Distributed Model Construction in Radio Interferometric Calibration

Sarod Yatawatta

ASTRON

The Netherlands Institute for Radio Astronomy,

Dwingeloo, The Netherlands

in collaboration with

Netherlands eScience Center,

Science Park, Amsterdam, The Netherlands

This work was supported by Netherlands eScience Center

(project DIRAC, grant 27016G05).

Introduction

- Calibration of radio telescopes: essential for correcting systematic errors (beam,ionosphere), removal of strong contaminating signals (foregrounds): for high quality imaging.
- Terabytes of data observed, data split into thousands of frequency channels, also stored at different locations in a network.
- Calibration solutions contain information about systematic errors.
- How do we build complete models for systematic errors in the data using calibration solutions?

LBA dipole

HBA dipole

LBA: low band (10-80 MHz), HBA: high band (100-240 MHz)

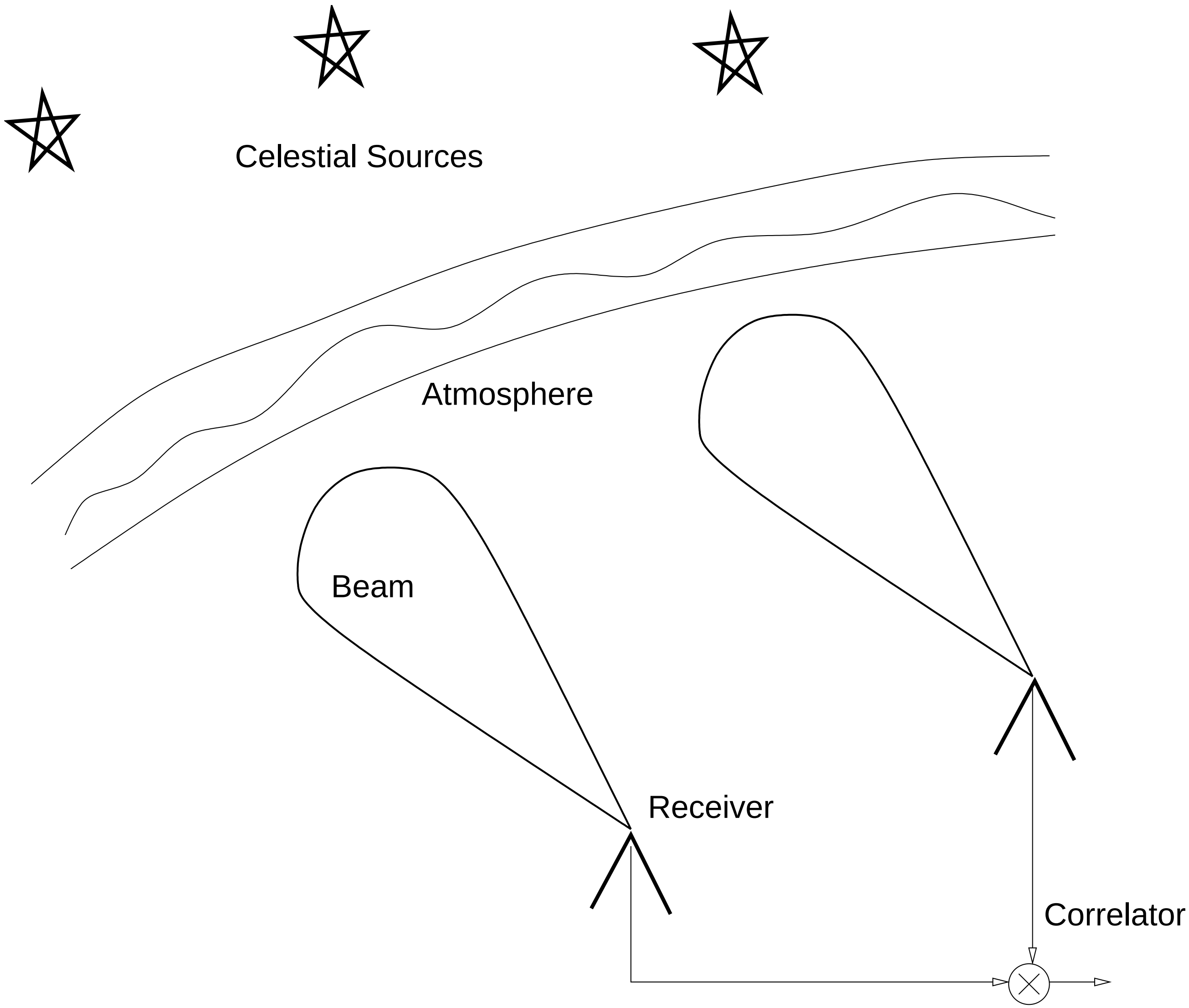
Antenna array

Phased array built using many dipoles

Modern radio telescopes

LOFAR core in the Netherlands

Radio interferometers



Two Receivers = Interferometer

Radio interferometry

We observe the Fourier transform of the sky. Major steps in radio astronomy:

- ☐ Correlation, Interference mitigation.
- ☐ Calibration:
  - ■ Estimate the systematic errors in the data and correct for them.
  - ■ Remove strong foreground sources to reveal weaker signals.
- ☐ Imaging and deconvolution:
  - ■ Convert observed Fourier space data into real space images.
  - ■ Remove errors due to incomplete sampling (deconvolution).
- ☐ Finally ... Science.

Calibration

$$\mathsf{V}_{pqf} = \mathsf{J}_{pf}\mathsf{C}_{pqf}\mathsf{J}_{qf}^H + \mathsf{N}_{pqf}$$

Observed data V_{pqf} at baseline p-q at frequency f, corrupted by systematic errors J_{pf} and J_{qf}.
Cost function

$$g_f(\mathsf{J}_f) = \sum_{p,q} \|\mathsf{V}_{pqf} - \mathsf{A}_p\mathsf{J}_f\mathsf{C}_{pqf}(\mathsf{A}_q\mathsf{J}_f)^H\|^2$$

where C_{pqf} is scalar, diagonal and

$$\mathsf{J}_f \stackrel{\triangle}{=} [\mathsf{J}_{1f}^T, \mathsf{J}_{2f}^T, \ldots, \mathsf{J}_{Nf}^T]^T, \quad \mathsf{A}_p \stackrel{\triangle}{=} [\mathbf{0}, \mathbf{0}, \ldots, \mathsf{I}, \ldots, \mathbf{0}]$$

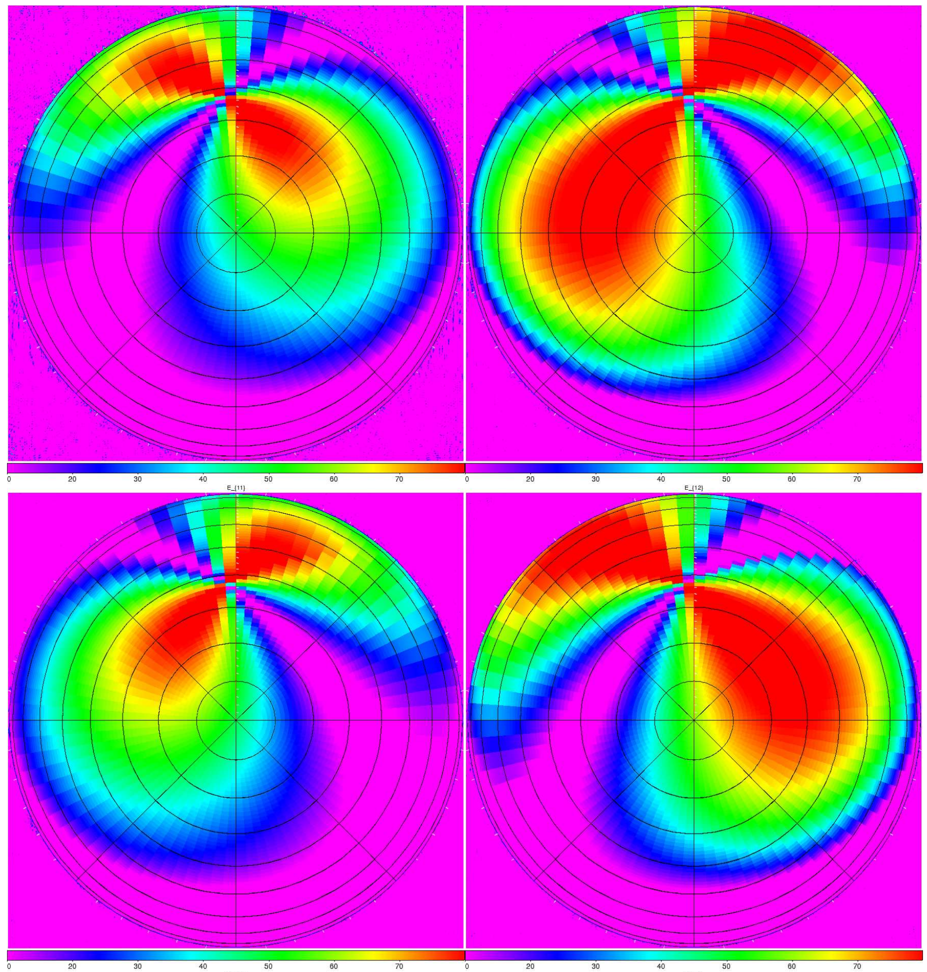
Calibration: minimizing $g_f(\mathsf{J}_f)$ to find J_f. Information about beam shape and ionosphere is hidden in J_f.

This work: building model X from calibration solutions J_f.

$$\mathsf{J}_f = \mathsf{X}\mathbf{\Phi}_{\alpha\beta f}$$

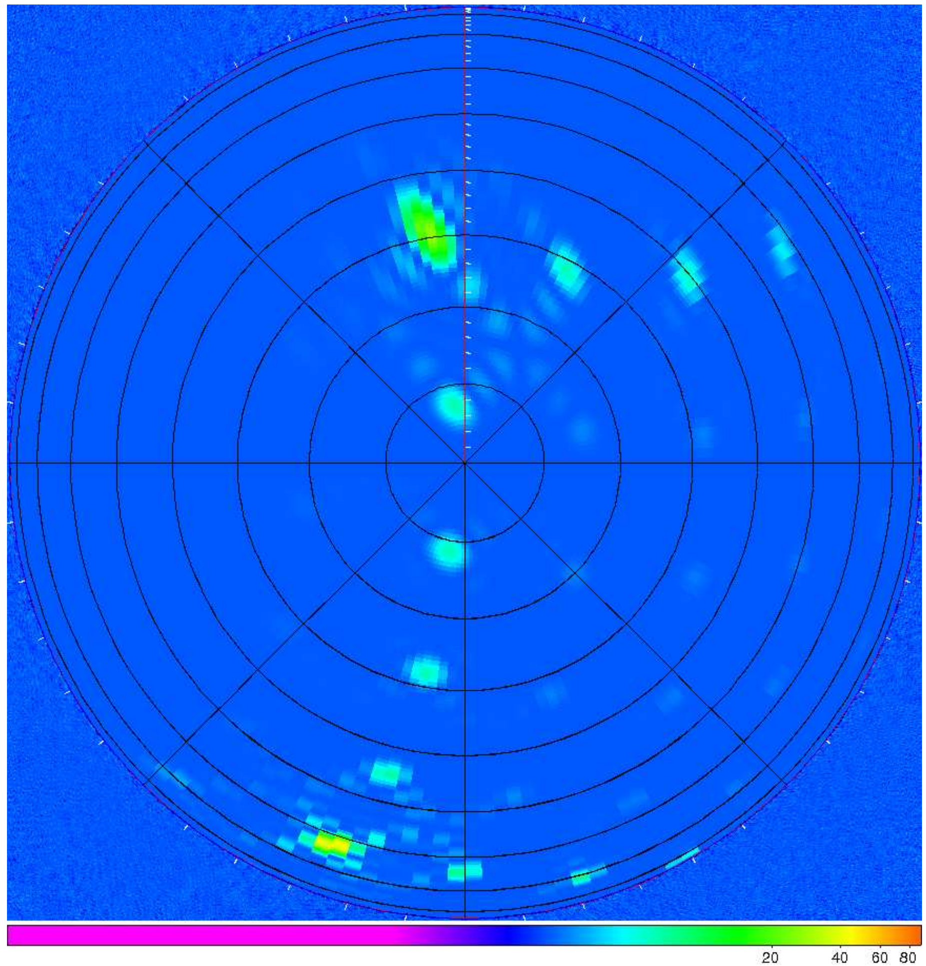
where (α, β) spatial, f frequency coordinates, $\mathbf{\Phi}_{\alpha\beta f}$: basis functions.

Dipole beam and the sky



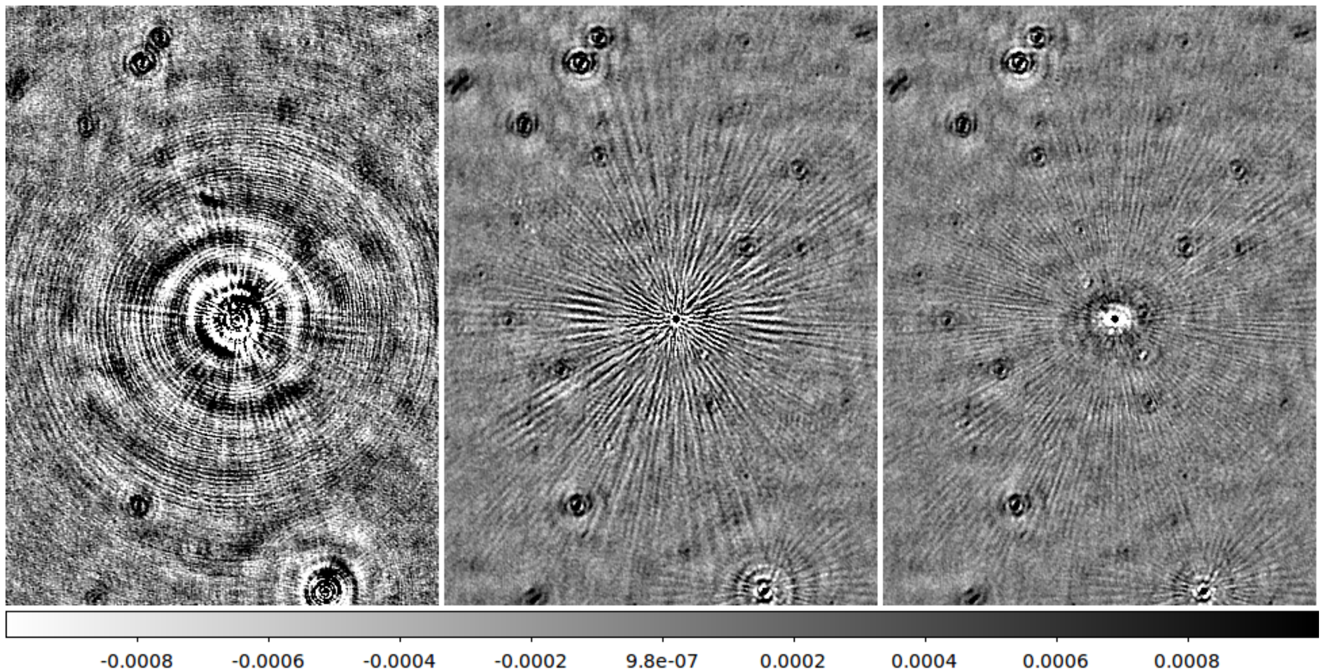
Magnitude of dipole beam projected onto the sky, zenith on top

Station beam



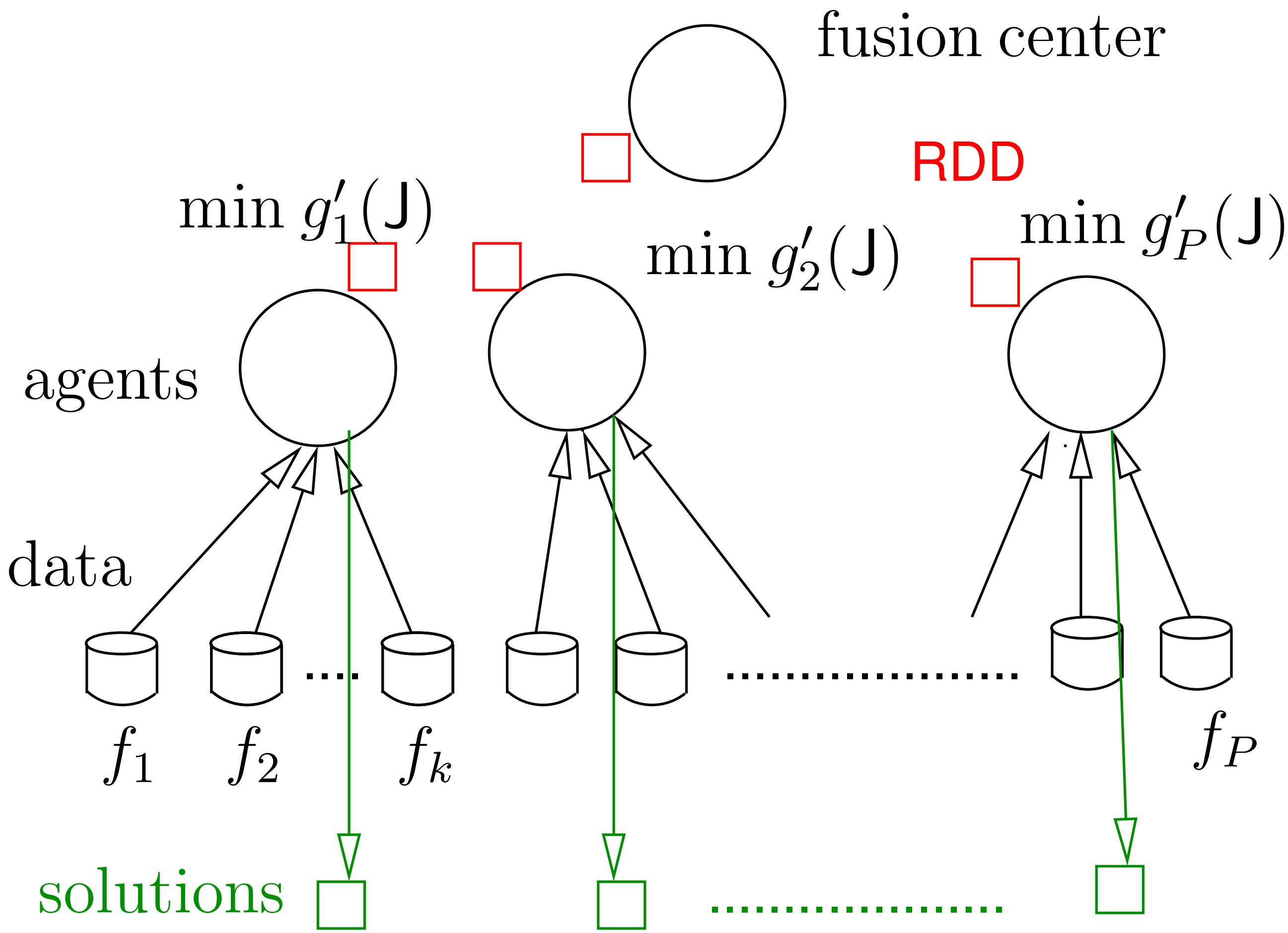
Station (array of dipoles) creates a focused beam, with sidelobes

Ionosphere



Atmospheric conditions in troposphere and ionosphere create errors. left: beam and ionospheric errors, middle: ionospheric errors only, right: after calibration

Distributed calibration



Data distributed across a cluster, calibration performed distributed and solutions also stored distributed.

Consensus optimization

Solutions affected by a unitary ambiguity, we have $\widehat{\mathsf{J}}_f = \mathsf{J}_f \mathsf{U}_f$ where U_f unknown unitary matrix. Eliminate unitary ambiguity by

$$\mathsf{A}_p \widehat{\mathsf{J}}_f \mathsf{C}_{pqf} (\mathsf{A}_q \widehat{\mathsf{J}}_f)^H = \mathsf{A}_p \mathsf{X} \mathbf{\Phi}_{\alpha\beta f} \mathsf{C}_{pqf} (\mathsf{A}_q \mathsf{X} \mathbf{\Phi}_{\alpha\beta f})^H$$

and find X satisfying this for all (α, β) and f. Using cost functions $h_j(\mathsf{X})$

$$\mathsf{X} = \arg\min_{\mathsf{X}} \sum_j h_j(\mathsf{X}) + \lambda \|\mathsf{X}\|^2 + \mu \|\mathsf{X}\|_1$$

using elastic net regularization to minimize over fitting (physically realistic solution). Caveat: not easy to solve directly.
Convert to a consensus problem as

$$\mathsf{X}_1, \mathsf{X}_2, \ldots, \mathsf{Z} = \arg\min_{\mathsf{X}_1, \ldots, \mathsf{Z}} \sum_j h_j(\mathsf{X}) + \lambda \|\mathsf{Z}\|^2 + \mu \|\mathsf{Z}\|_1$$

subject to $\mathsf{X}_j = \mathsf{Z}$ for all j.

ADMM

Augmented Lagrangian

$$L(\mathsf{X}_{f_1},\ldots,\mathsf{Z},\mathsf{Y}_{f_1},\ldots) = \sum_j h_j(\mathsf{X}_j) + \|\mathsf{Y}_j^H(\mathsf{X}_j - \mathsf{Z})\| + \frac{\rho}{2}\|\mathsf{X}_j - \mathsf{Z}\|^2 + \lambda\|\mathsf{Z}\|^2 + \mu\|\mathsf{Z}\|_1$$

Iterative optimization with $n = 1, 2, \ldots$

- ☐ Locally optimize to find

$$(\mathsf{X}_j)^{n+1} = \arg\min_{\mathsf{X}_j} L_j\left(\mathsf{X}_j, (\mathsf{Z})^n, (\mathsf{Y}_j)^n\right)$$

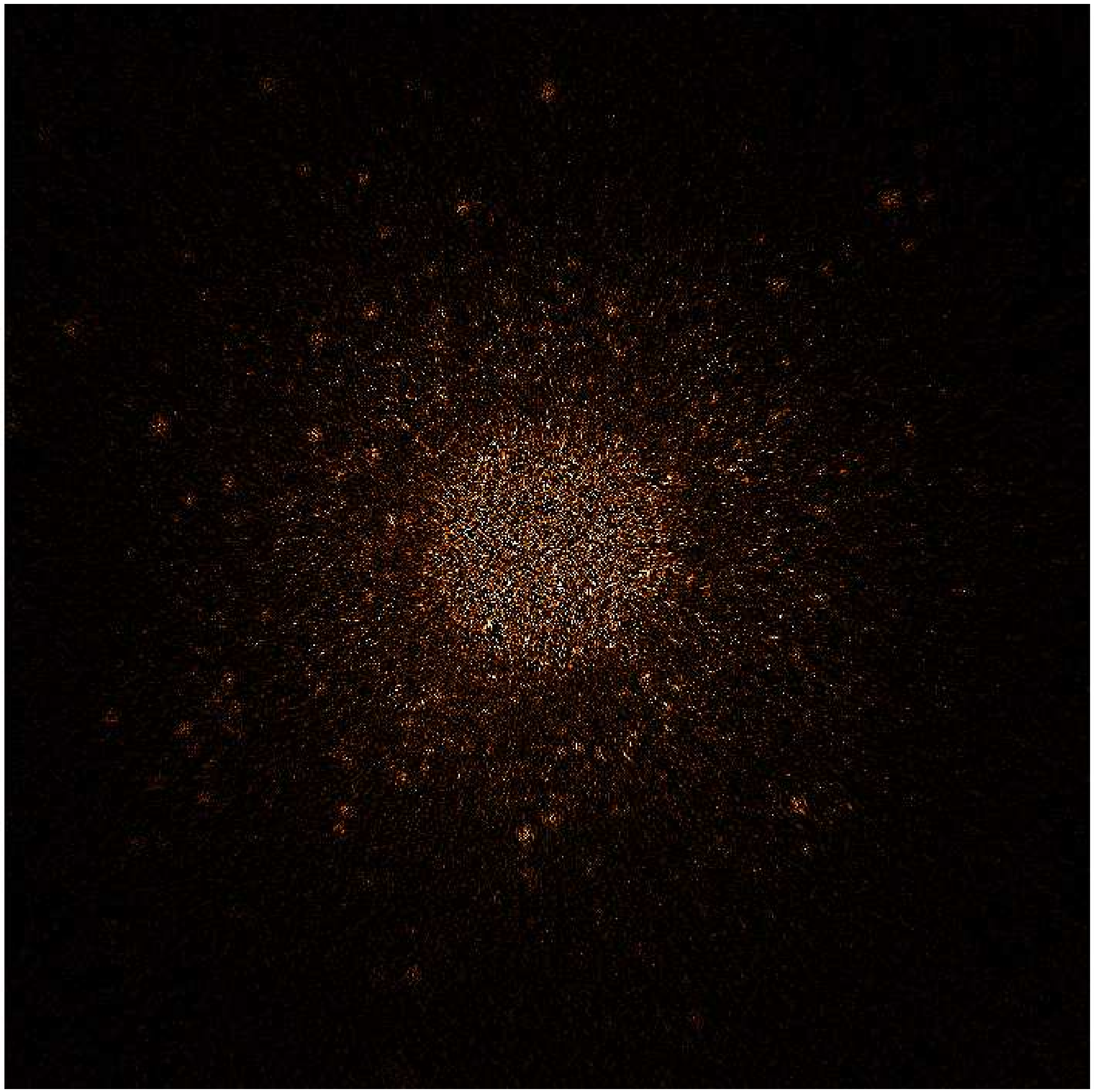
- ☐ Globally average and soft threshold (closed form solution)

$$(\mathsf{Z})^{n+1} = \arg\min_{\mathsf{Z}} \sum_j L_j\left((\mathsf{J}_j)^{n+1}, \mathsf{Z}, (\mathsf{Y}_j)^n\right)$$

- ☐ Locally update Lagrange multiplier

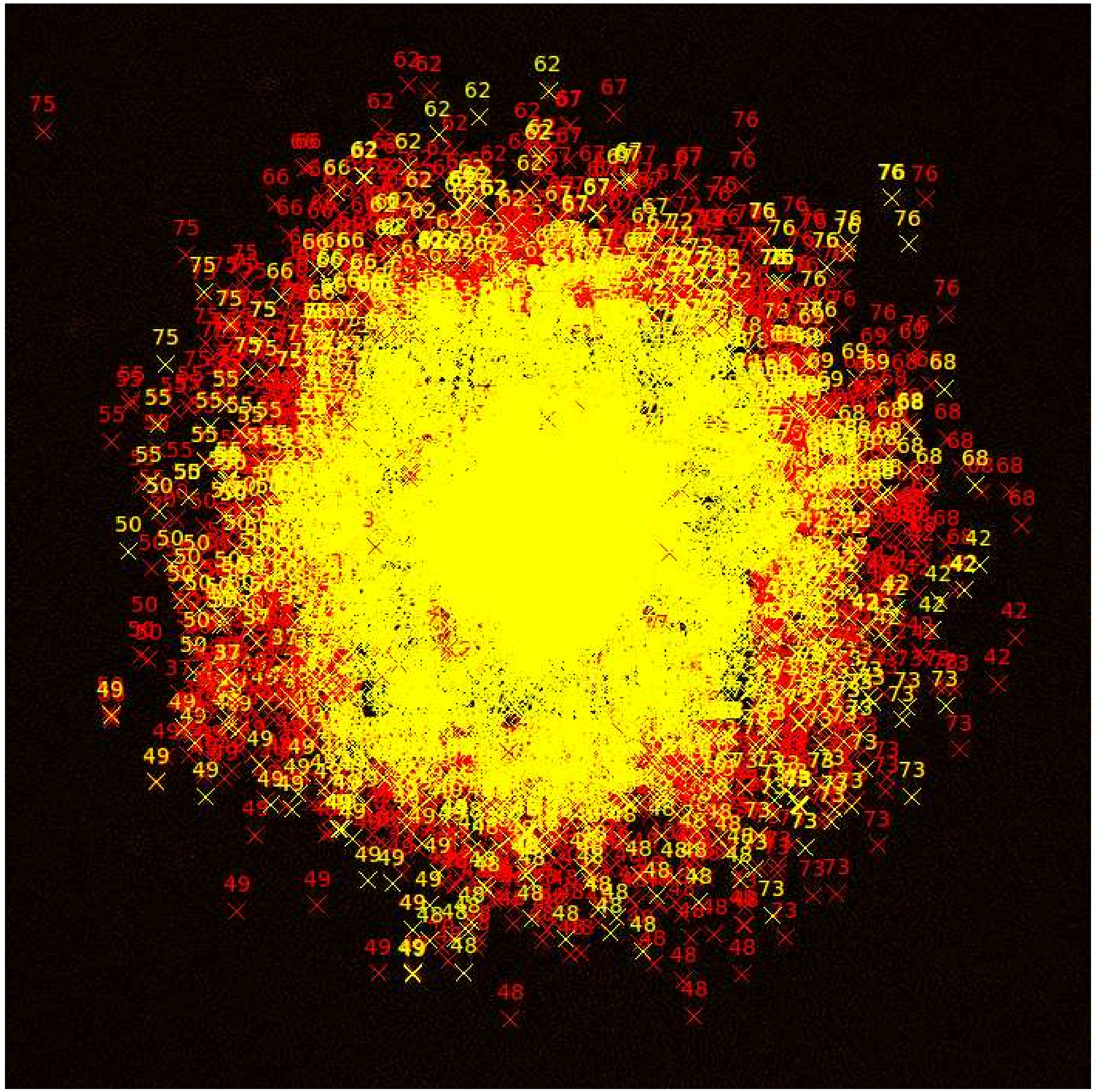
$$(\mathsf{Y}_j)^{n+1} = (\mathsf{Y}_j)^n + \rho((\mathsf{X}_j)^{n+1} - (\mathsf{Z})^{n+1})$$

A typical sky model



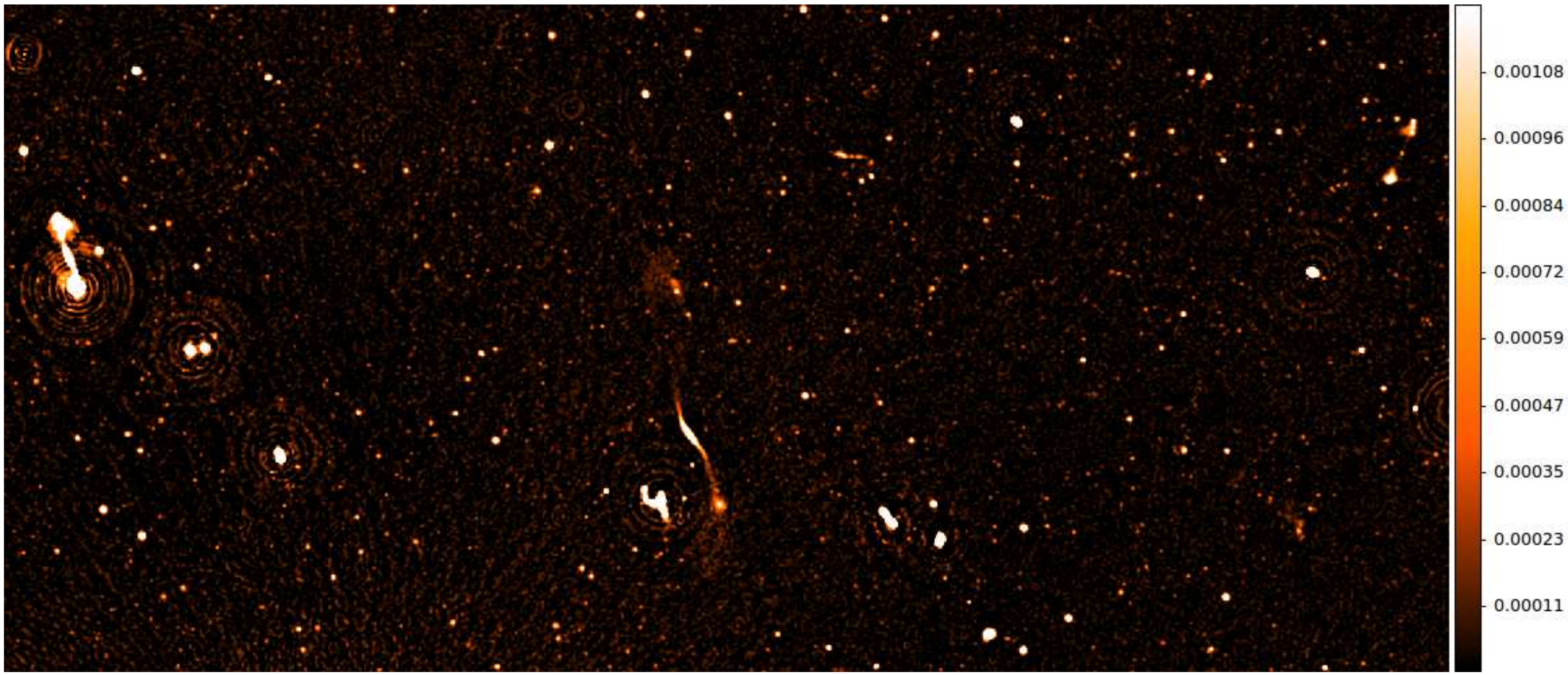
40×40 sq. deg. image, more sources appear in the center of the beam

Sources in calibration model



About 8000 sources covering a 10×10 square degrees

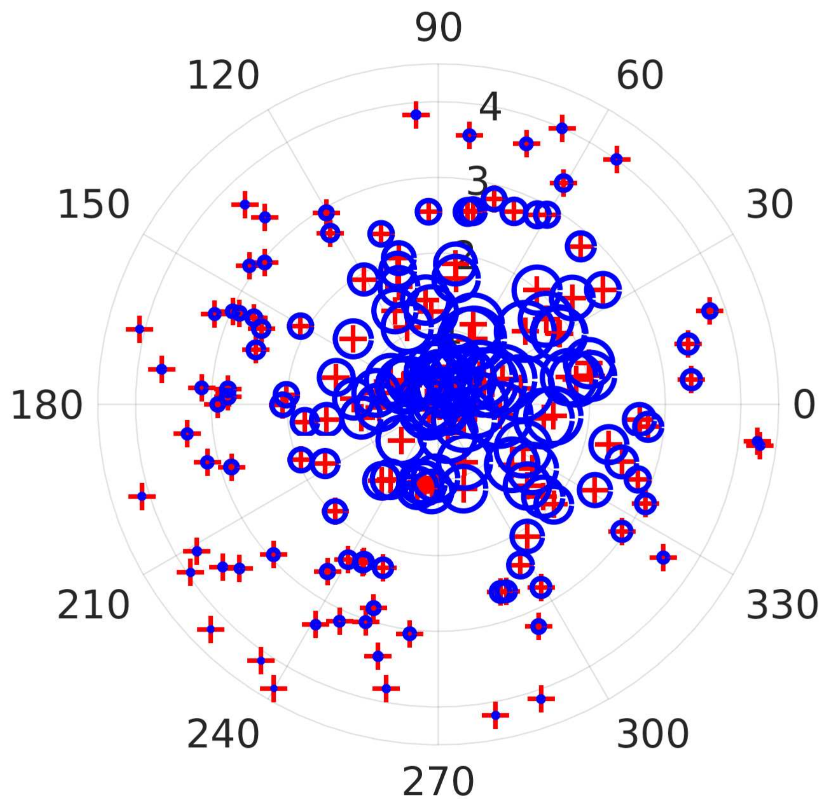
Small area with high sensitivity



Small area (about $1/1000$) of the full field of view, each source gives a unique C_{pqf}

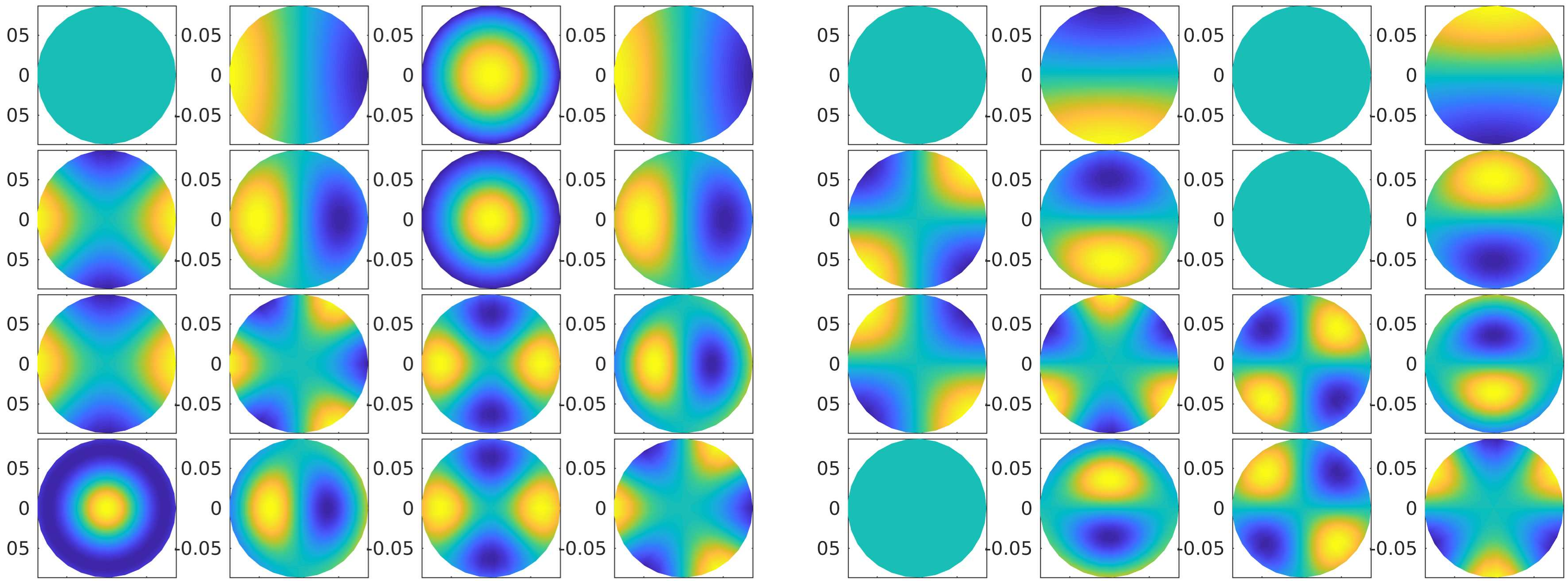
Simulation

- Simulate an array with $N = 6$ receivers, data corrupted by simulated beam/ionospheric errors.
- How well can we build models for systematic errors?



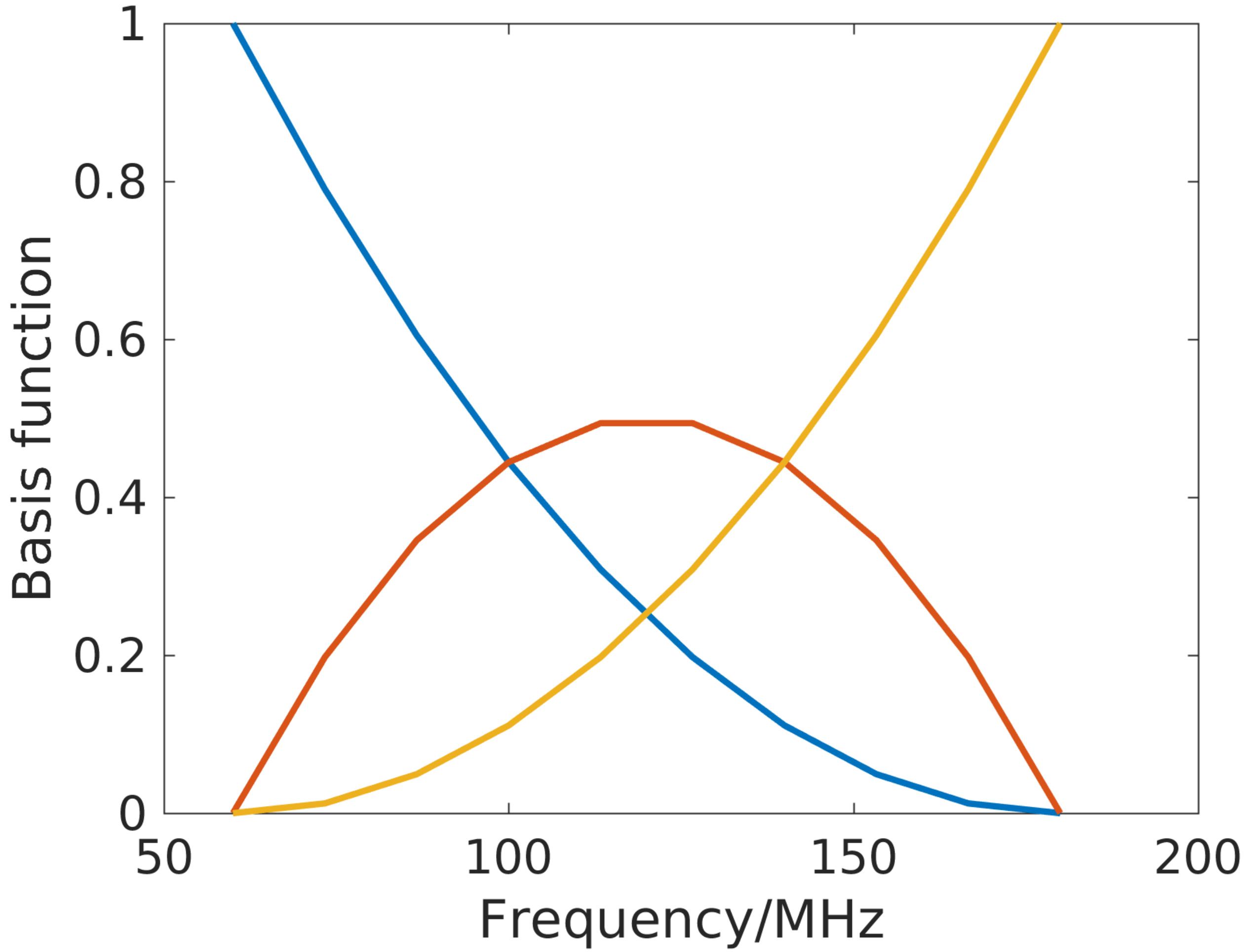
Sky model, > 100 distinct directions in the sky, each giving a sampling point

Spatial basis functions



Spherical harmonic basis (left) real (right) imaginary parts

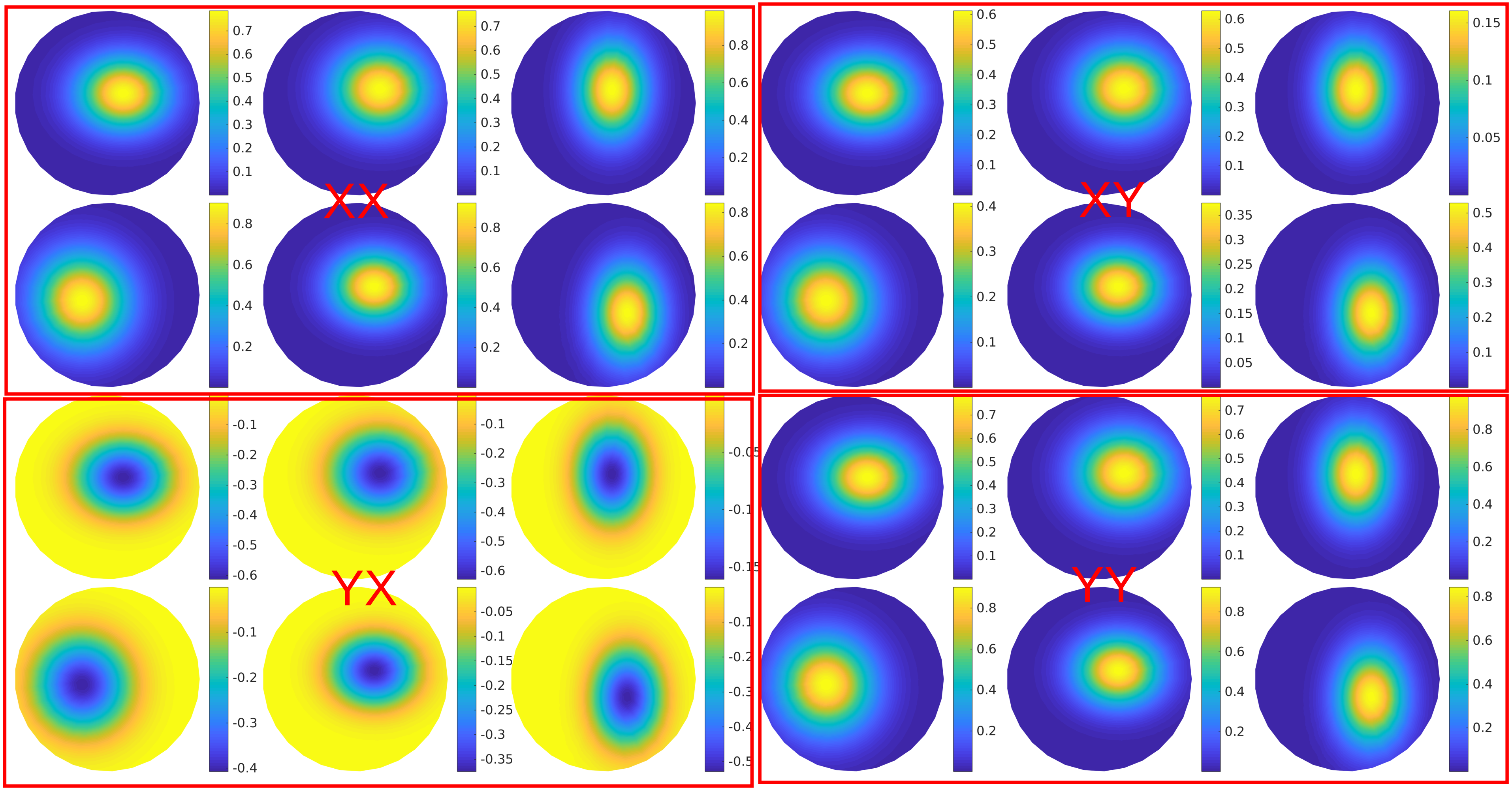
Frequency basis functions



Bernstein basis in frequency, 3 polynomials

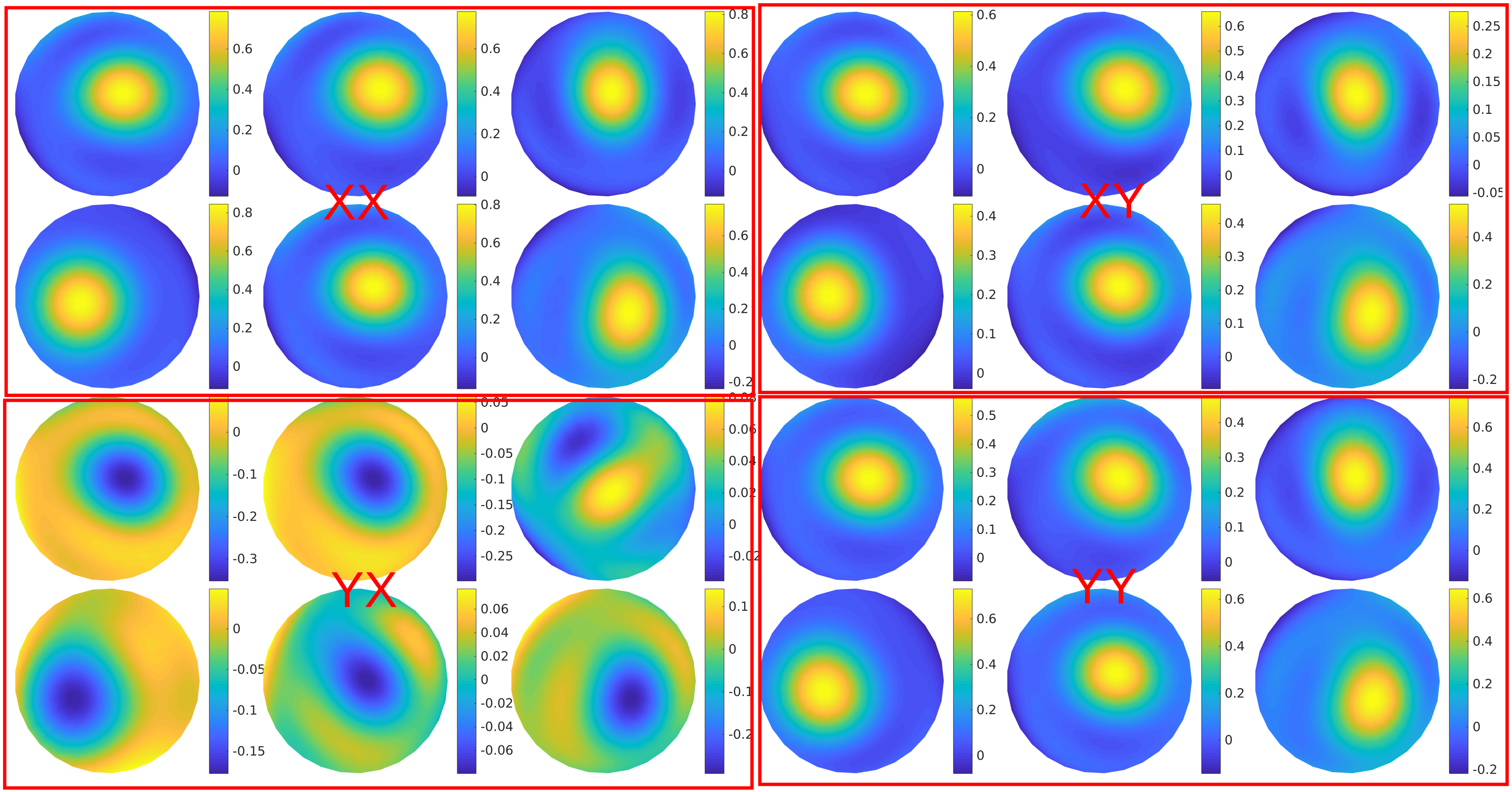
Finally, $\mathbf{\Phi}_{\alpha\beta f} =$ the product of spatial and frequency bases.

True systematic errors (real part)



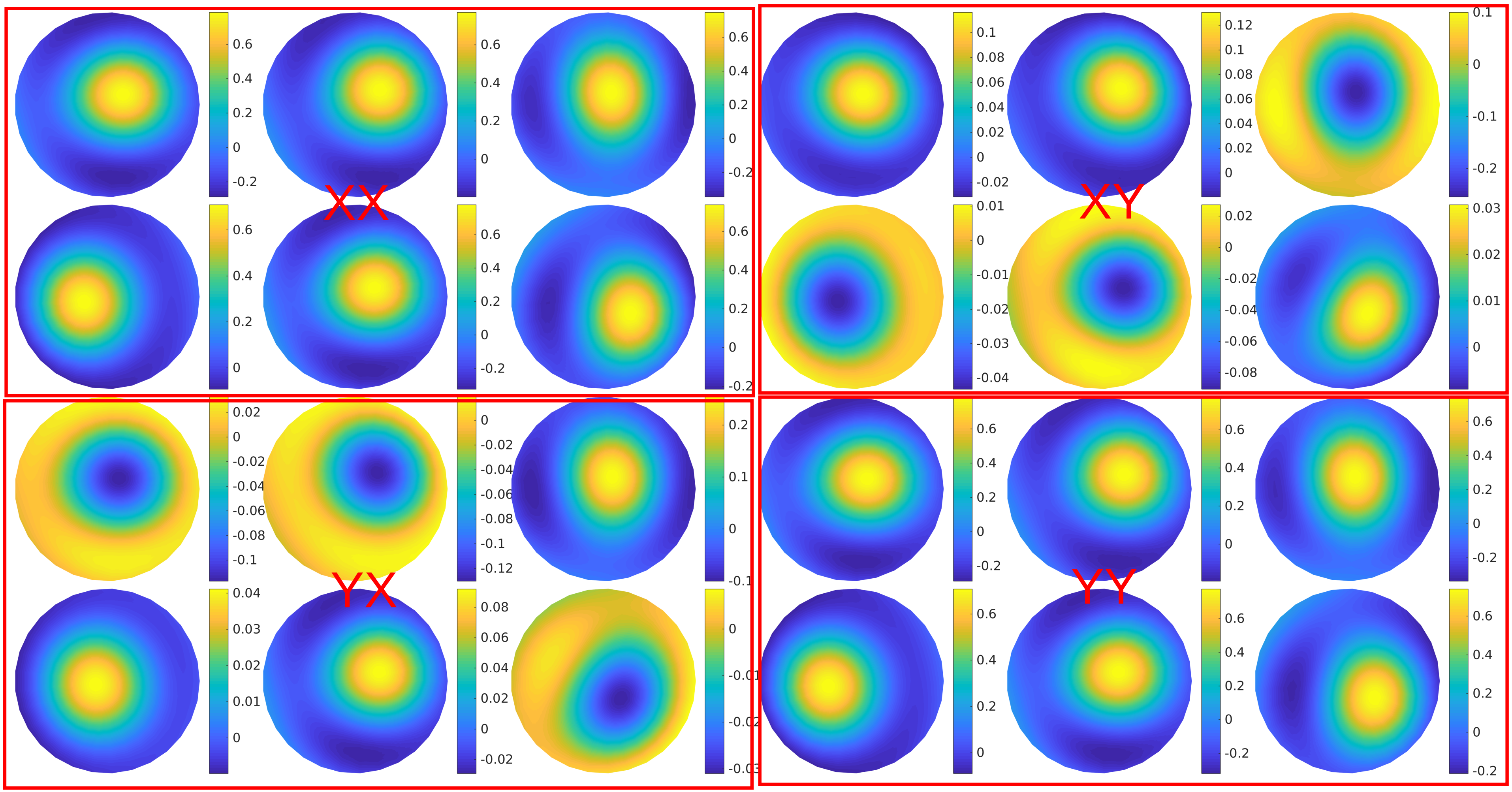
XX,XY,YX,YY polarizations

Linear model (real part)



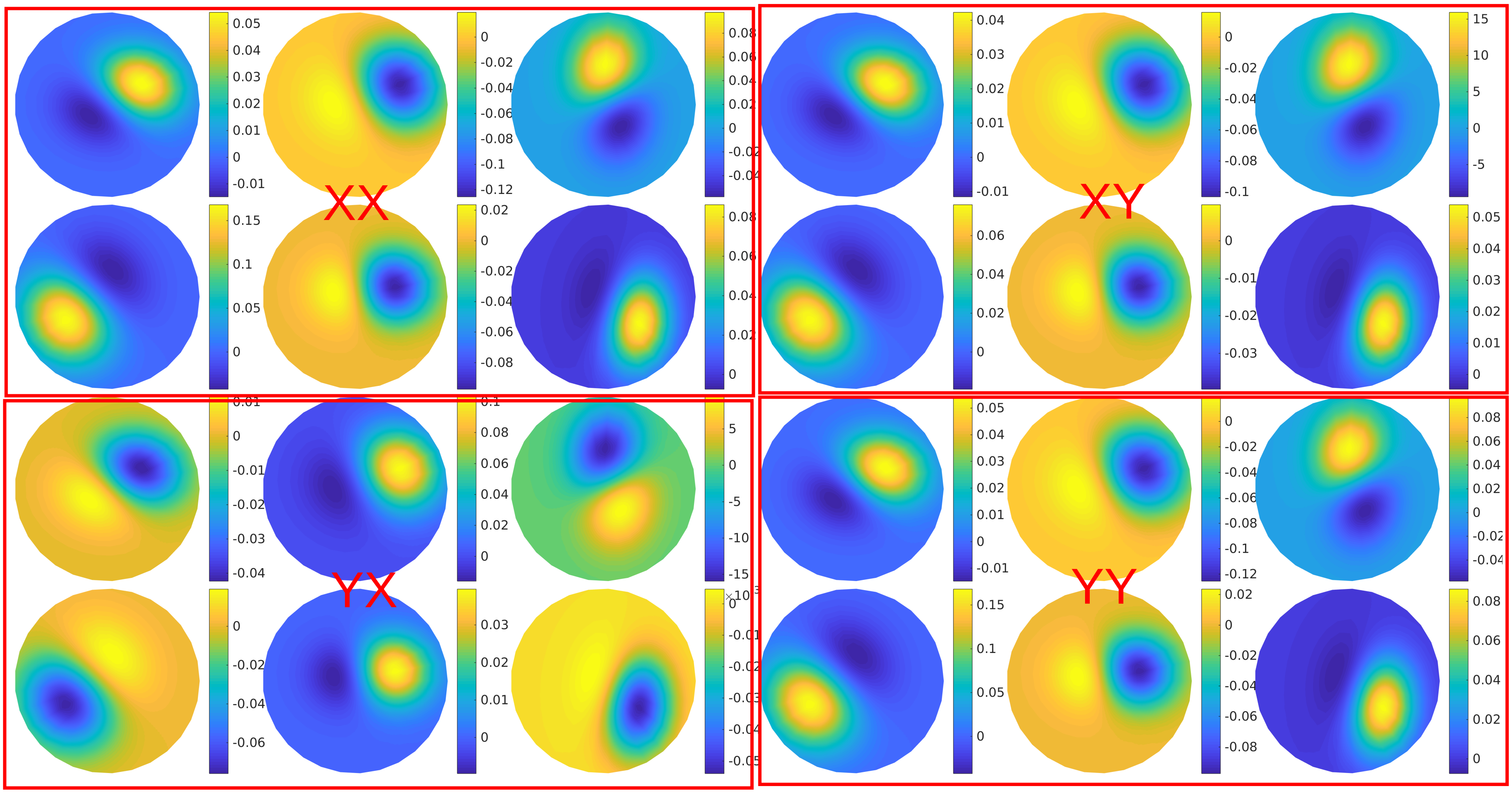
XX,XY,YX,YY polarizations

Distributed model (real part)



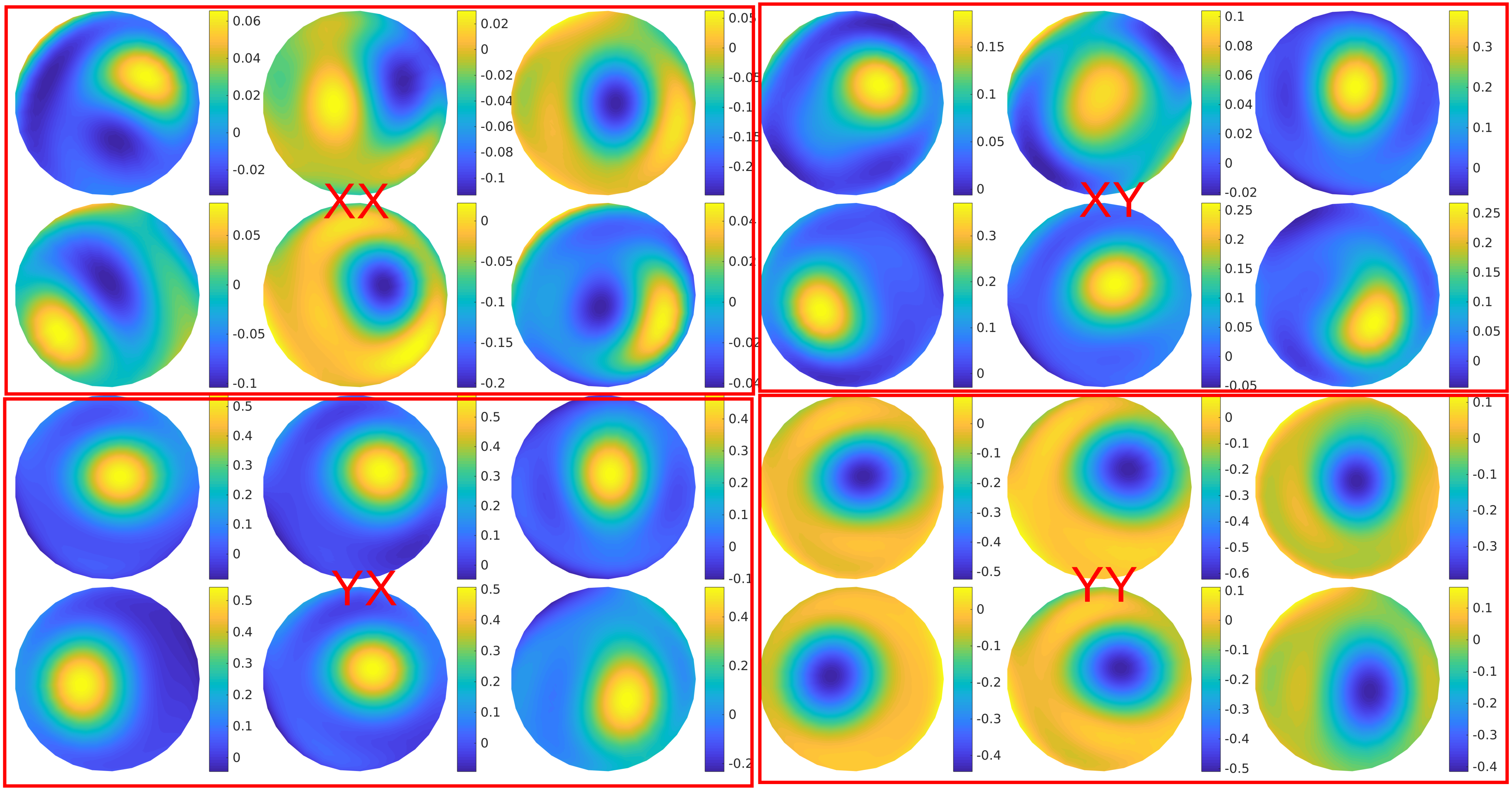
XX,XY,YX,YY polarizations

True systematic errors (imaginary part)



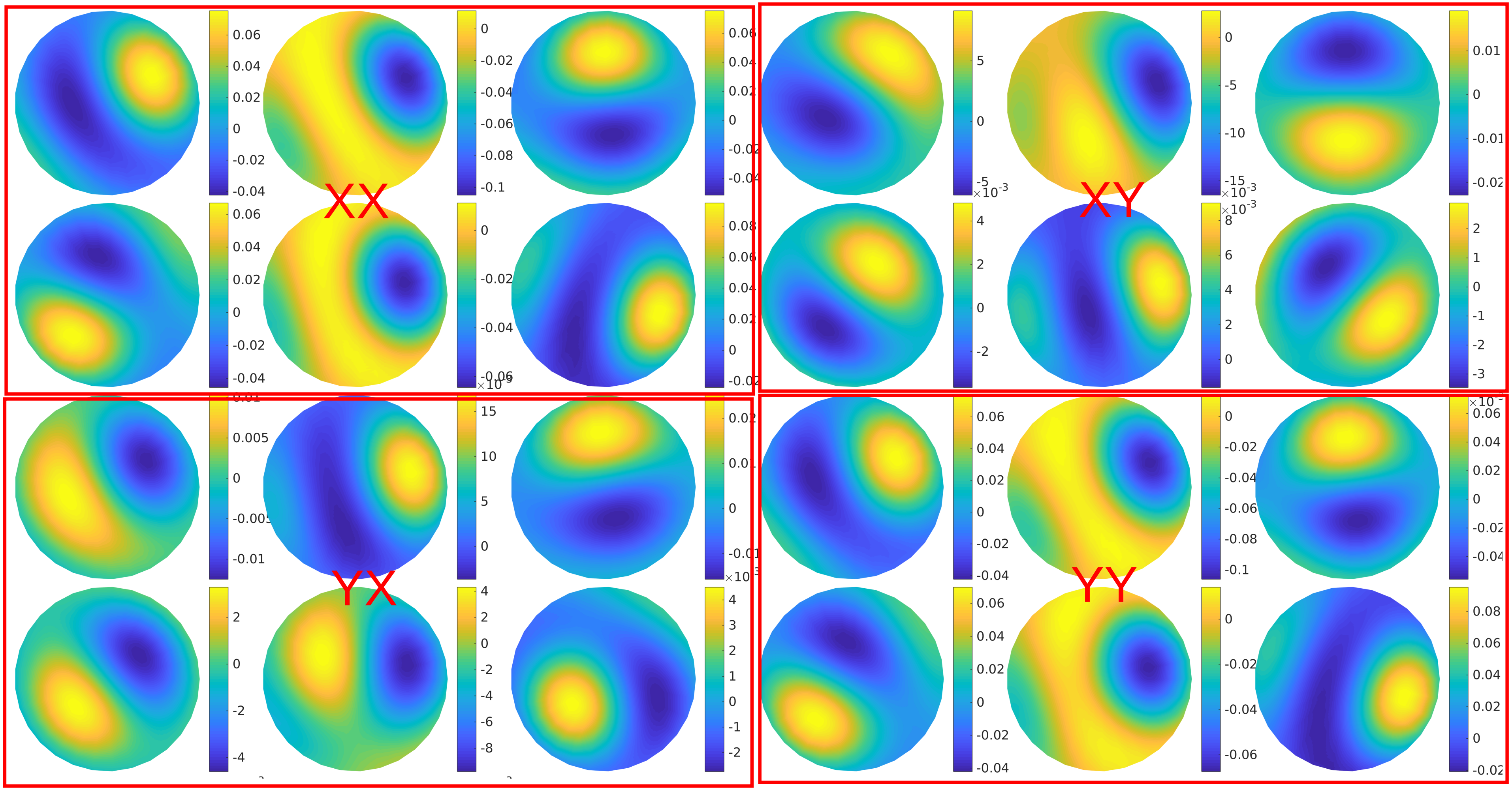
XX,XY,YX,YY polarizations

Linear model (imaginary part)



XX,XY,YX,YY polarizations

Distributed model (imaginary part)



XX,XY,YX,YY polarizations

Conclusions

- ☐ Distributed algorithm for construction of distributed models for ionosphere and beam shape: computationally efficient.
- ☐ Elastic net regularization gives best results.