ASTRON
Netherlands Institute for Radio Astronomy

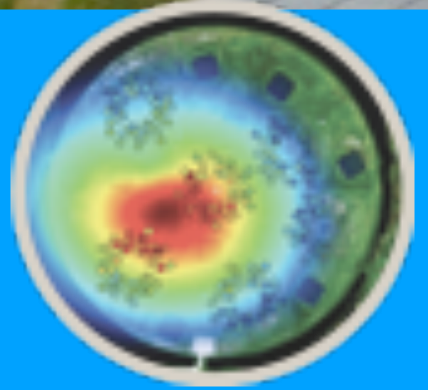
STELLAR

LOFAR

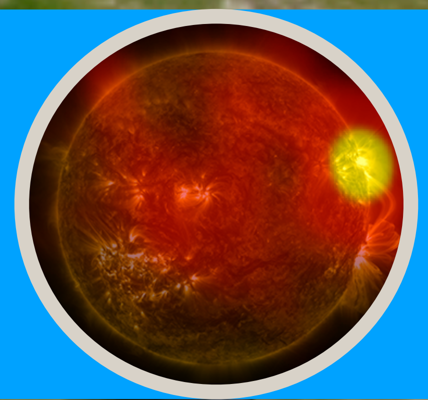
T1: Removal of Instrumental Effects in LOFAR Data & the LOFAR Solution Tool

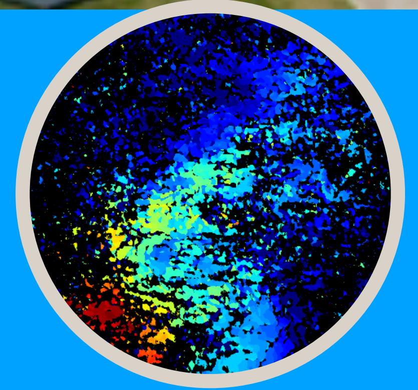
Henrik Edler

22/03/2021



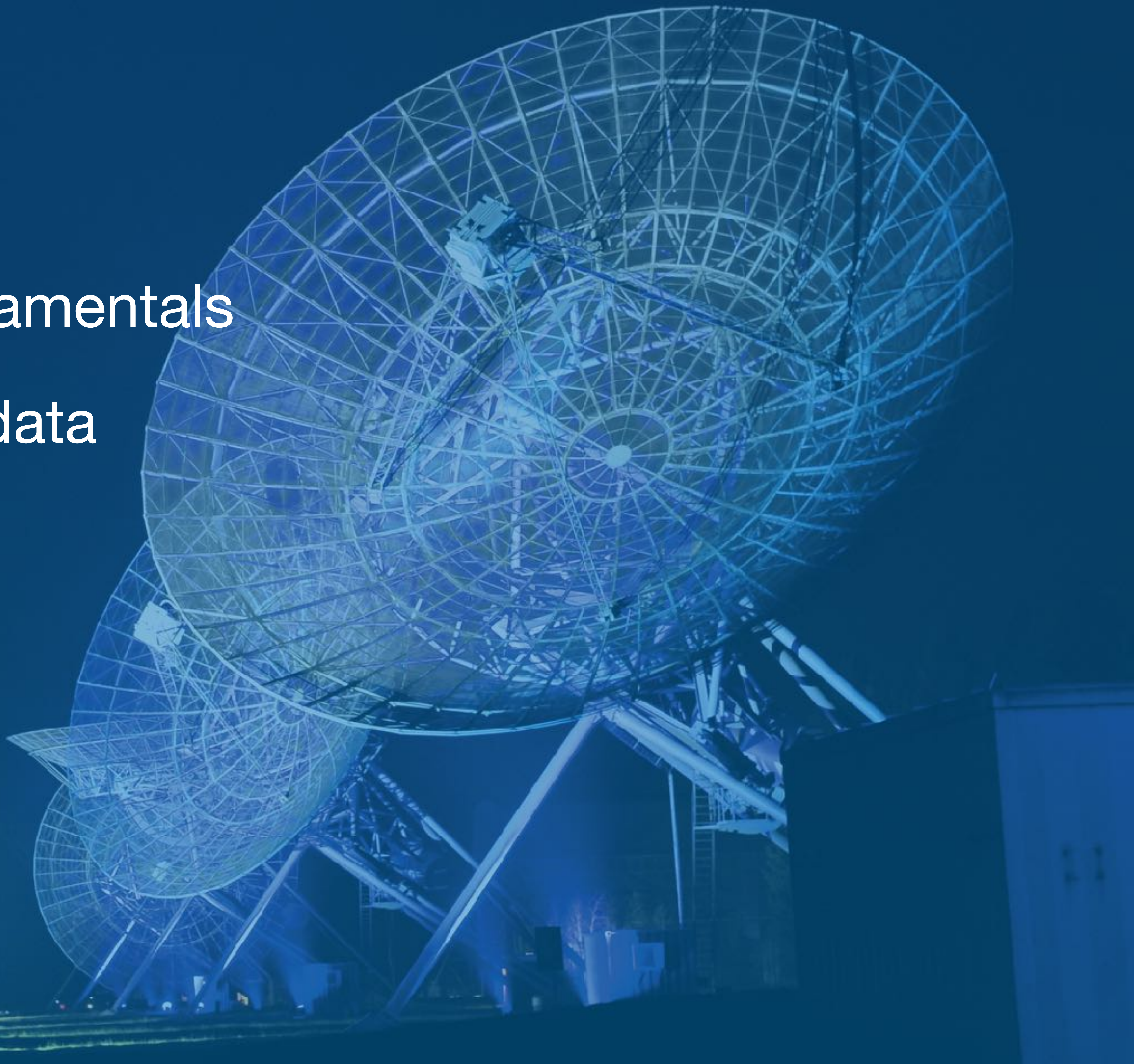
6th LOFAR Data School





Content

1. Brief recap of calibration fundamentals
2. Systematic effects in LOFAR data
3. The LOFAR solution tool
4. Calibration tutorial

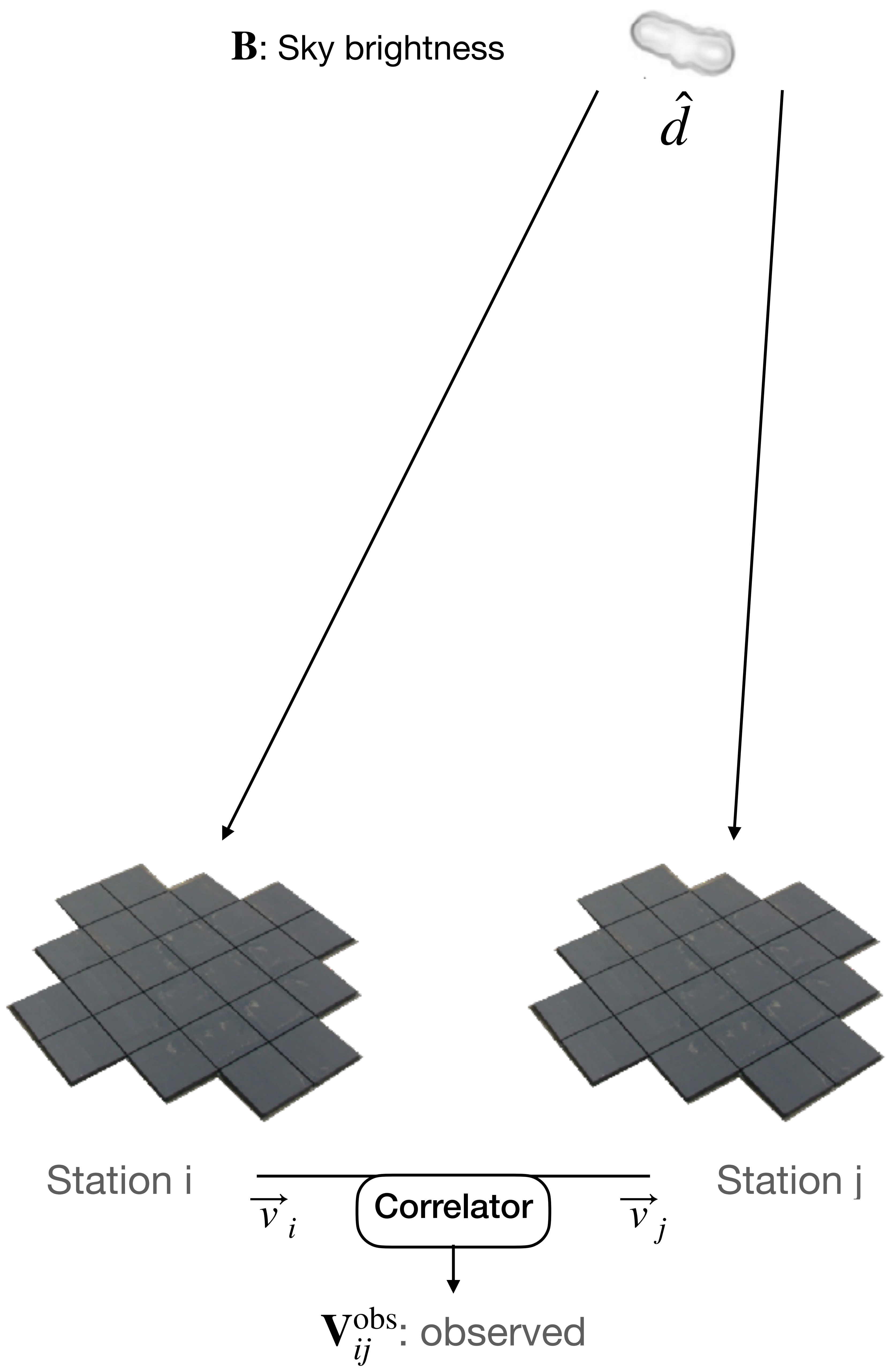


Calibration Basics

- Radio interferometry: measure visibilities
- Reconstruct sky brightness distribution from complex visibilities by applying the radio interferometer measurement equation:

$$\mathbf{V}_{ij} = \int_{4\pi} \mathbf{J}_i(\hat{d})\mathbf{B}(\hat{d})\mathbf{J}_j^{\dagger}(\hat{d})\mathrm{d}\Omega$$

- signal vector $\vec{e} = \begin{pmatrix} e_x \\ e_y \end{pmatrix}$ ➡ $\mathbf{B} = \langle \vec{e}\,\vec{e}^{\dagger} \rangle$
- voltage vector $\vec{v} = \begin{pmatrix} v_a \\ v_b \end{pmatrix}$ ➡ $\mathbf{V}_{ij} = 2\langle \vec{v}_i \vec{v}_j^{\dagger} \rangle$
- **But**: Need to determine Jones matrices to describe systematic effects!



For review of the RIME & Jones formalism check:

- J. P. Hamaker, J. D. Bregman and R. J. Sault 1996, A&AS
- O. Smirnov 2011, A&A

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Calibration Basics

- Need **Jones-matrices** to recover true visibility data :

observed visibility
between stations i and j

$$\mathbf{V}_{ij}^{obs} = \mathbf{J}_i \mathbf{V}_{ij}^{\text{true}} \mathbf{J}_j^{\top}$$

- Jones matrices: 2x2 matrices: $\overrightarrow{v} = \mathbf{J}\overrightarrow{e}$

- Total Jones matrix is product of individual effects in physical order

$$\mathbf{J} = \mathbf{J}_{\text{clock}} \times \mathbf{J}_{\text{bandpass}} \times \mathbf{J}_{\text{leak}} \times \mathbf{J}_{\text{beam}} \times \mathbf{J}_{\text{iono}} \times \ldots$$

- Some examples for Jones matrices:

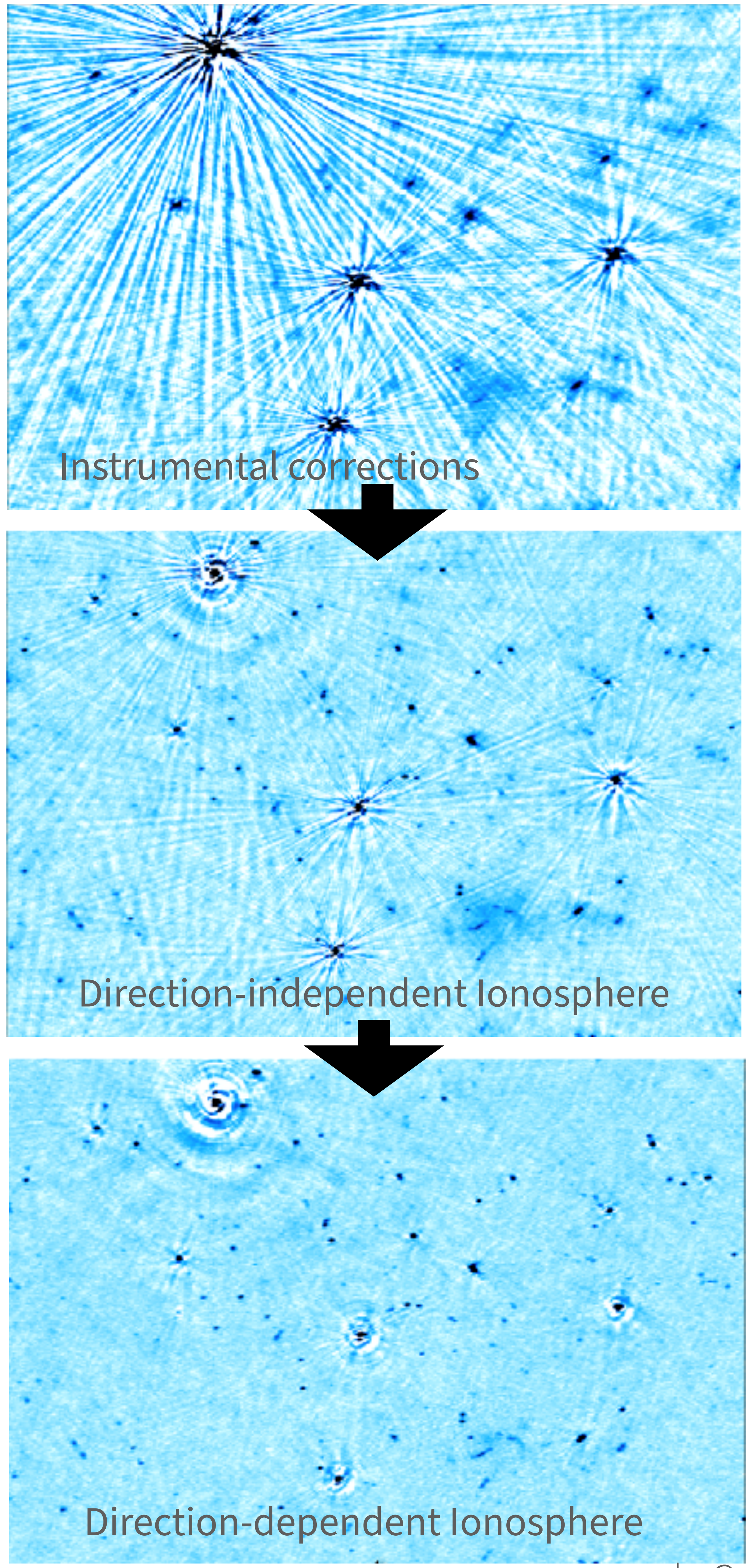
$$\mathbf{J}_{\text{rot}} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \quad \mathbf{J}_{\text{bandpass}} = \begin{pmatrix} g_X & 0 \\ 0 & g_Y \end{pmatrix} \quad \mathbf{J}_{\text{pol. misalignment}} = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i \nu \Delta t} \end{pmatrix} \quad \mathbf{J}_{\text{clock}} = e^{2\pi i \nu t} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{J}_{\text{full-Jones}} = \begin{pmatrix} a_{xx}e^{i\phi_{xx}} & a_{xy}e^{i\phi_{xy}} \\ a_{yx}e^{i\phi_{yx}} & a_{yy}e^{i\phi_{yy}} \end{pmatrix}$$

"Jones-scalar"

Calibration Basics

Why do we need to calibrate?

- We want only “trustable” emission
- Create images with scientific value
- Accurate astrometry
- Align flux scale, e.g. for spectral studies



de Gasperin

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Systematic Effects in LOFAR data

Systematic effect	Type of Jones matrix[a]	Ph/Amp/Both[b]	Frequency dependency	Direction dependent?	Time dependent?
Clock drift	Scalar	Ph	$\propto \nu$	No	Yes (many seconds)
Polarisation alignment	Diagonal	Ph	$\propto \nu$	No	No
Ionosphere - 1st ord. (dispersive delay)	Scalar	Ph	$\propto \nu^{-1}$	Yes	Yes (few seconds)
Ionosphere - 2sn ord. (Faraday rotation)	Rotation	Both	$\propto \nu^{-2}$	Yes	Yes (few seconds)
Ionosphere - 3rd ord.	Scalar	Ph	$\propto \nu^{-3}$	Yes	Yes (few seconds)
Ionosphere - scintillations	Diagonal	Amp	–	Yes	Yes (few seconds)
Dipole beam	Full-Jones	Both	–	Yes	Yes (minutes)
Bandpass	Diagonal	Amp	–	No	No

Instrumental effects

Isolated and discussed in *de Gasperin et al. 2019, A&A*

Calibrator source

Find solutions for instrumental effects:
- Clock drift
- Polarization misalignment
- Bandpass

Transfer solutions

Target field

- Solve for direction-dependent effects (ionosphere)

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

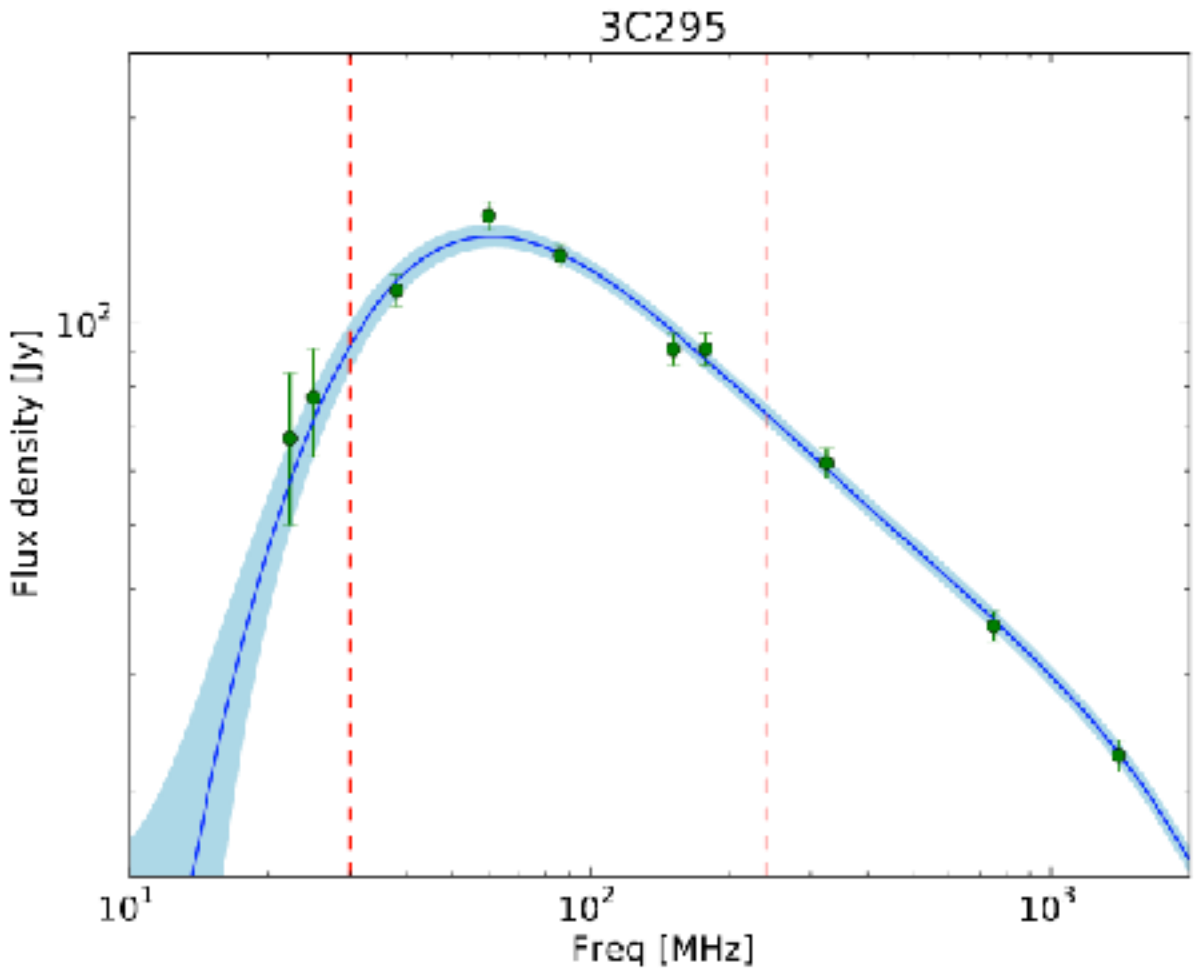
6th LOFAR Data School

LOFAR Calibrators

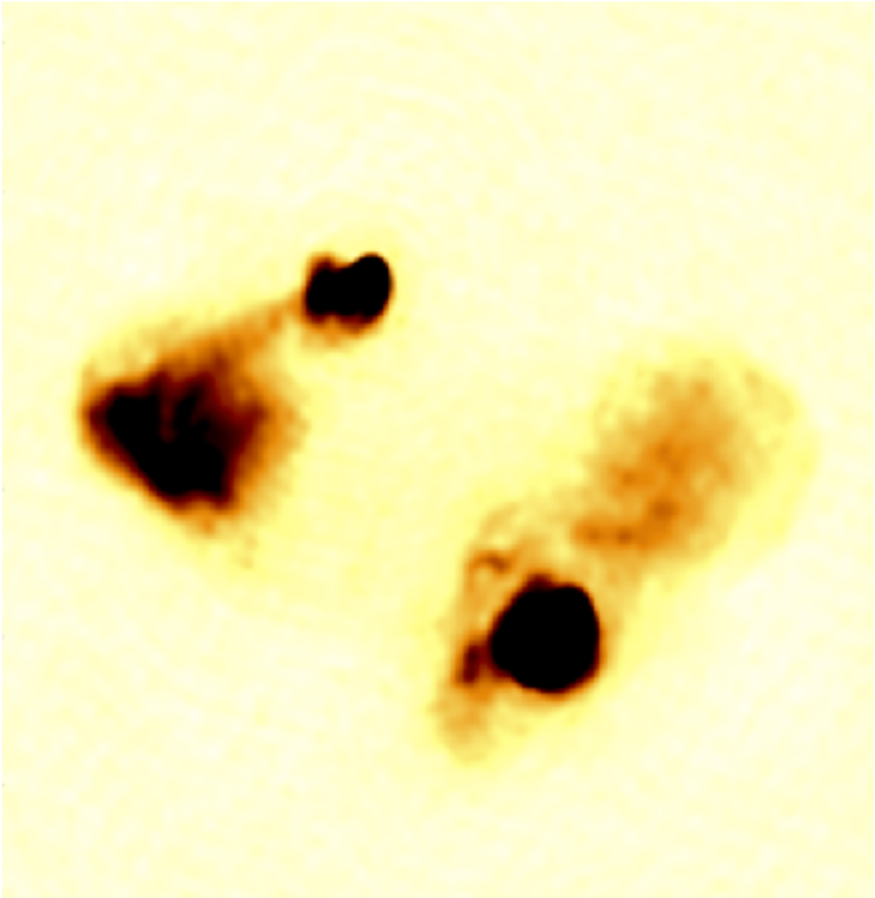
- Calibrator source requirements:
  - Bright (dominate the field)
  - Simple morphology
  - Good model
  - Well-defined spectrum
- Only very few such sources for LOFAR!

Name	S_{150MHz} [Jy]	Morphology
3C48	64.768	Point
3C147	66.738	Point
3C196	83.084	Double
3C286	27.477	Point
3C295	97.763	Double
3C380	77.352	Point+Diffuse
CygA	10690.0	FRII

G. H. Heald et al. 2015



Scaife & Heald 2012



3C196 at sub-arcescond resolution (de Bruyn)

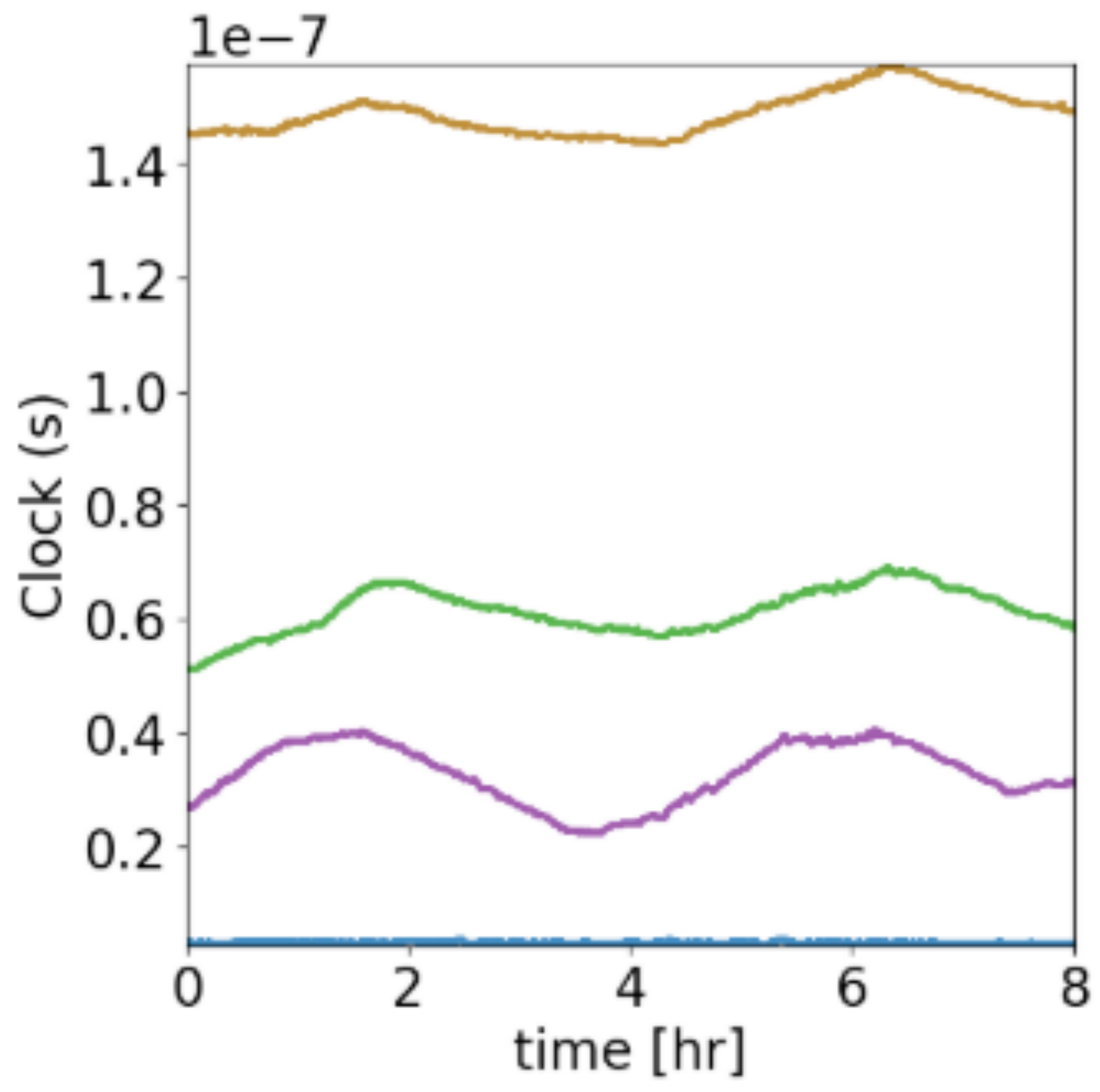
Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

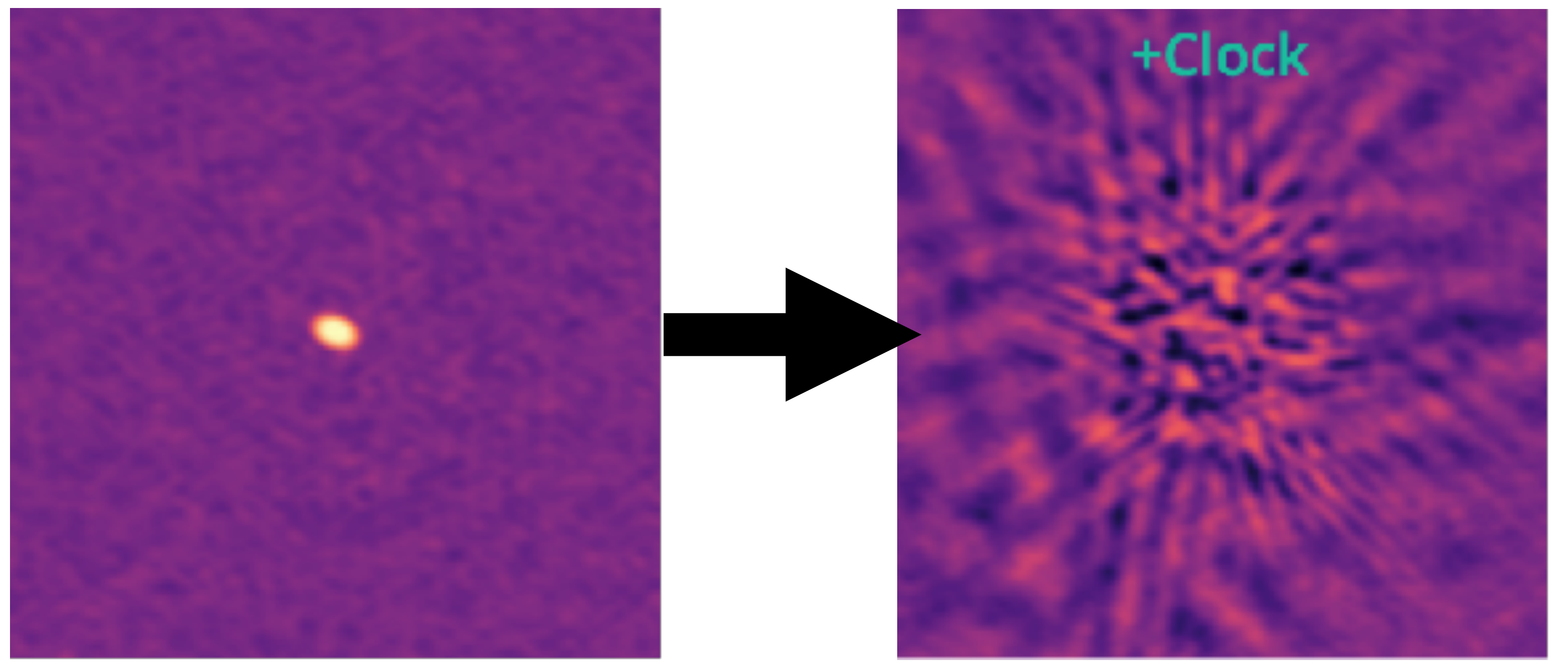
Instrumental Systematic Effects

Clock drift

- Remote stations have individual GPS-synchronized clocks
    - These clocks drift around O(10ns/h)
- Scalar phase error: $\Delta\phi = 2\pi\nu\Delta t$
    - Characteristic frequency dependence $\propto \nu$!

$$\mathbf{J}_{\text{clock}} = e^{2\pi i \nu t} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$





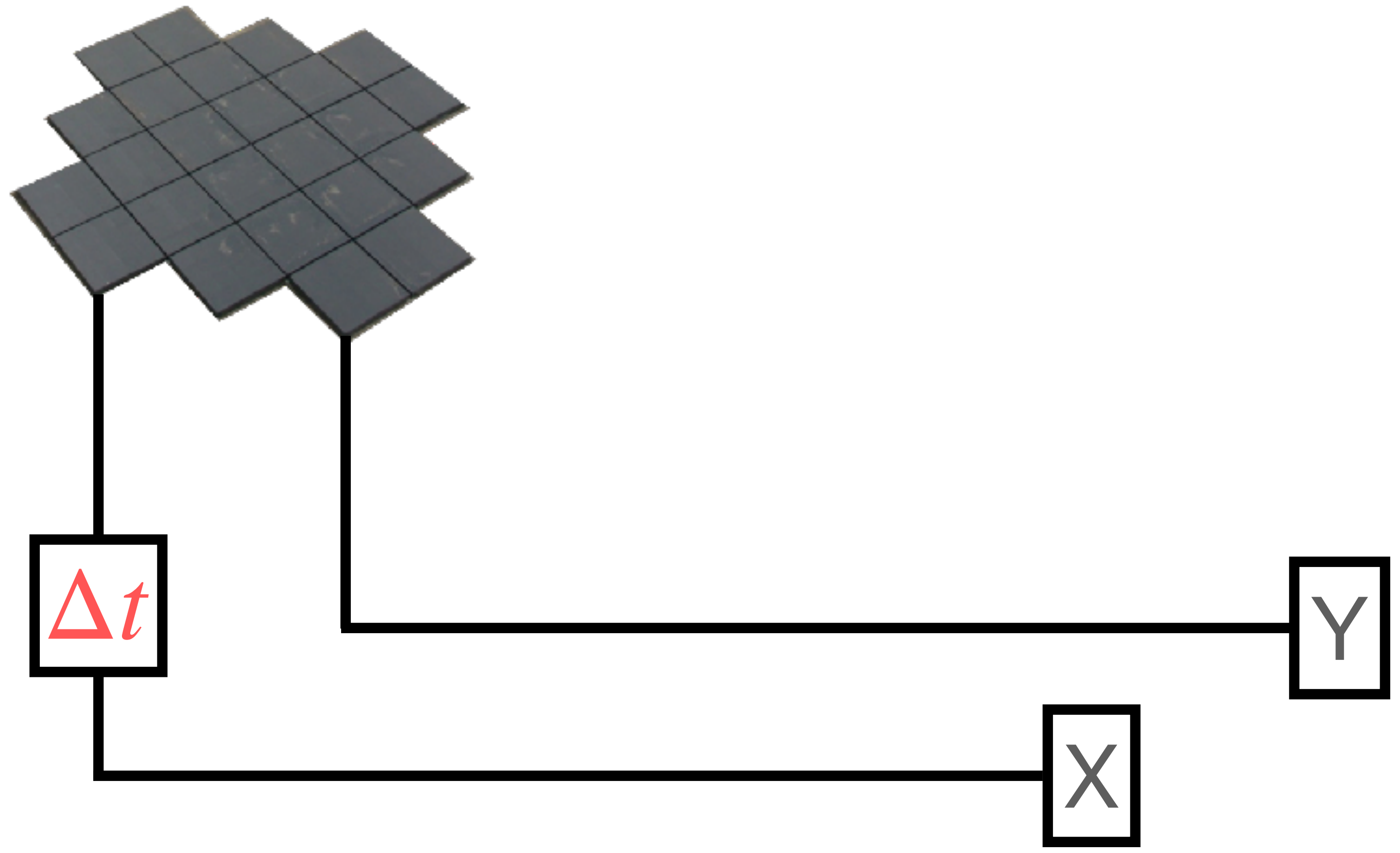
The effect of a simulated clock error on a point source

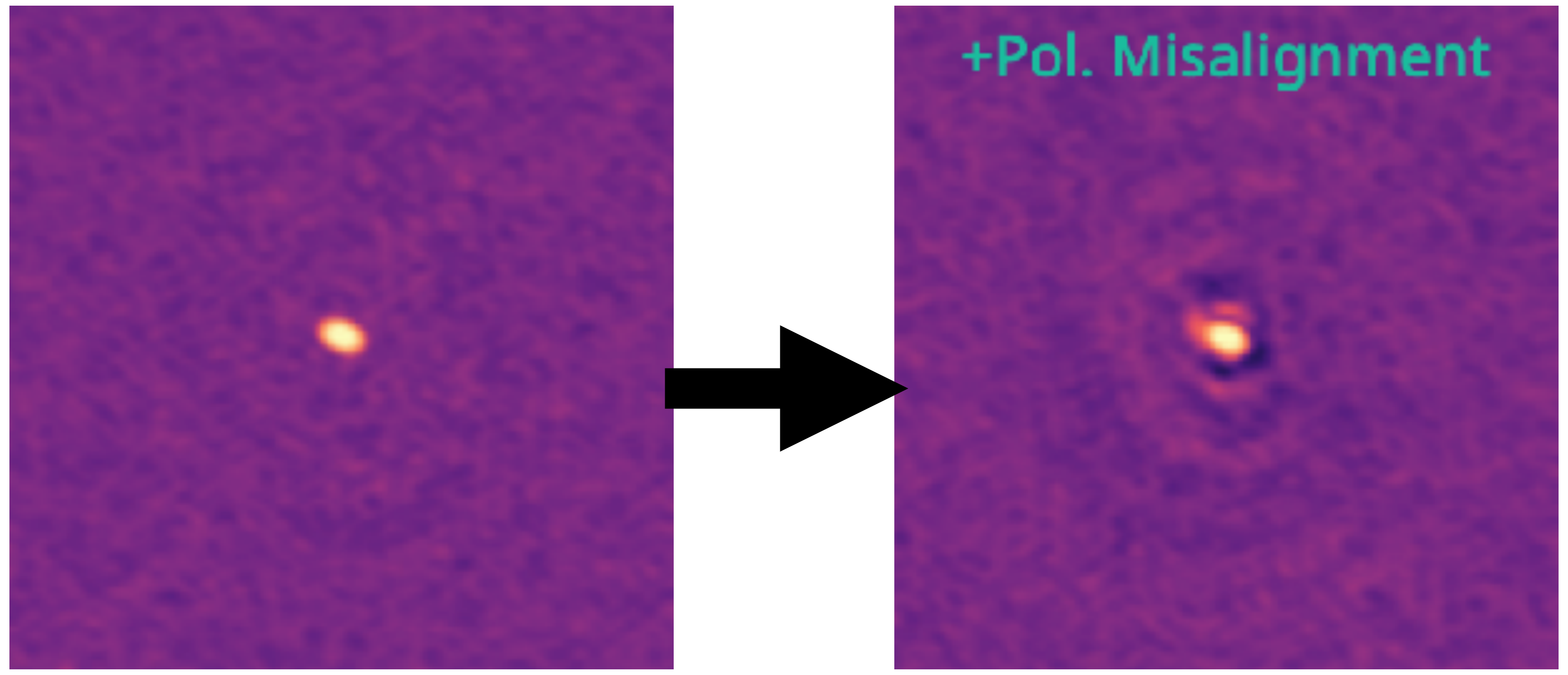
Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Instrumental Systematic Effects

Polarization Misalignment

- Some LOFAR stations show the presence of a constant timing-delay between the X and Y polarization signal O(ns)
- Attributed to calibration of station electronics
- Phase-only diagonal matrix: $$\mathbf{J}_{\text{pol. misalignment}} = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i \nu \Delta t} \end{pmatrix}$$





The effect of simulated pol. misalignment on a point source

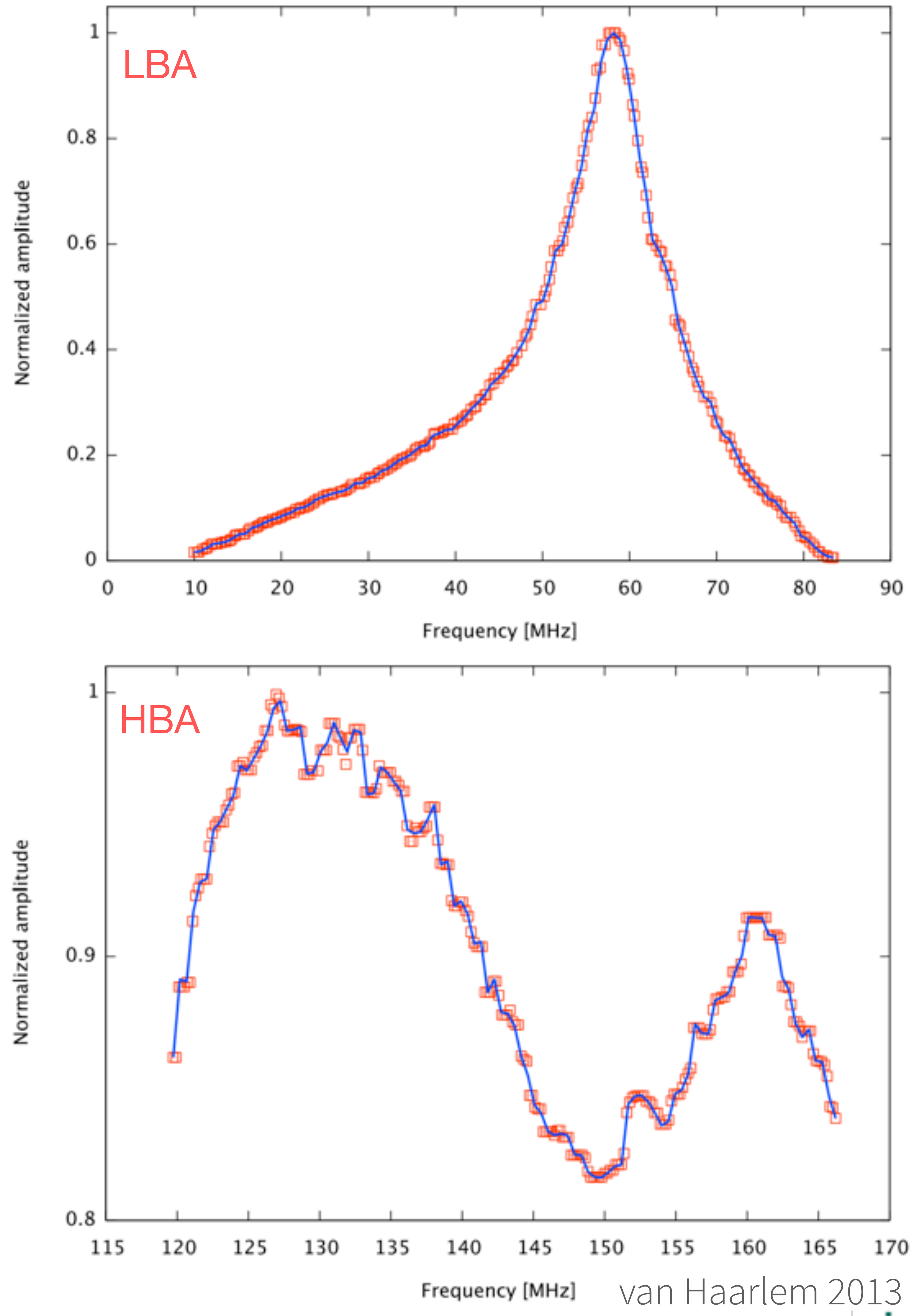
Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Instrumental Systematic Effects

Bandpass

- Frequency dependence of the instrument response
- Largely shaped by dipoles
- Small deviations between stations
- Real valued diagonal matrix

$$\mathbf{J}_{\text{bandpass}} = \begin{pmatrix} a_{xx} & 0 \\ 0 & a_{yy} \end{pmatrix}$$

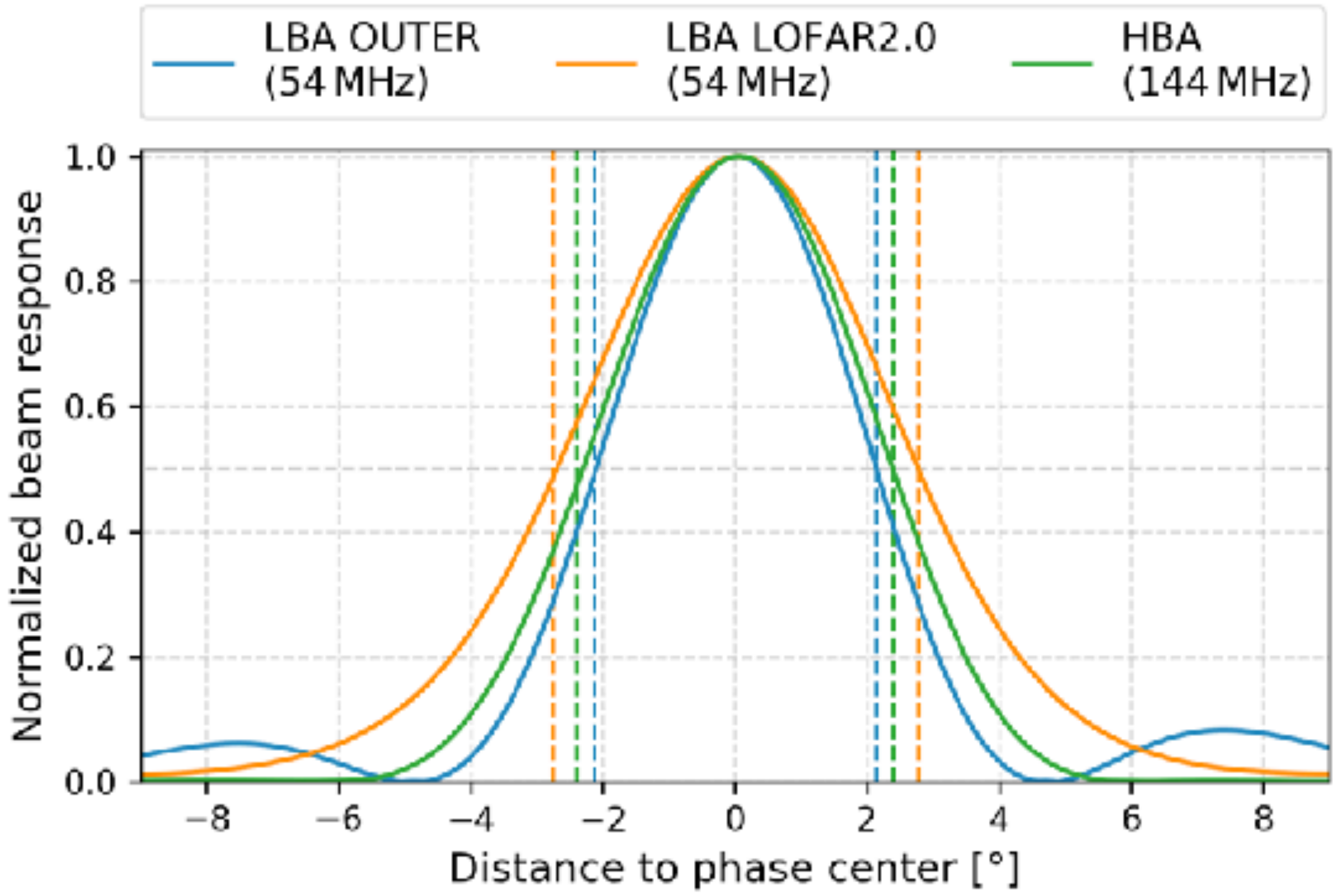


van Haarlem 2013

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Instrumental Systematic Effects

Primary Beam



- Direction-dependence of the system response
- A model of the beam is used to correct for it
- LOFAR: "array of arrays", primary beam has two components
  - Element beam (dipole response, full-jones)
  - Array factor (beam-forming, scalar)

$$\mathbf{J}_{\text{element beam}} = \begin{pmatrix} a_{xx}e^{i\phi_{xx}} & a_{xy}e^{i\phi_{xy}} \\ a_{yx}e^{i\phi_{yx}} & a_{yy}e^{i\phi_{yy}} \end{pmatrix}$$

$$\mathbf{J}_{\text{array factor}} = ae^{i\phi}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

DP^3-Default Pre-Processing Pipeline

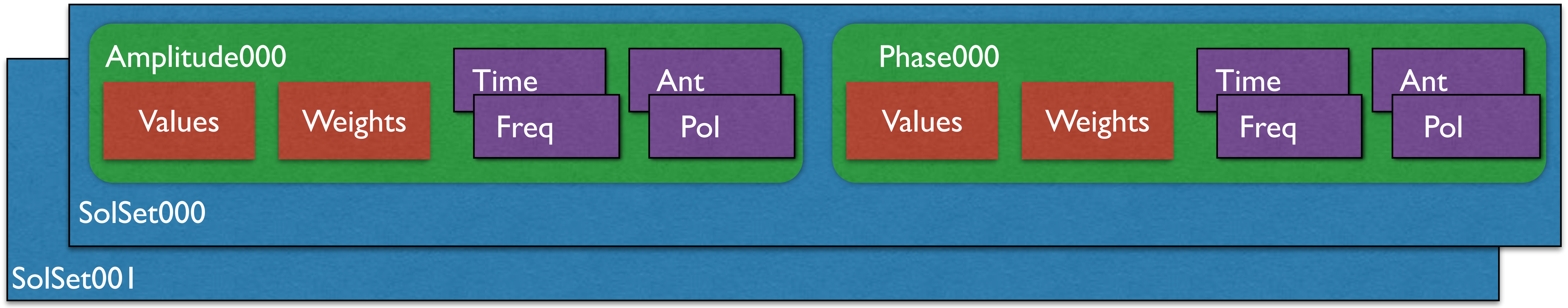
- Carry out tasks on data in a “pipelined” manner - no intermediate I/O
- Basic data operations, such as averaging, prediction, phase shifting…
- But: much more than *just* a pre-processing pipeline:
  - Features steps from raw data reading to advanced calibration algorithms!
- Not the only calibration software used for LOFAR (killMS, SAGEcal,…)
- Documentation: https://www.astron.nl/citt/DP3/index.html

LoSoTo

https://github.com/revoltek/losoto

The LOFAR Solution Tool

- Software to modify calibration solutions (extracting, inspection, smoothing…)
- Implements the H5parm data format for calibration solutions
  - HDF5 files following a defined structure



  - Solution table types: amplitude, phase, clock, TEC, rotation measure
- Command-line software in python :

```
Usage: losoto.py [-v|-V] h5parm parset [default: losoto.parset]
```

```
Ncpu = 0

#[bkp]
#operation = DUPLICATE
#soltab = sol000/phase000
#soltabOut = phaseOrig000

[align]
soltab = sol000/phase000
operation = POLALIGN
soltabOut = polalign
average = True
replace = True
fitOffset = True
minFreq = 30e6
refAnt = 'CS001HBA0'


[plotAlign]
operation = PLOT
soltab = sol000/polalign
axesInPlot = [time,freq]
axisInTable = ant
axisDiff = pol
plotFlag = True
prefix = plots-phase/ph-align_
refAnt = 'CS001HBA0'
minmax = [-3.14,+3.14]

[residual]
operation = RESIDUALS
soltab = sol000/phase000
soltabsToSub = polalign

[plotPr]
operation = PLOT
soltab = sol000/phase000
axesInPlot = [time,freq]
axisInTable = ant
axisDiff = pol
plotFlag = True
prefix = plots-phase/ph-residuals_
refAnt = 'CS001HBA0'
minmax = [-3.14,+3.14]
```

↑ example parset

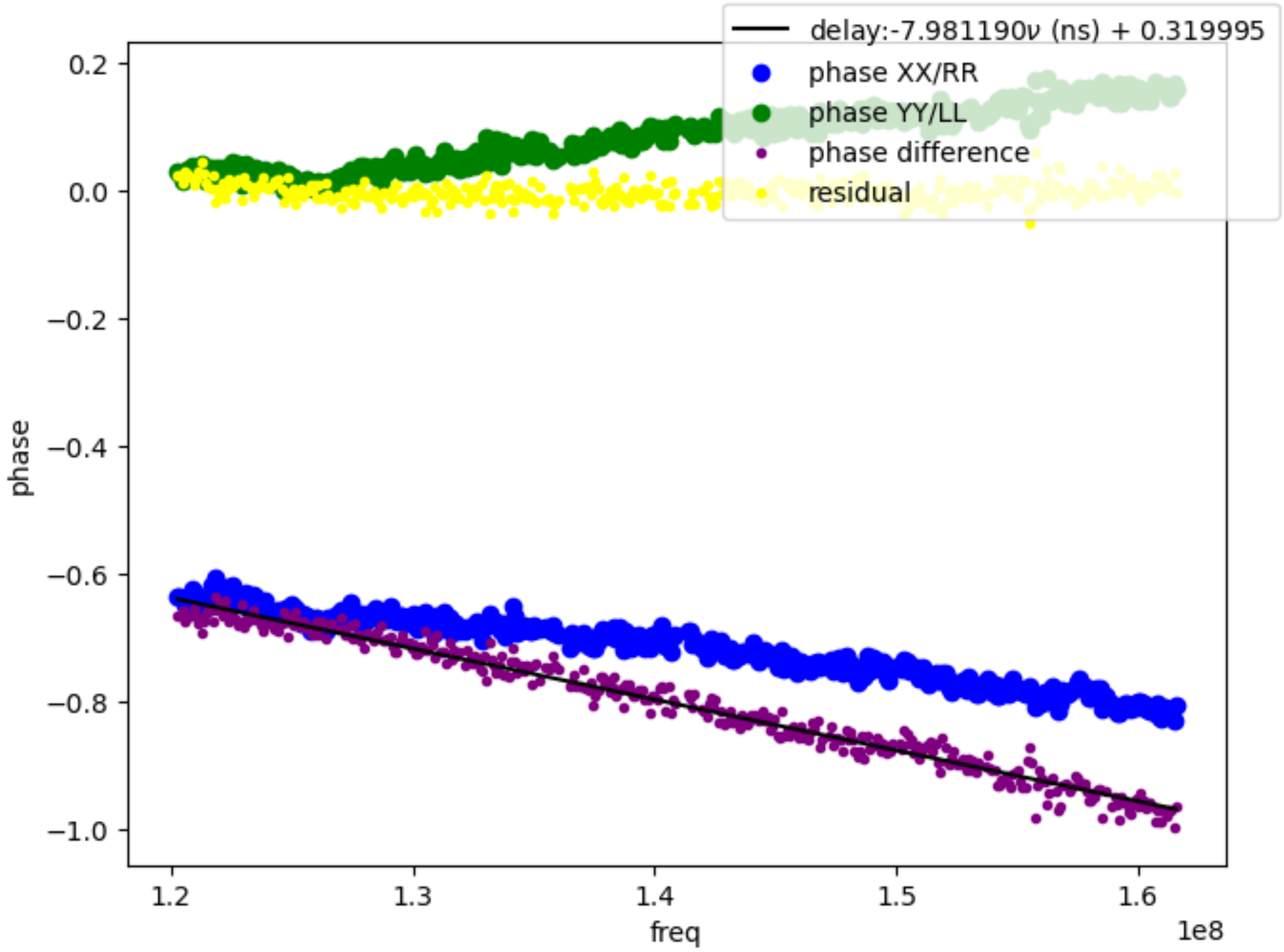
6th LOFAR Data School

LoSoTo

The POLALIGN Operation

- Extract the time-delay between the X and Y polarization from phase solutions
- Phase solutions: $n_{stations} \times n_{times} \times n_{freqs} \times n_{pol}$
- Dominant phase errors (ionospheric and clock delay) effect both polarizations equally
- Work with the phase difference: $\Delta\phi = \phi_{XX} - \phi_{YY}$
- Fit a delay term $2\pi\nu\Delta t$ to this phase difference
- Average in time
- Results in $n_{stations}$ time delays

➡much less free parameters!



Other LoSoTo Operations

- ABS: take absolute value
- CLIP: clip solutions around median
- **CLOCKTEC**: extract clock and TEC from phases
- DIRECTIONSCREEN: fit a screen to TEC and phase values
- **DUPLICATE**: create identical copy of solution table
- FARADAY: extract rotation measure from solution table
- FLAG: flag data
- FLAGEXTEND: flag data surrounded by flags
- FLAGSTATION: flag a certain station/antenna
- INTERPOLATE: regrid and interpolate data along an axis
- INTERPOLATEDIRECTIONS: spatial interpolation of solutions in multiple directions

- LOFARBEAM: fill solution table with beam response
- NORM: normalize solutions
- **PLOT**: plot solution tables
- PLOTSCREEN: screen plotting
- REFERENCE: reference values to station
- REPLICATEONAXIS: replace values along axis with a certain slice
- **RESIDUALS**: take difference of two solution tables
- REWEIGTH: change weights
- **SMOOTH**: apply various smoothing methods along axes
- TEC: estimate TEC from phase solutions

You will make use of the **boldface operations** in the tutorial

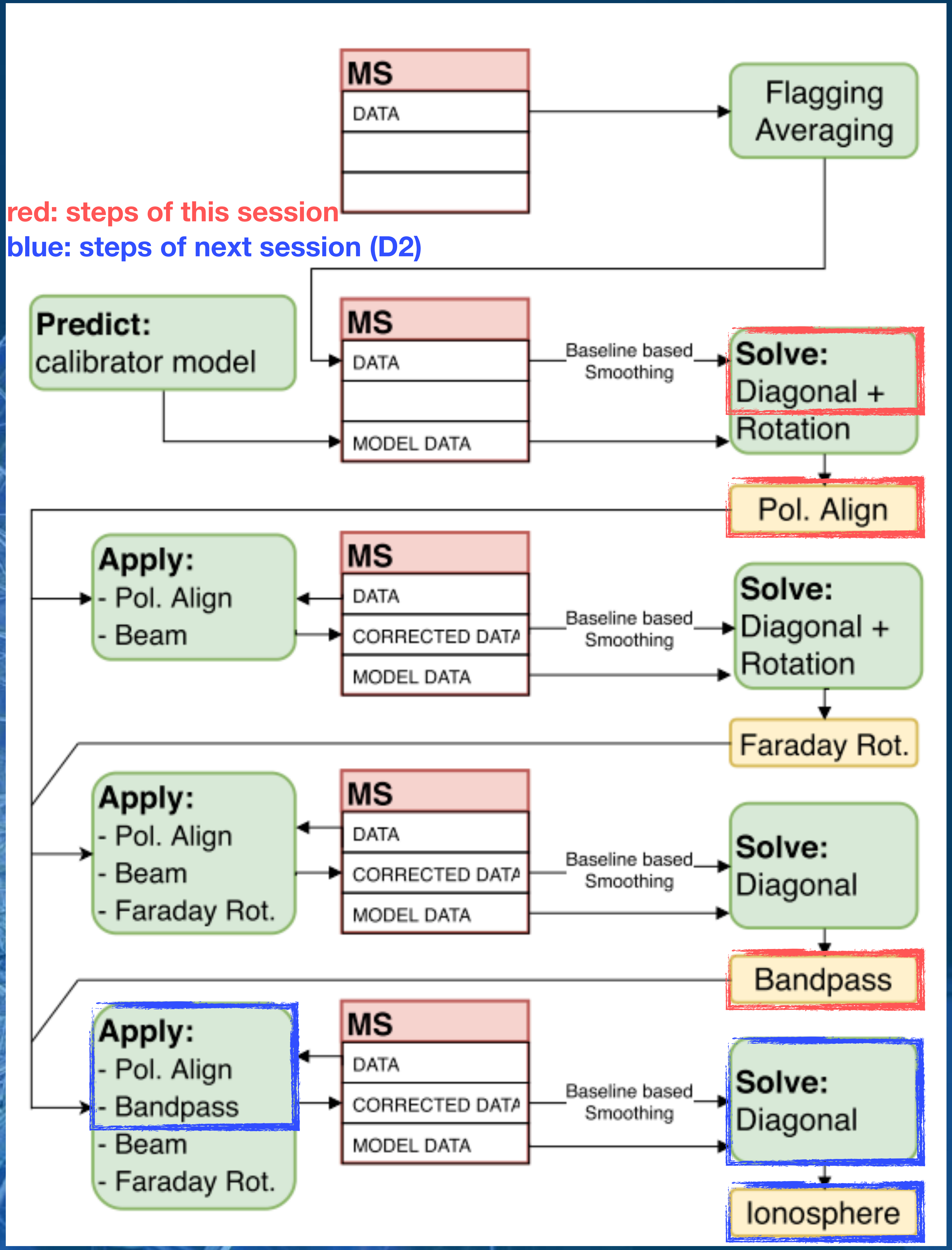
Calibration Tutorial

What we are going to do:

- Reproduce steps from the LOFAR calibrator pipeline (PreFACTOR)
- We will derive solutions for the **polarization alignment** and **bandpass** from a calibrator observation

Observation	
Target	3C295
Time	10 mins
Frequency	120-160 MHz
Time resolution	Averaged to 8s
Frequency resolution	Averaged to 0.1 MHz

Disclaimer: This example ignores a few effects (beam, differential Faraday rotation) that must be taken into account in reality!



PreFactor pipeline as described in *de Gasperin et al. 2019*

➡ see also session T2 by A. Drabent!

First Steps

- Download and unzip the tutorial material (LDS21_systematics+simulations.simg, removal_of_systematics.zip, see tutorial instructions). If this singularity image does not work on your machine, you can alternatively try lds-img.sif.
- Unpack the tutorial data (red monospaced font highlights terminal input):
  - `unzip /path/to/removal_of_systematics.zip`
- Enter the singularity (see instructions)
  - `singularity shell --bind /path/to/removal_of_systematics /path/to/LDS21_systematics+simulations.simg`
- Inside the singularity, navigate to the tutorial data
  - `cd /path/to/removal_of_systematics`
- Test the software (if you get an '`illegal instruction error`', you have to built the singularity on your machine, check the Dockerfile in the slack channel #t1-losoto)
  - `DPPP -v`
- Check the content of the folder:
  - `ls`

```
user@host:~/lofar_data_school/removal_of_systematics$ ls
3C295-simple.skymodel calibrator_3C295.MS   parsets
```

Sky model file

Measurement set (observation data)

Folder containing the parameter sets we will use

Predict the Model

- To find the solutions, we need a model of out calibrator source. This model is contained in the ASCII file `3C295-simple.skymodel`. Check the content of this file:
  - `cat 3C295-simple.skymodel`

```
# (Name, Type, Patch, Ra, Dec, I, ReferenceFrequency='150.e6', SpectralIndex, MajorAxis, MinorAxis, Orientation) = format
, , 3C295, 14:11:20.31, +52.12.10.00
3c295A, POINT, 3C295, 14:11:20.49, +52.12.10.70, 48.8815, , [-0.582, -0.298, 0.583, -0.363], 0, 0, 0
3c295B, POINT, 3C295, 14:11:20.79, +52.12.07.90, 48.8815, , [-0.582, -0.298, 0.583, -0.363], 0, 0, 0
```

- This is a simple yet sufficient model consisting of two point sources
- This model is in terms of flux density - you need to predict the corresponding visibilities
- First convert to binary format which can be read by DPPP:
  - `makesourcedb outtype="blob" format="<" in=3C295-simple.skymodel out=3C295-simple.sourcedb`
- Now you can use DPPP to fill a new data column (“MODEL_DATA”) with predicted visibilities:
  - `DPPP parsets/DPPP-predict.parset`

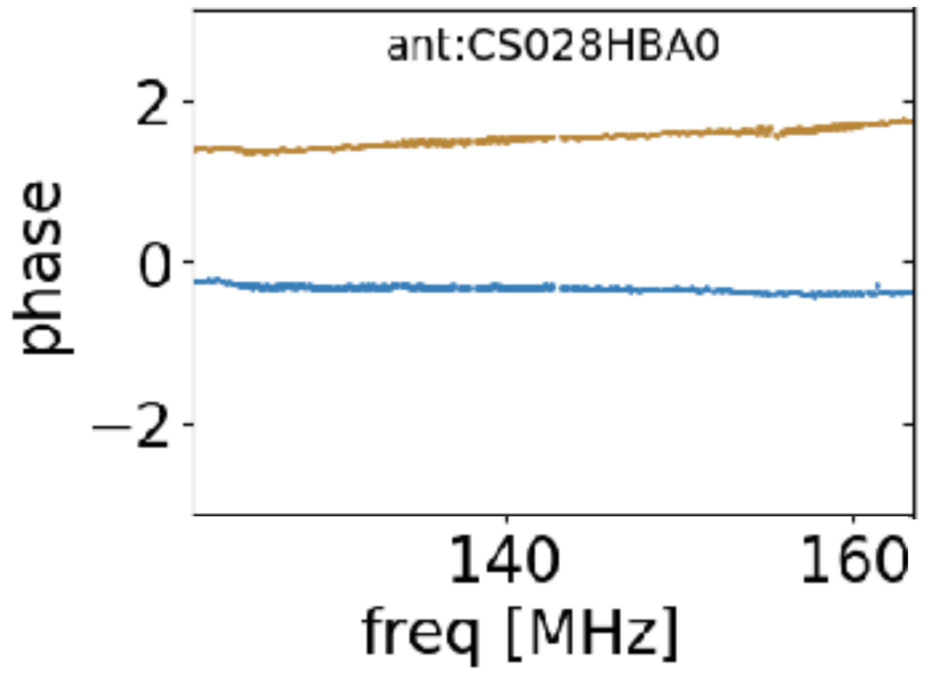
Solve against the model

- Now you can fit the Jones-matrixes to the model. We will solve for one diagonal matrix per channel/time/station $\left\| \mathbf{V}_{ij}^{\rm obs} - \mathbf{J}_i \mathbf{V}_{ij}^{\rm model} \mathbf{V}_j^{\dagger} \right\|$
- Run DP3 (this might take a few minutes, for a full data set many hours):
  - `DPPP parsets/DPPP-solve.parset`
- This created a h5parm file called `solutions.h5`. Get a quick overview of the file using LoSoTo:
  - `losoto --info -v solutions.h5`
- As you see, the file contains one solution set and two solution tables: **phase000** and **amplitude000**
- (This also created a file `solutions.h5-axes_values.txt` containing the time steps, frequency information etc. of the solutions.)

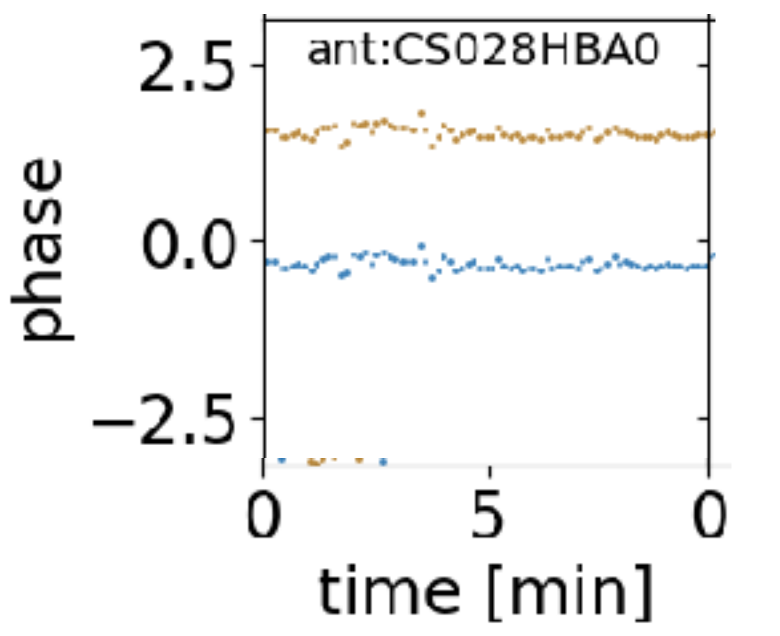
Inspect the phase solutions

- Next, plot the phase solutions with LoSoTo:
  - `losoto solutions.h5 parsets/losoto-plot-phases.parset`
- The resulting figures are in the `plots-phase` folder. Open and inspect the plots (you probably want to open them outside of the singularity)
- Lots of panels! Each panel shows phase solutions vs time and/or frequency for one station

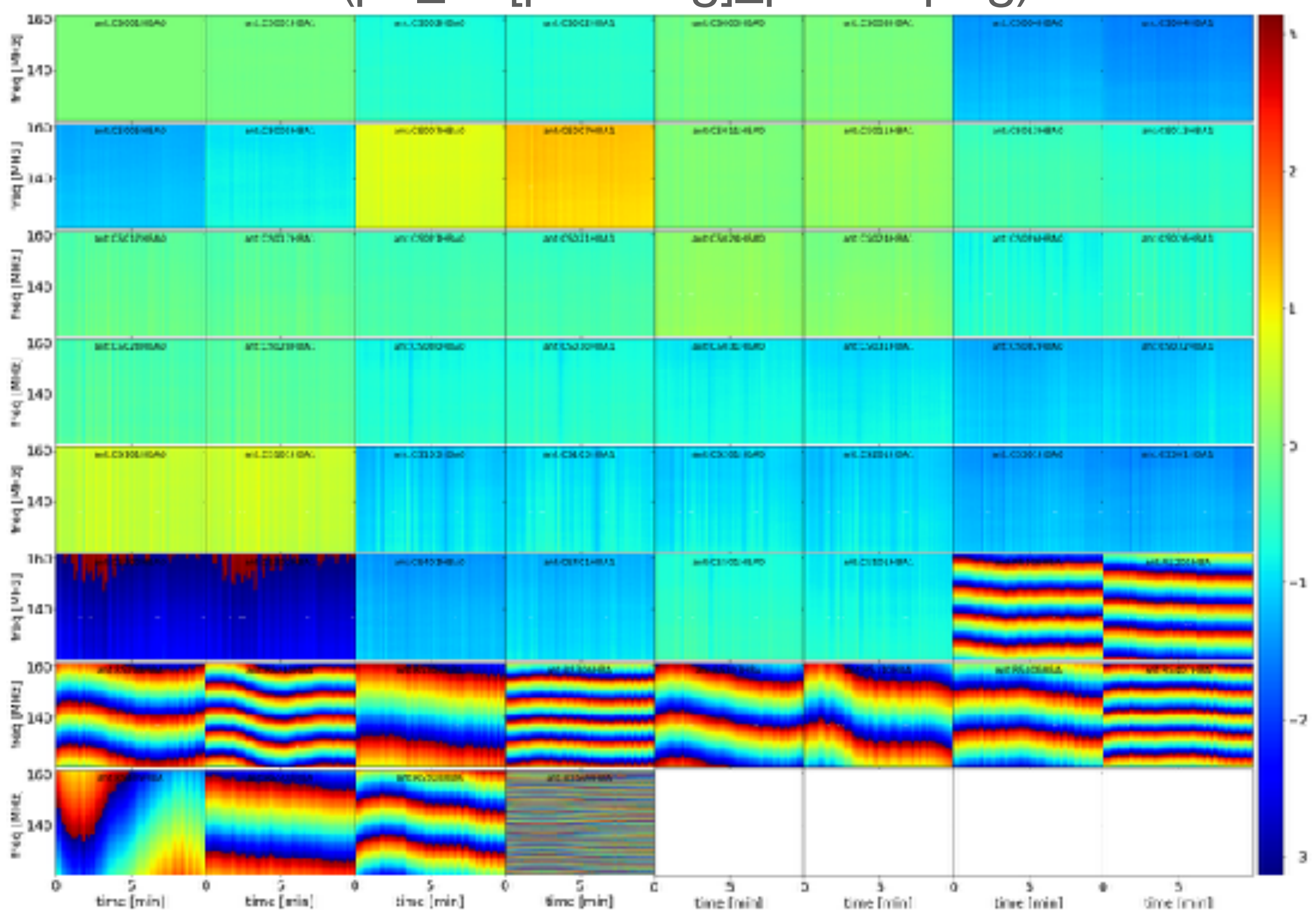
Pol. Offset **not** constant in freq.



Pol. Offset constant in time!

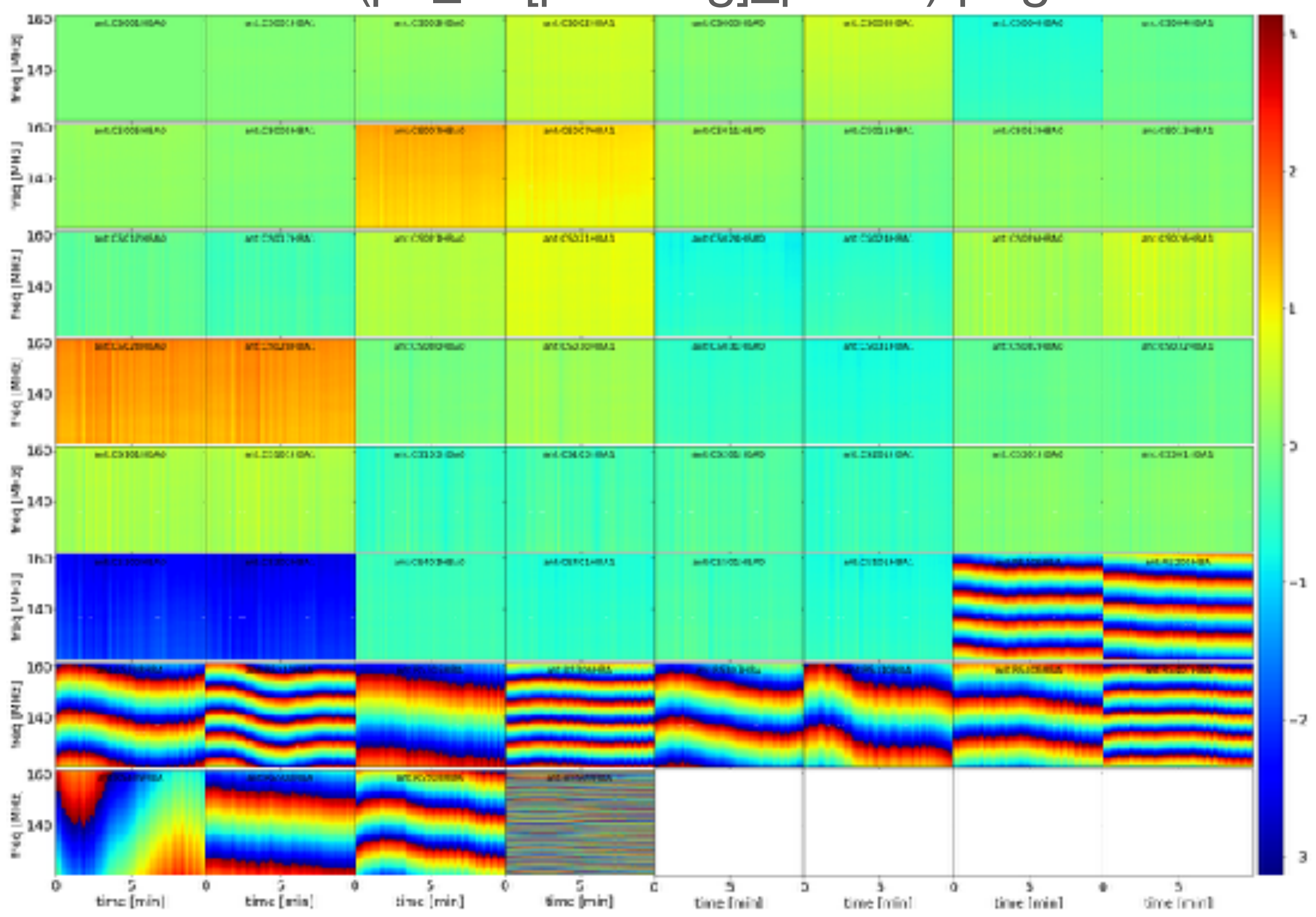


Pol. X (ph_dir[pointing]_polXX.png)



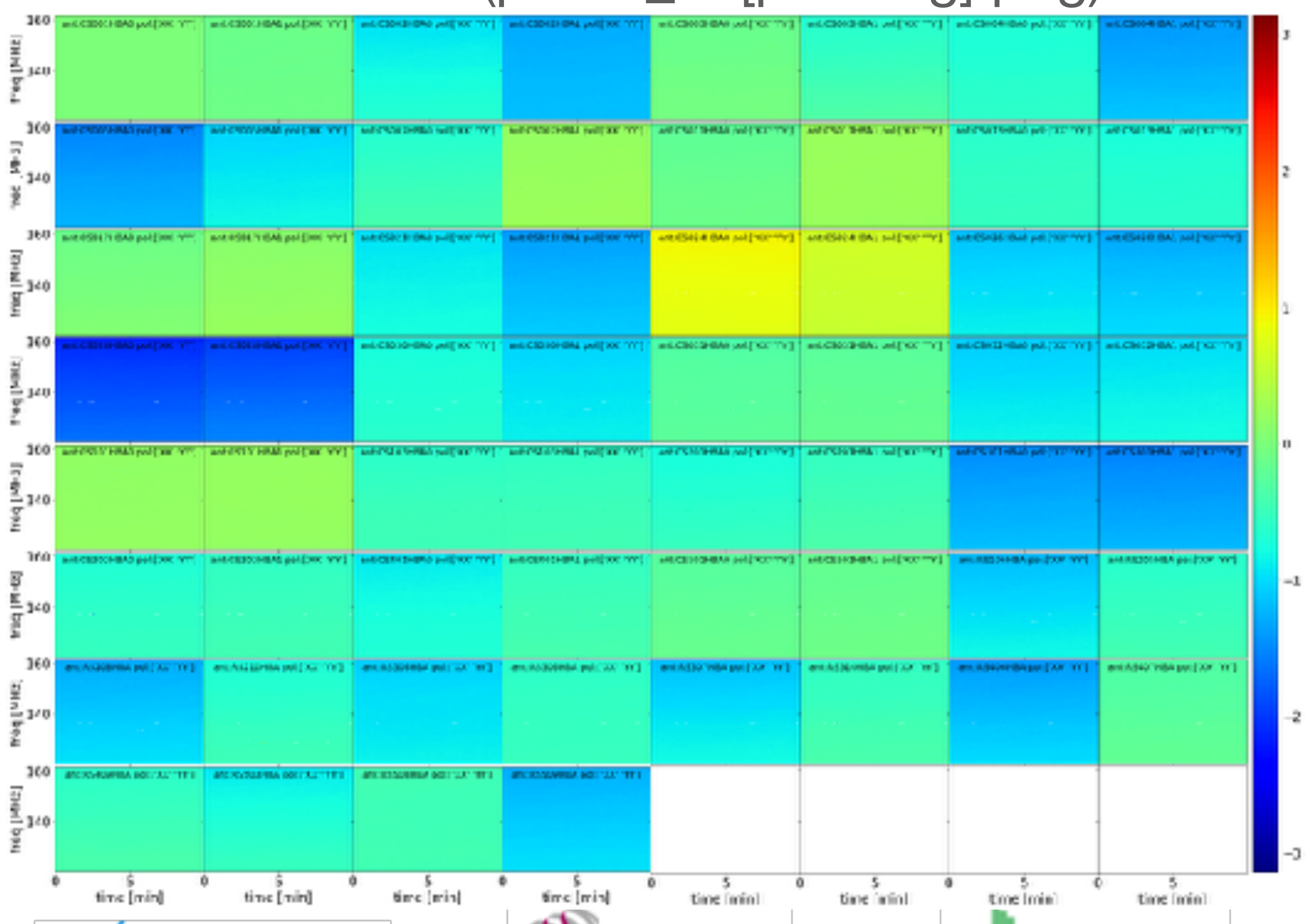
\-

Pol. Y (ph_dir[pointing]_polYY).png



=

Difference (ph-dif_dir[pointing].png)



ASTRON LOFAR STELLAR

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

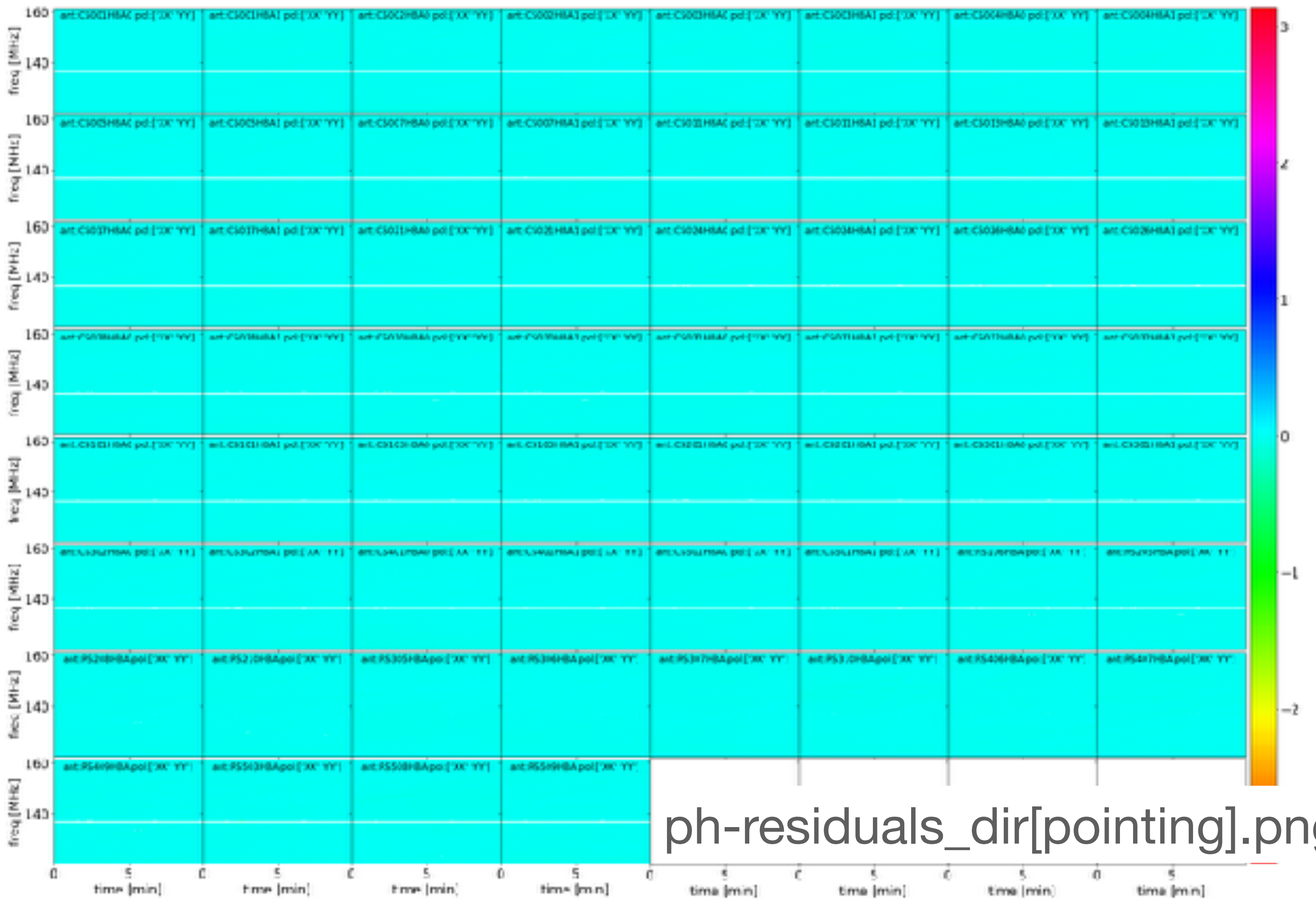
Polarization alignment

- Now, we run the polalign operation to fit the time-delays between the X and Y polarisation for each station to the phase solutions stored in phase000. The parset will also plot us the residuals.

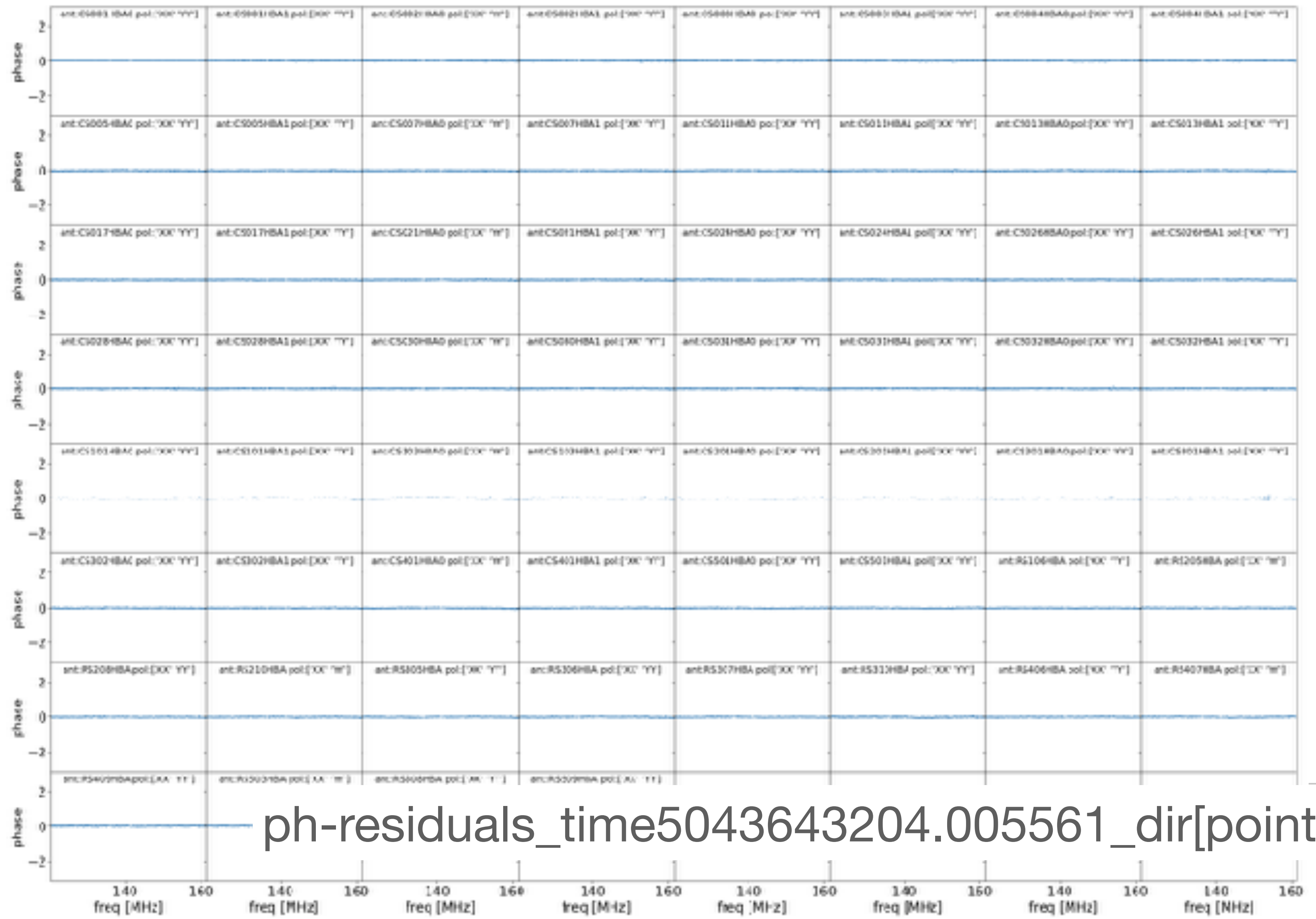
  - `losoto solutions.h5 parsets/losoto-polalign.parset`

- This step will also log the time delay found for each antenna, you will see that they are order of 1 ns

- Look at the residual plots, this is the difference of the phase solutions and the phases corresponding to the fitted delays



ph-residuals_dir[pointing].png

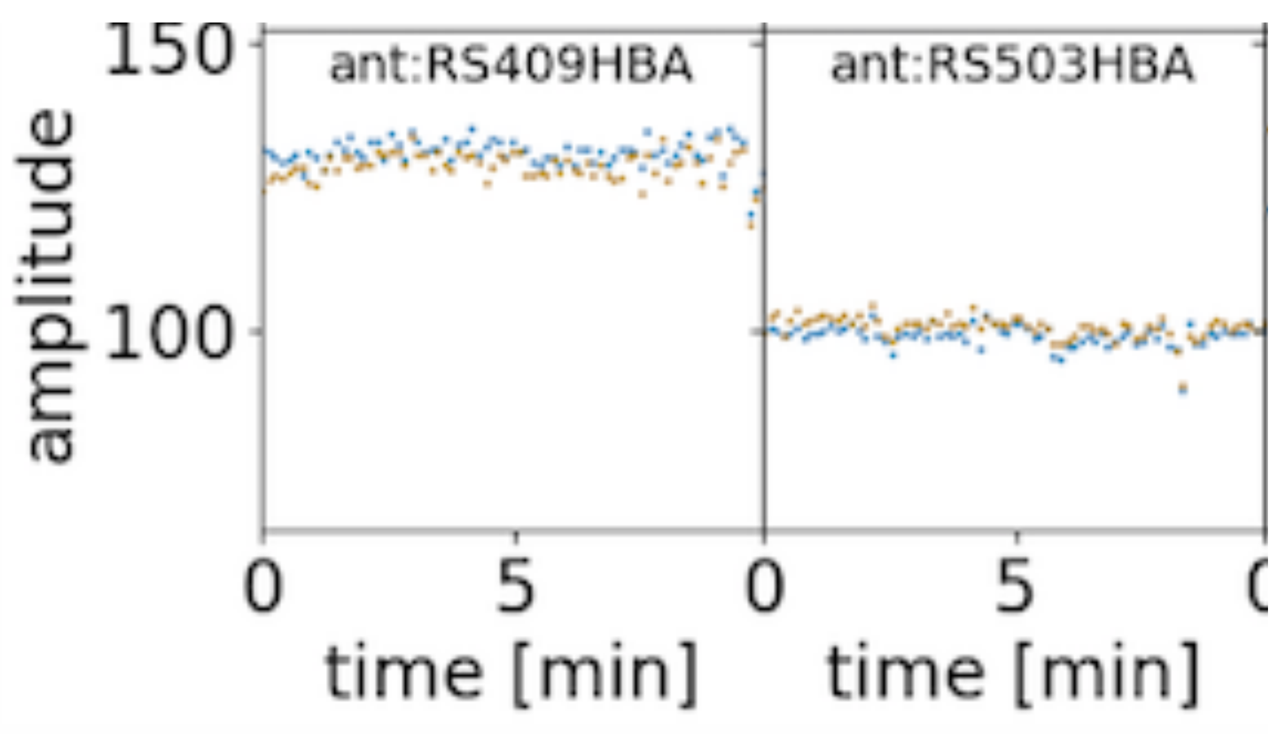


ph-residuals_time5043643204.005561_dir[pointing].png

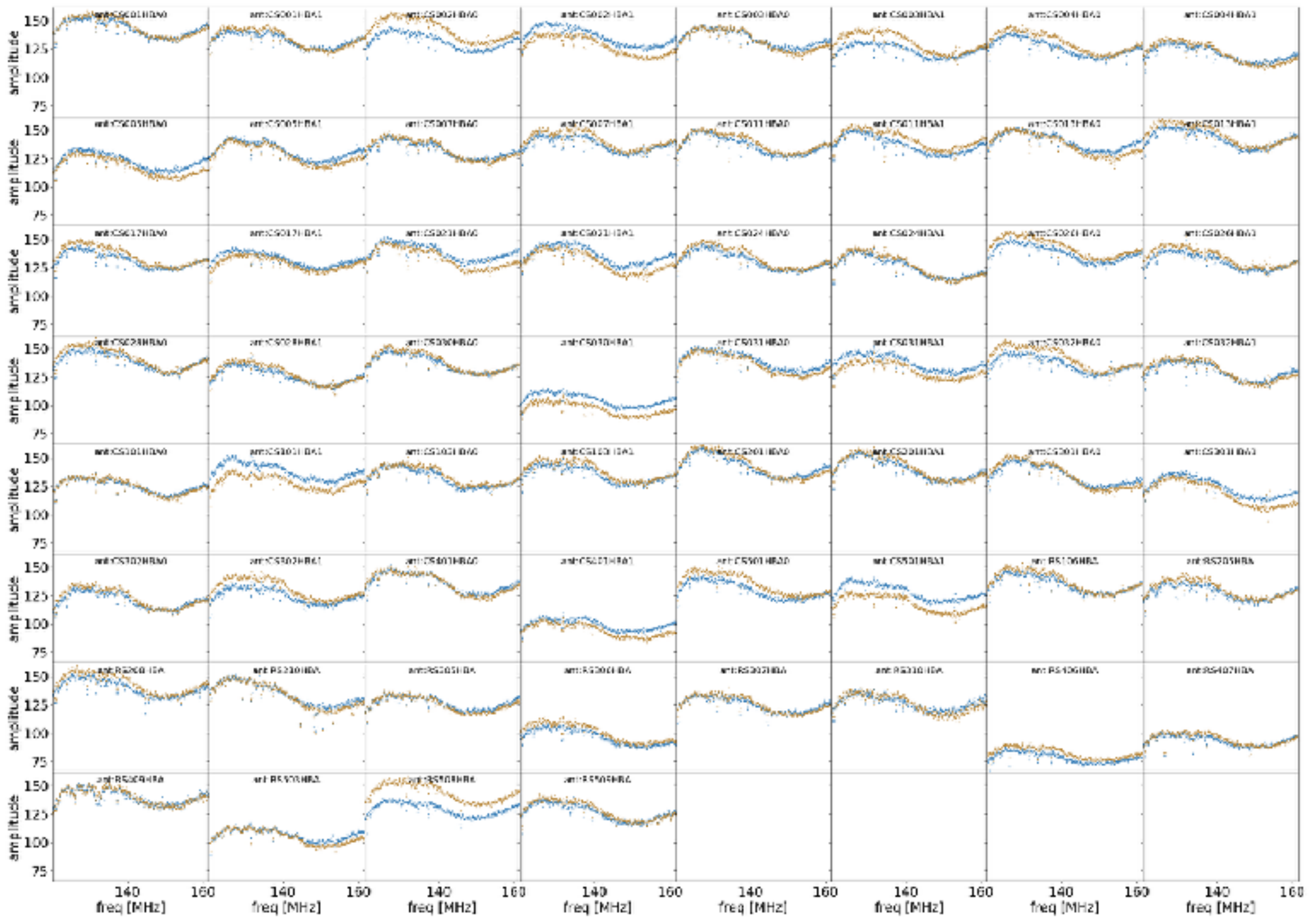
Question: How many parameters do we have in the phase solutions? To how many parameters did we reduce this with the polalign operation? (Use the `losoto --info` option)

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

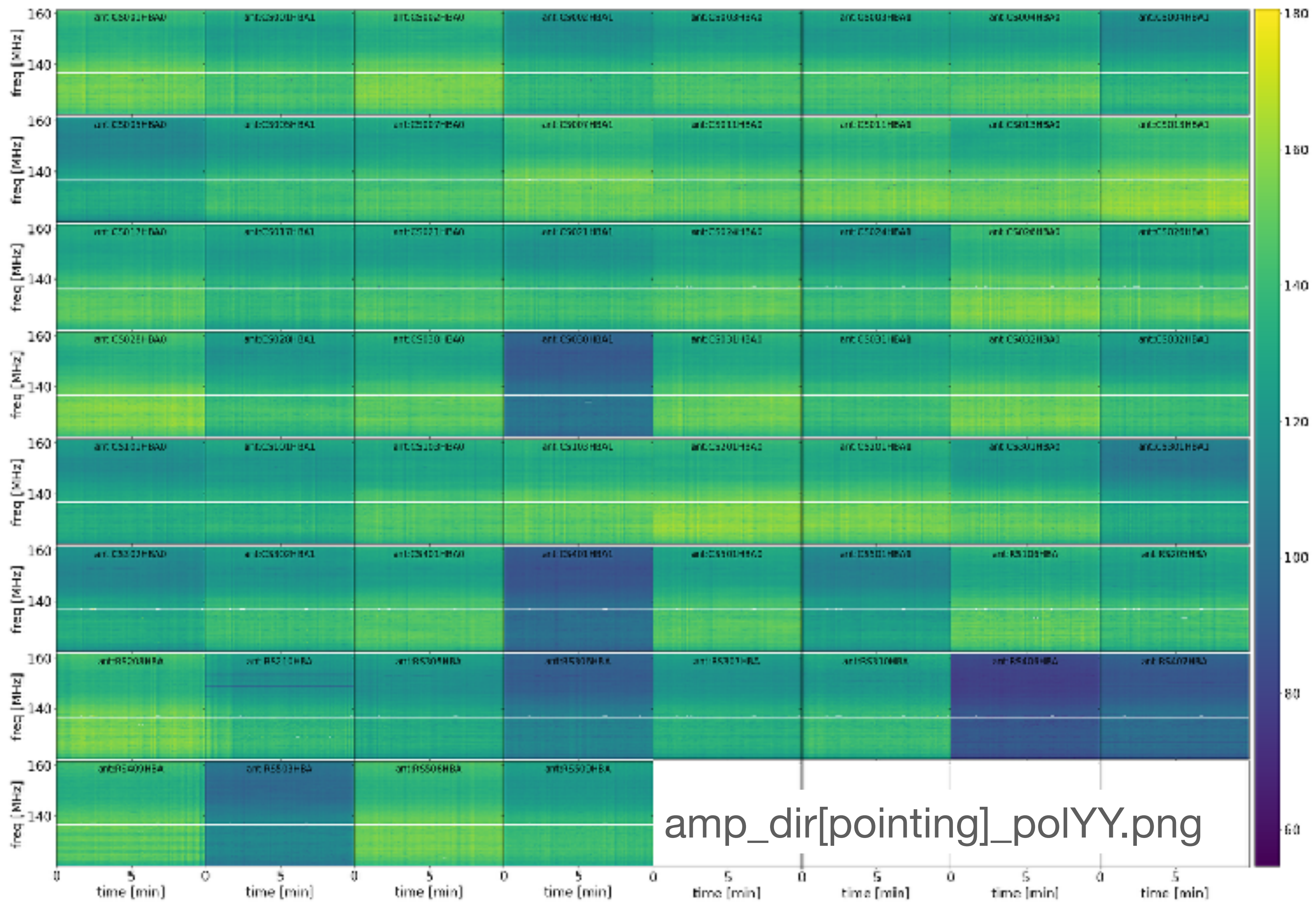
The amplitude solutions



- Now let's plot the solutions stored in amplitude000
  - `losoto solutions.h5 parsets/losoto-plot-amplitudes.parset`
- The resulting plots are in plots-amplitude/
- ~constant in time, characteristic shape in frequency



amp_time5043643204.005561_dir[pointing].png

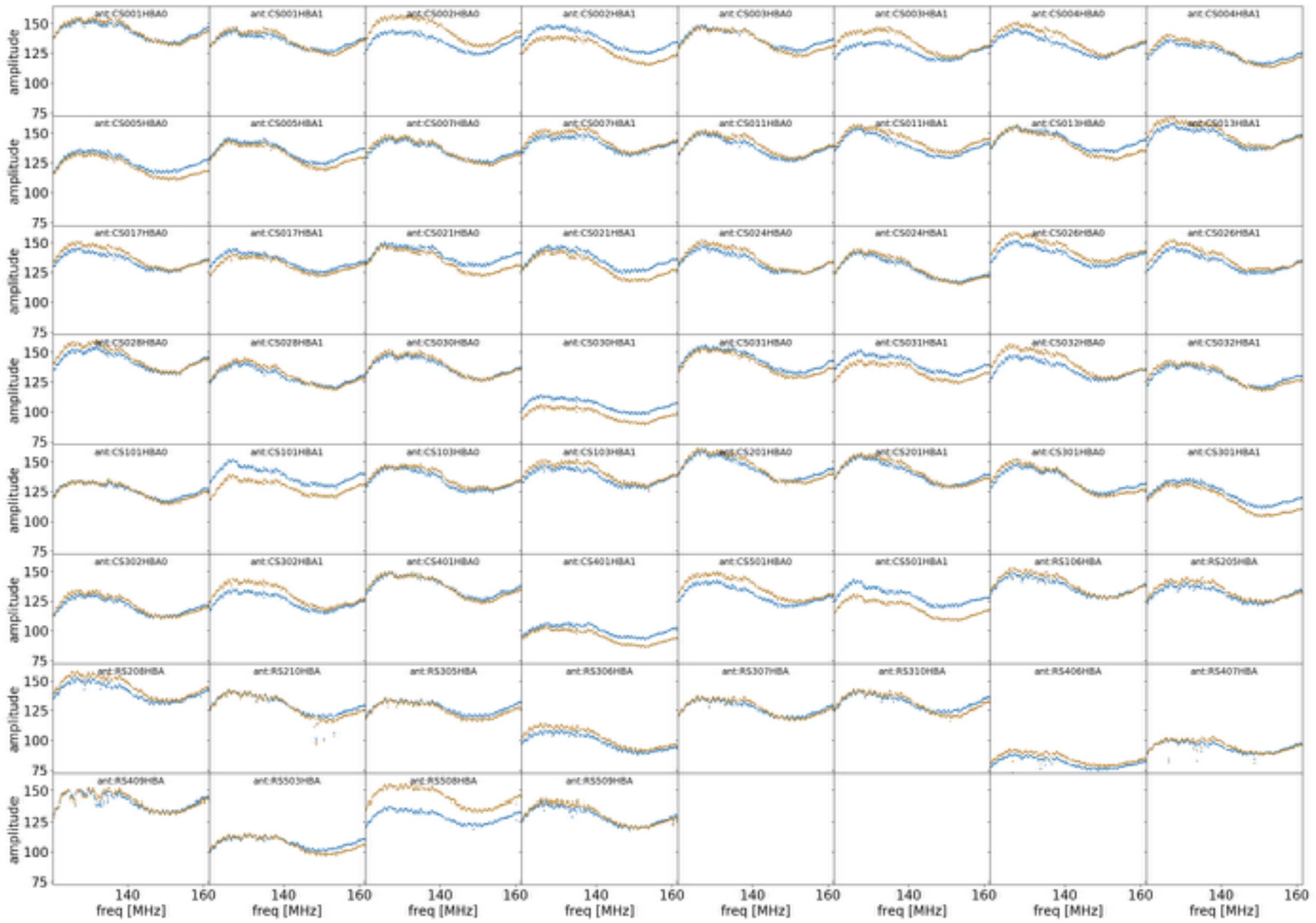


amp_dir[pointing]_polYY.png

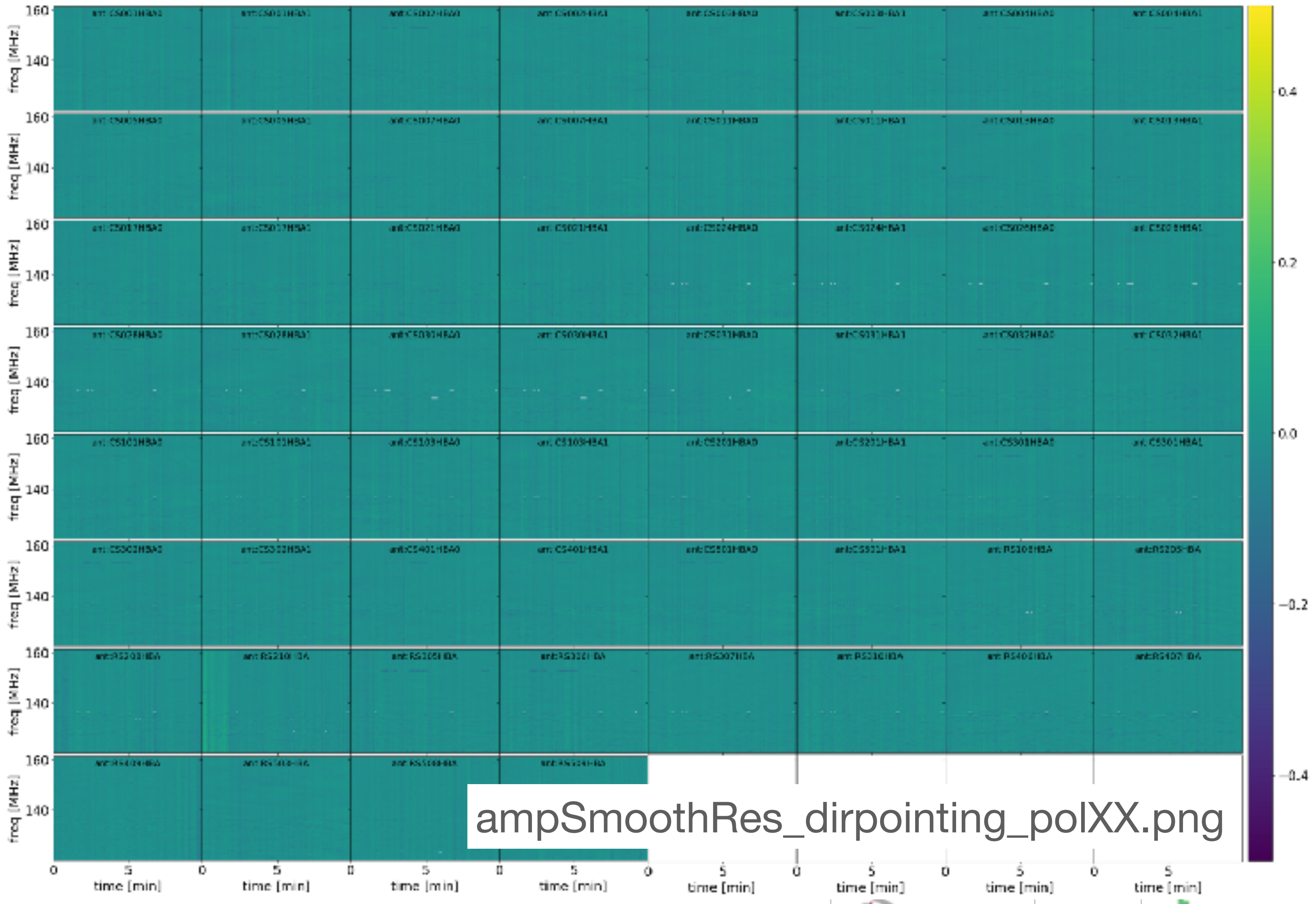
Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Extract the bandpass

- The amplitude solutions are noisy, so we take the mean in time for each station and each polarization using the SMOOTH operation.
  - `losoto solutions.h5 parsets/losoto-bandpass.parset`
- Check the diagnostic plots



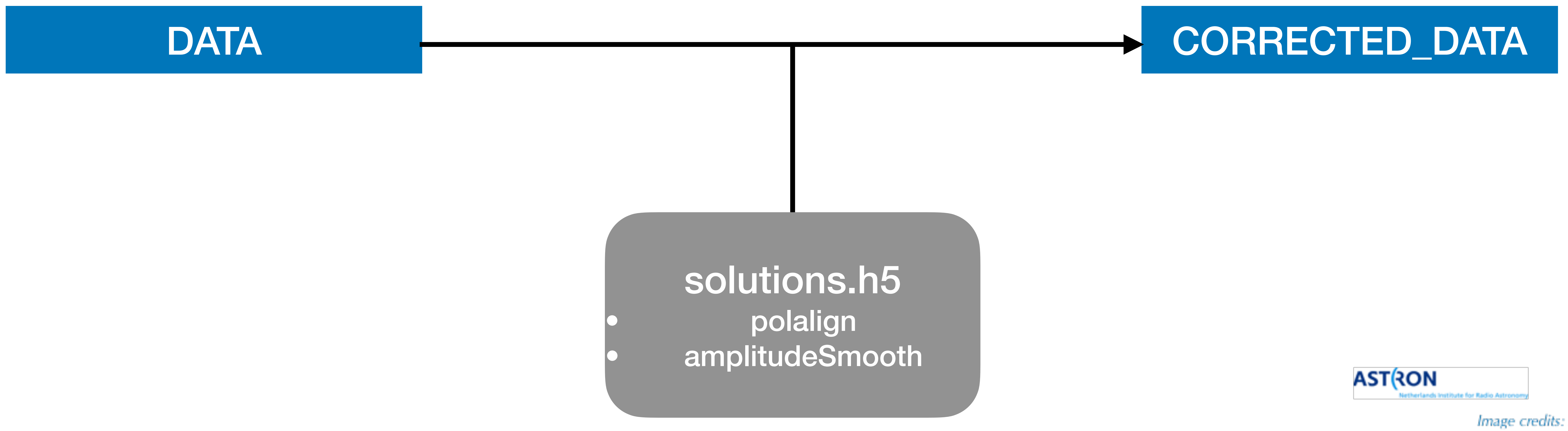
ampSmoothRes_time5043643204.005561_dirpointing.png



ampSmoothRes_dirpointing_polXX.png

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

- **If you are interested and have time left, you can further work on this data to extract ionospheric (and clock) systematics, this procedure is explained in the session D2 by Maaijke Mevius**
- You will need the folder `parsets-ct` which contain additional parameter sets -> see Slack
- Place the folder `parsets-ct` in `removal_of_systematics`
- Now apply the corrections you derived to the data in preparation for further calibration steps.
- This creates a new column with the name `CORRECTED_DATA`:
  - `DPPP parsets-ct/DPPP-apply.parset`

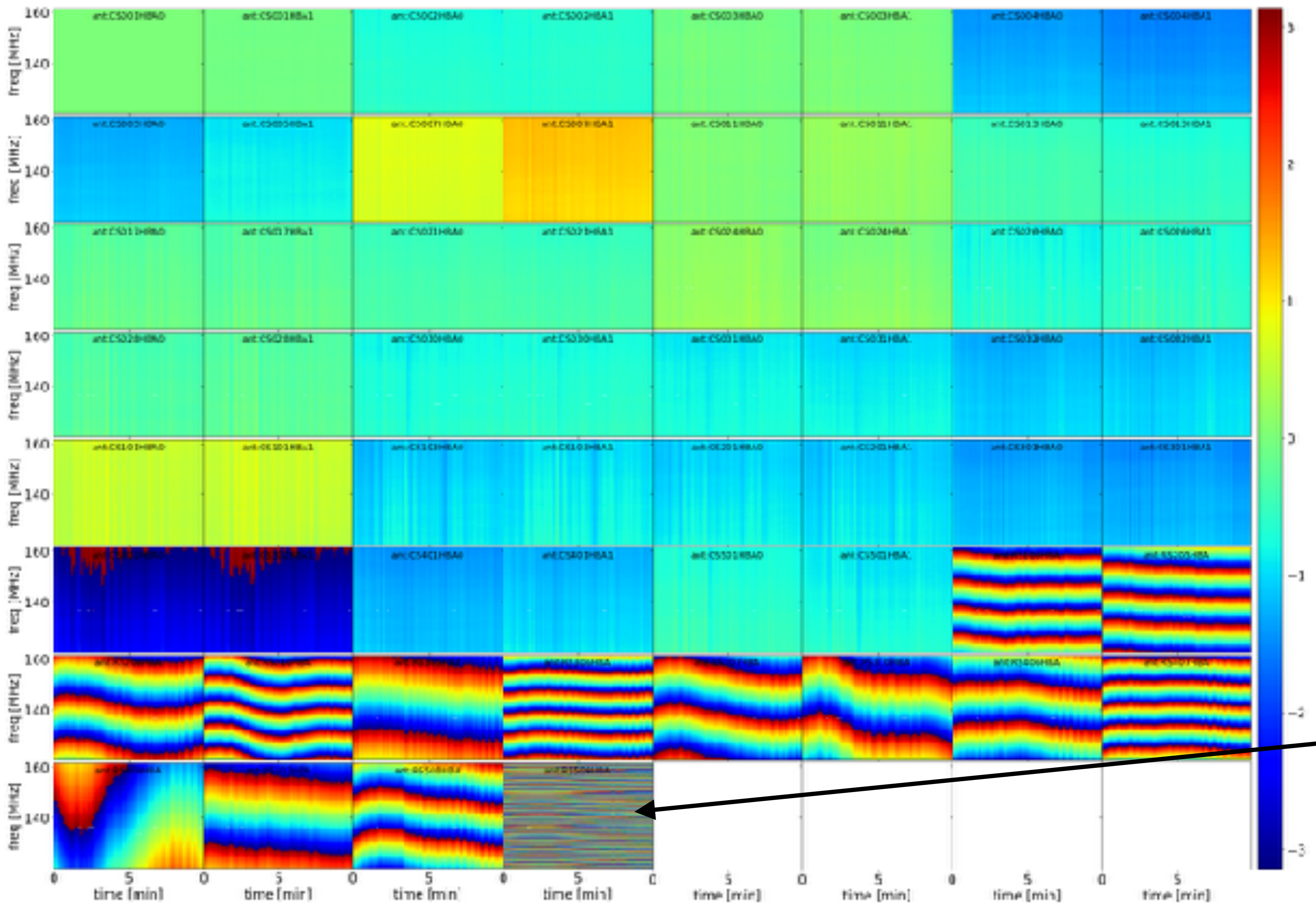


Another solution step

These steps accompany session (D2) by M. Mevius

- We will solve for scalar phases, and then simultaneously fit the clock delay as well as the first order ionospheric corruption to the solutions
  - `DPPP parsets/DPPP-solve.parset msin.datacolumn=CORRECTED_DATA sol.mode=phaseonly sol.h5parm=solutions-clocktec.h5`
- plot the solutions:
  - `losoto solutions-clocktec.h5 parsets-ct/losoto-plot-phases-clocktec.parset`
- The resulting plots are in `plots-clocktec/`

Command-line arguments overwrite parameter set arguments



(This is likely a very strong clock delay)

ph_dirpointing_polXX.png

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Clock-TEC-separation

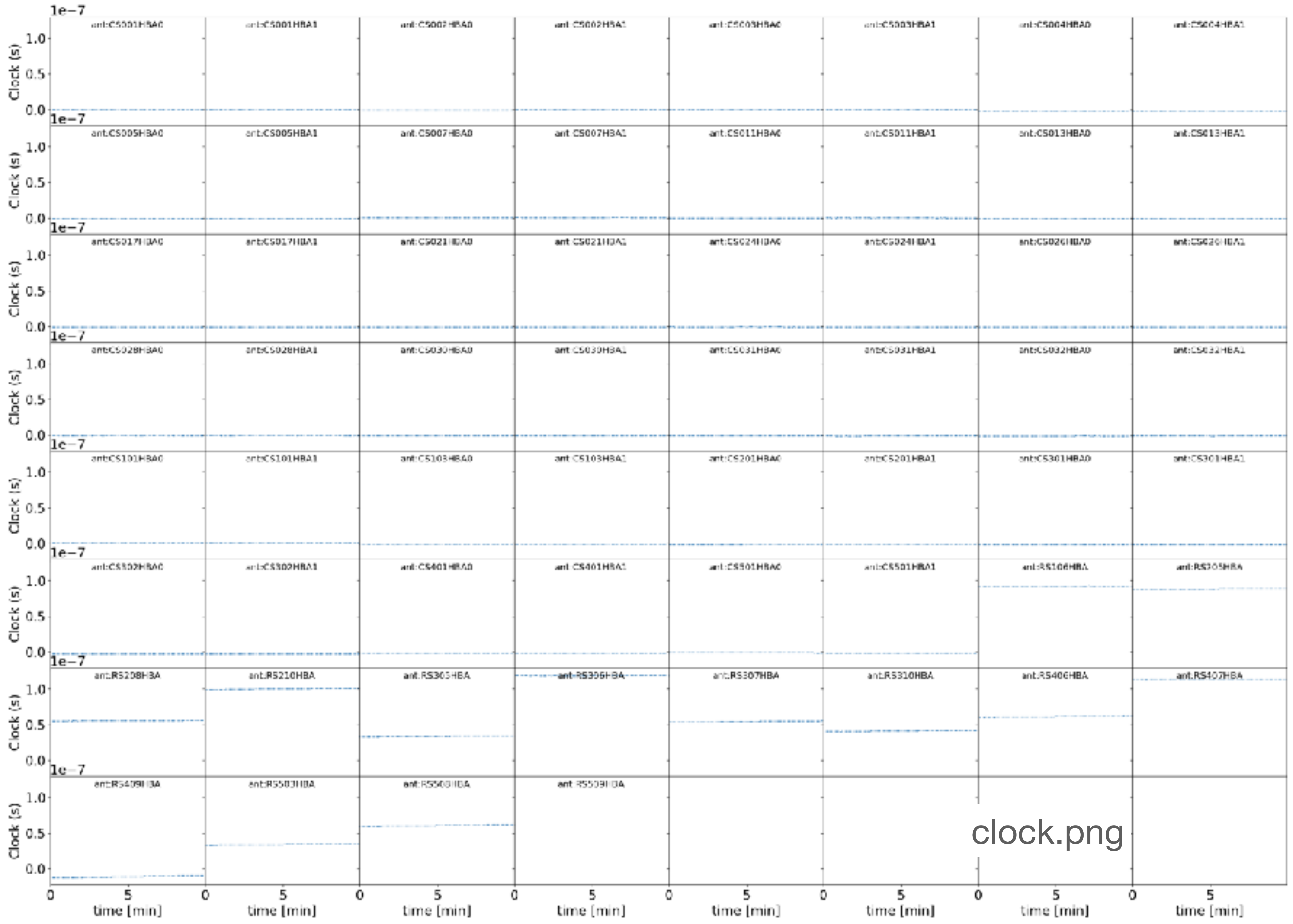
These steps accompany session (D2) by M. Mevius

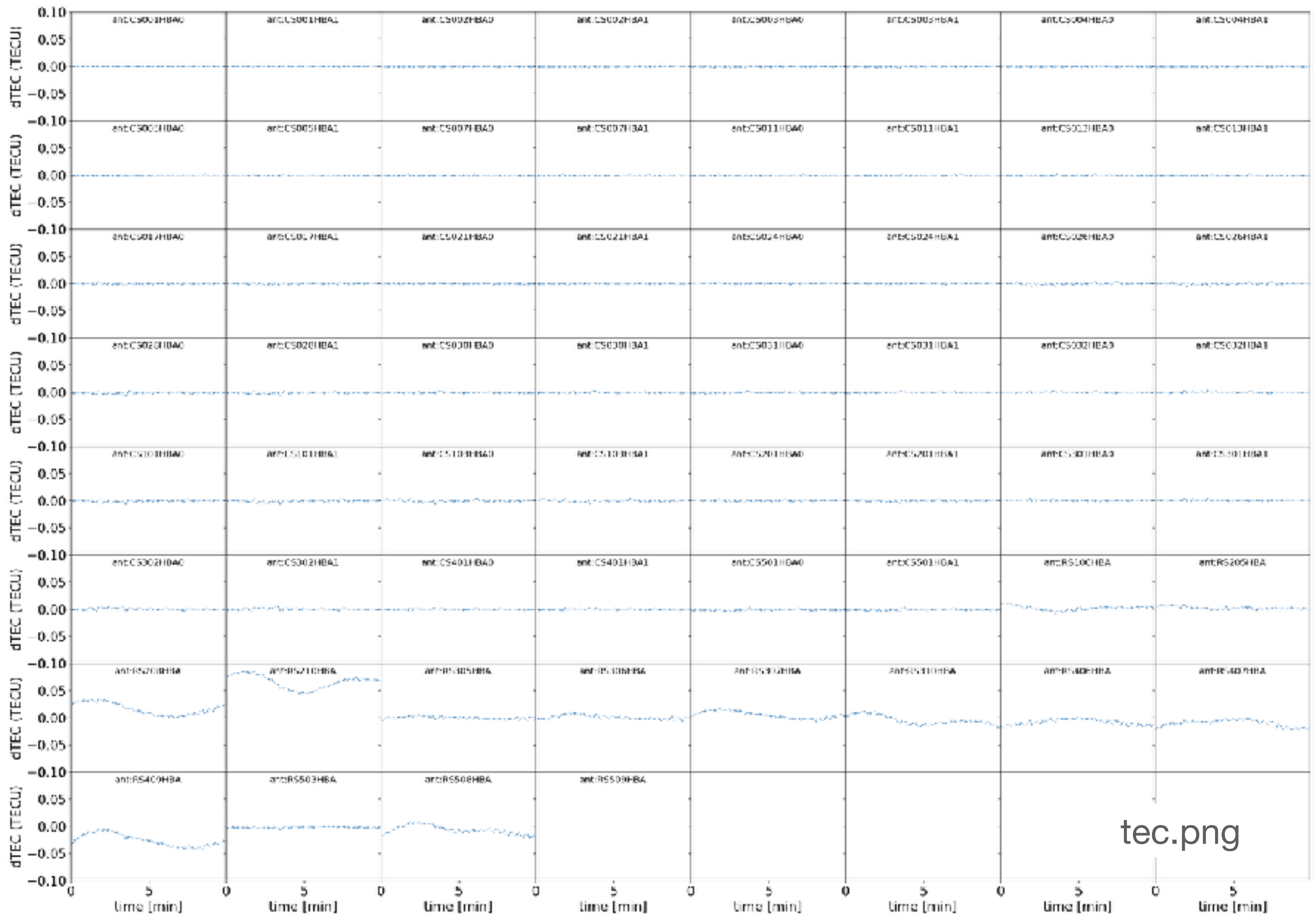
- Clock and first order ionospheric phase errors have different characteristic frequency dependence: $\Delta\phi_{\text{clock}} \propto \Delta t \nu$ $\Delta\phi_{\text{TEC}} \propto \Delta\text{TEC}\nu^{-1}$
- This can be exploited to extract underlying physical effects from phases:
  - `losoto solutions-clocktec.h5 parsets-ct/losoto-clocktec.parset`

Clock/TEC-separation results

These steps accompany session (D2) by M. Mevius

- The last operation also produced plots of the delays, TEC and residuals in `plots-clocktec/`





(clock is time-variable, but on longer time scales)

Image credits: A. Corstanje, F. Sweijen, C. Van Eck, P. Zucca

Final remarks

- Follow the tutorial steps during the hands-on session later on
- If you have questions to the tutorial, you can ask them in the slack channel
- If you want to learn a bit more, you can take a look at the parameter sets in the `parsets/` folder and compare the settings to the documentations at:
    - https://www.astron.nl/citt/DP3/index.html
    - http://revoltek.github.io/losoto/losoto.html