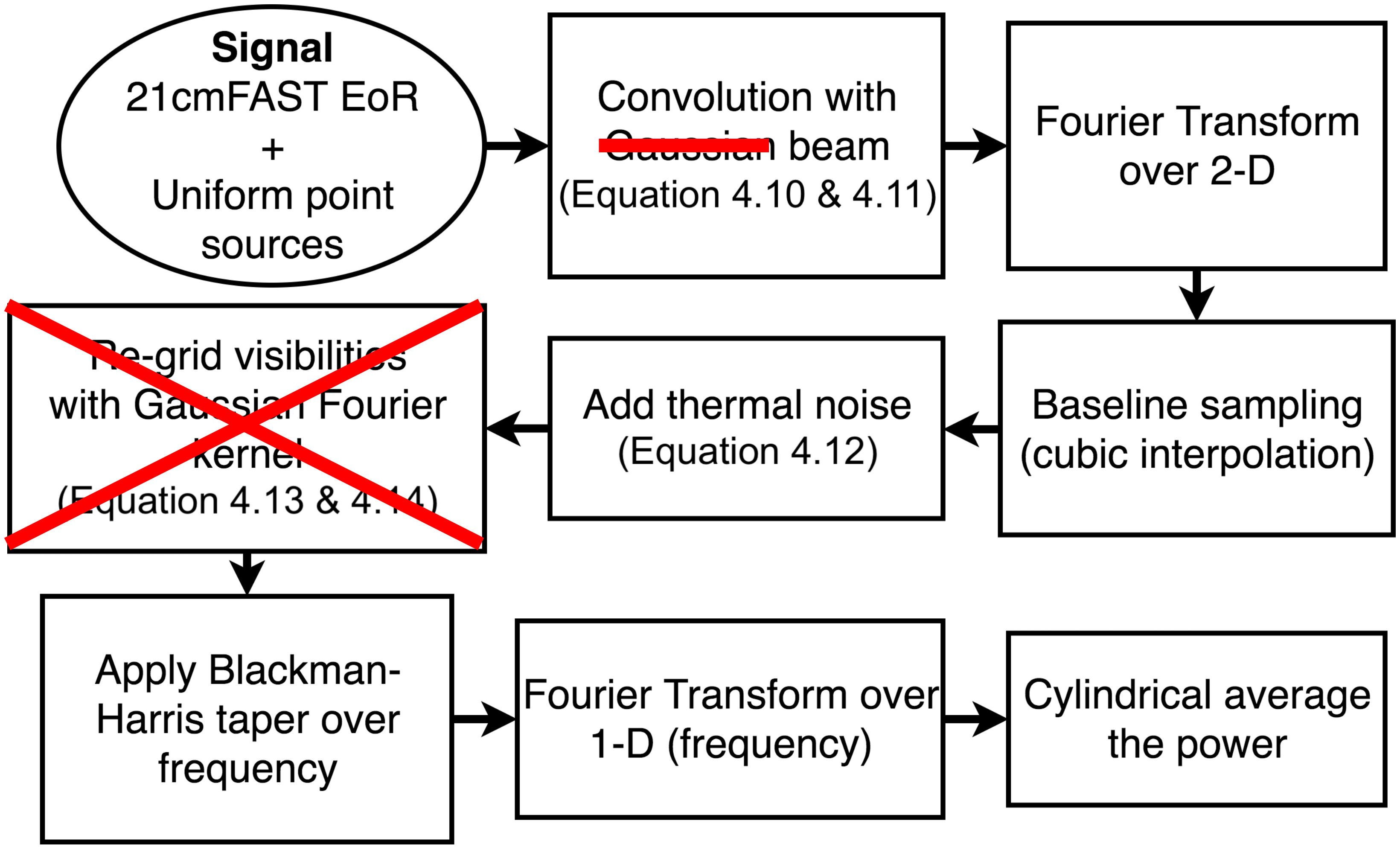


Nasirudin et. al (in prep.)

Step 5

Covary astrophysical, cosmological, and beam parameters following the prior distribution

Interferometric Framework



Nasirudin et. al (2020)

Next Steps

Investigate if there is an impact on the **parameter constrain**

- Incorporate into 21cmMC/py21cmmc_fg framework

- Compare posteriors with and without forward-modelled beam errors

- Possibly use GPU to speed things up

Conclusion

Develop a flexible forward modelling EoR parameter estimation framework that will include:

- Astrophysics and cosmology
- Galactic and extragalactic foregrounds
- Realistic beam perturbation
- Instrumental and other systematics

The framework will be **modular**: users can provide best guess and set of priors

KPCA reconstruction of large database of beam simulations provides an empirical basis for beam errors