Durham University

# LOFARm to table

## A broad overview from the antennas to images

Frits Sweijen, LOFAR Data School 2026, ASTRON

# <obligatory standing between you and lunch joke>

# Contents


## Telescope to archive

Correlation
Preprocessing

## Archive to image - general concept

Instrumental corrections
Direction-independent calibration
Direction-dependent calibration

# More in-depth imaging talks this week

**Systematic effects in LOFAR - Cristina Cordun**

**LOFAR self-calibration - Christian Groeneveld**

**Quality assessment of LOFAR data - Aleksandar Shulevski**

**Widefield imaging - Andre Offringa**

**VLBI with LOFAR - me + tutorial by Roland Timmerman**

**LBA with LiLF - Jort Boxelaar**

# The key data concepts

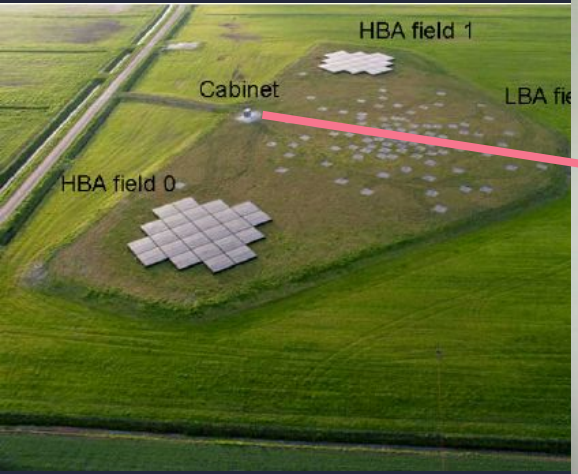
**Correlation:** creating the actual measurements, i.e. visibilities, that we work with

**Calibration:** correct visibilities for systematics, ionospheric and other corruptions

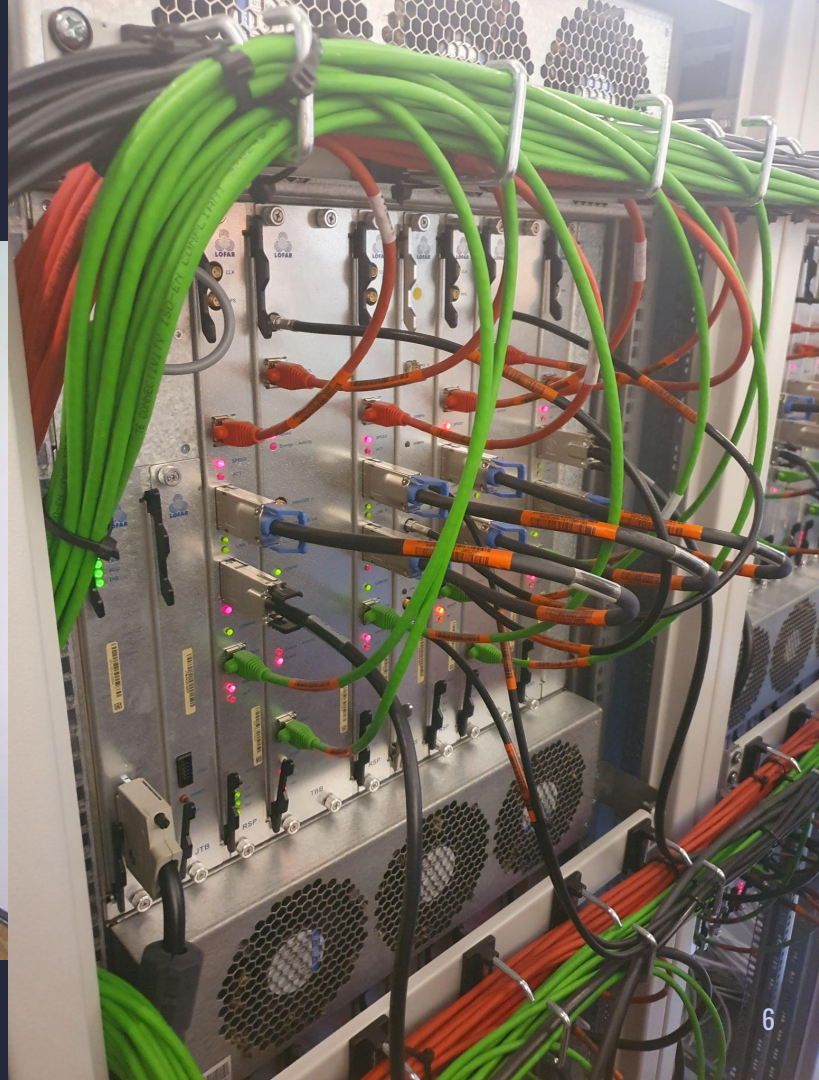
**Imaging:** transform data from the visibility domain to sky image domain

**Editing:** flagging, removing altering etc. of data

# Station cabinet





# Station cabinet

**Analogue-to-digital conversion (ADC)**

**Station beamforming**
N.B. for HBA there is analogue beamforming for the individual tiles

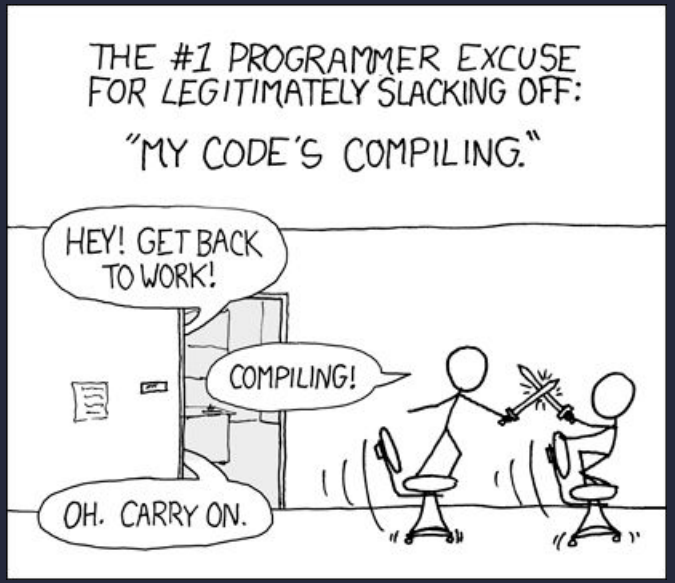
**Transient buffer board (TBB)**
Short 5s raw data buffer at 5 ns resolution for transient events
1.5 TB/s for a 24-tile core station if 16 bit voltages

**Subband formation**
195.3125 kHz wide with 767.939 Hz channels
5.12 μs time sampling (200 MHz clock)

# Imaging data flow




1. Get data
2. Follow these six easy steps
3. ???
4. profit




XKCD 303

# Correlation

**Voltage data real time from stations to correlator in Groningen**

3 Gb/s per station x 54 stations = 162 Gb/s data rate (approx. due to varying station sizes)

**Cross-correlation and narrow-band RFI excision**

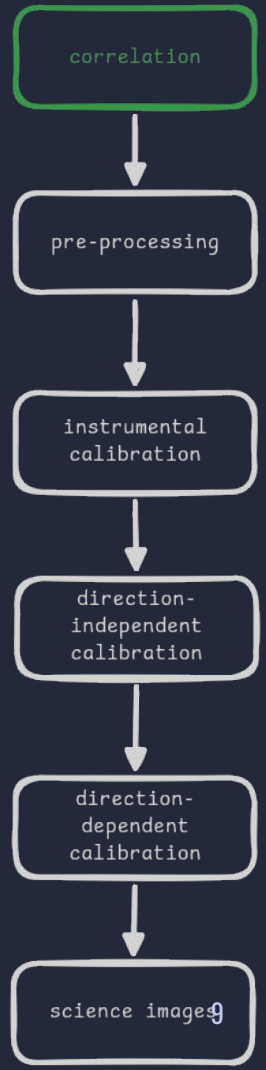
195 kHz wide subbands with 0.8 Hz channels for high precision RFI excision

**Initial data averaging**

3 kHz frequency resolution

1 s time resolution

~16 TB for an 8 hour HBA observation (assuming Dysco compression)



# Correlation

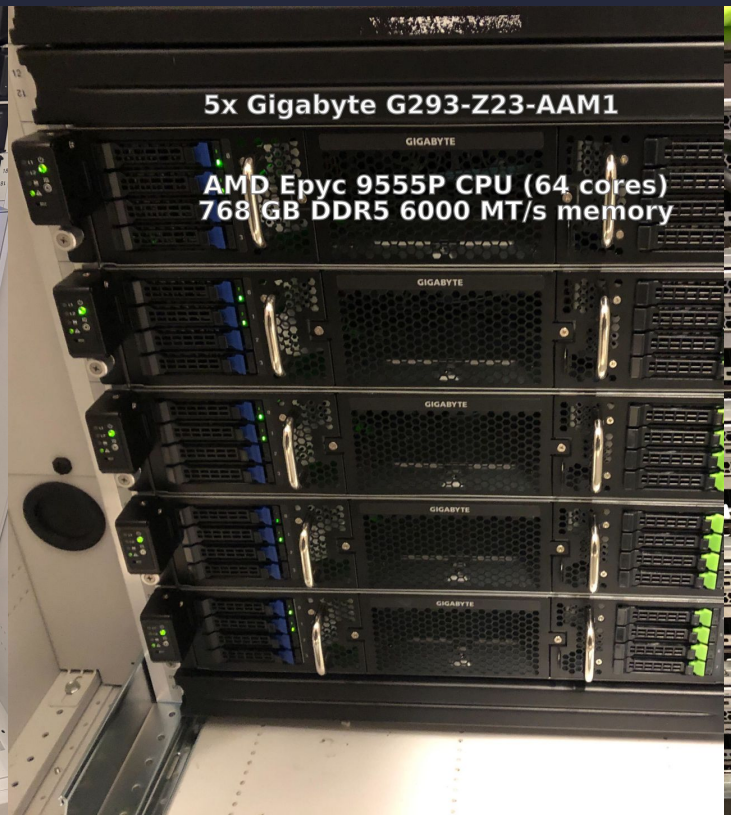





# Correlation

COBALT (COrrelator and BeAmformer for the LOFAR Telescope; see e.g. Broekema+2018 for details)






# Correlation

$$V_{ij} = A\, e^{i\varphi}$$

# Correlation

Visibility on baseline i-j

$$V_{ij} = A\, e^{i\varphi}$$

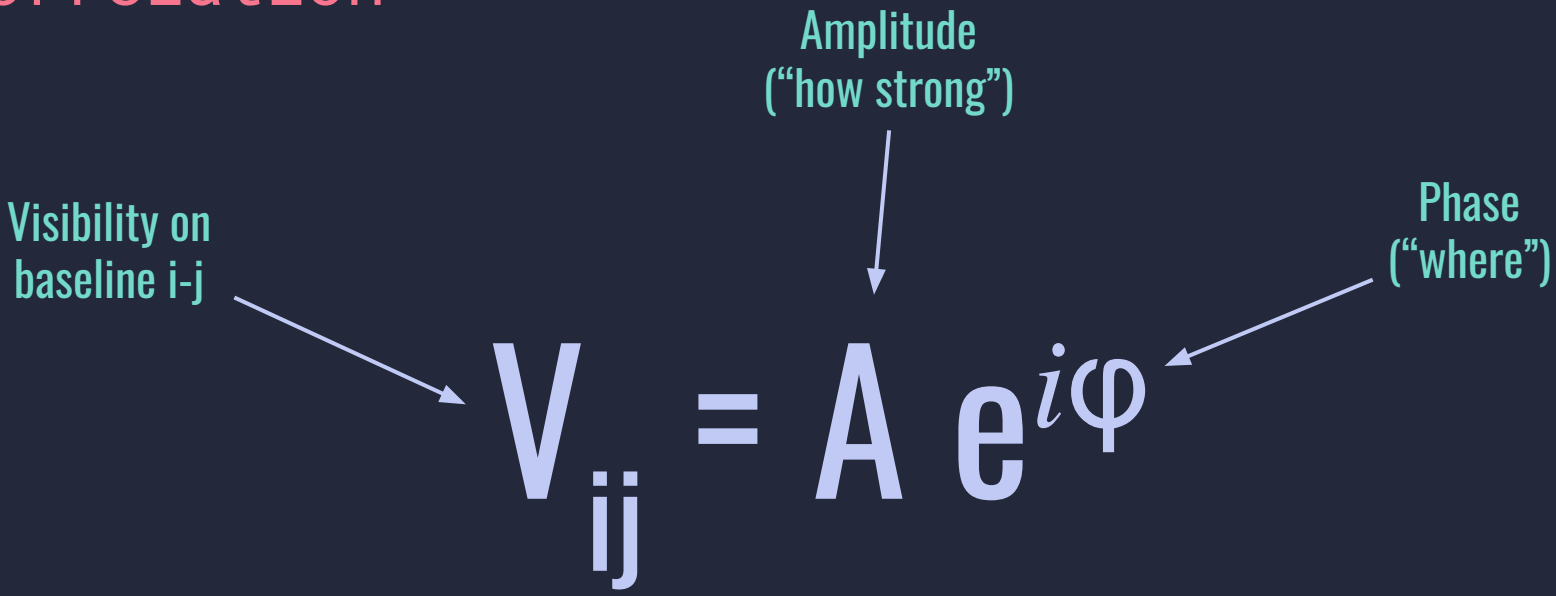
# Correlation

Amplitude
("how strong")

Visibility on
baseline i-j

$$V_{ij} = A\, e^{i\varphi}$$

# Correlation



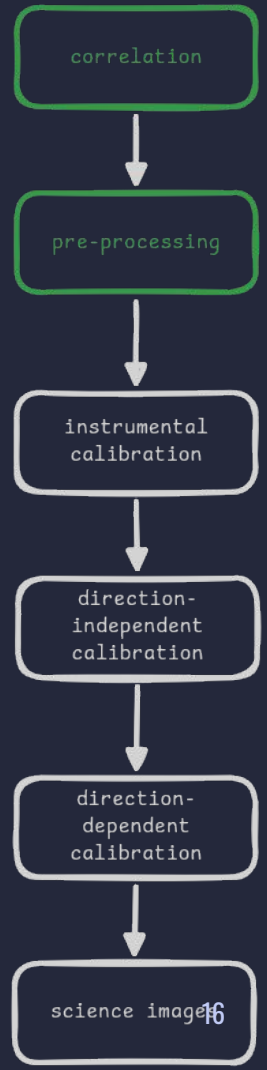
# Preprocessing

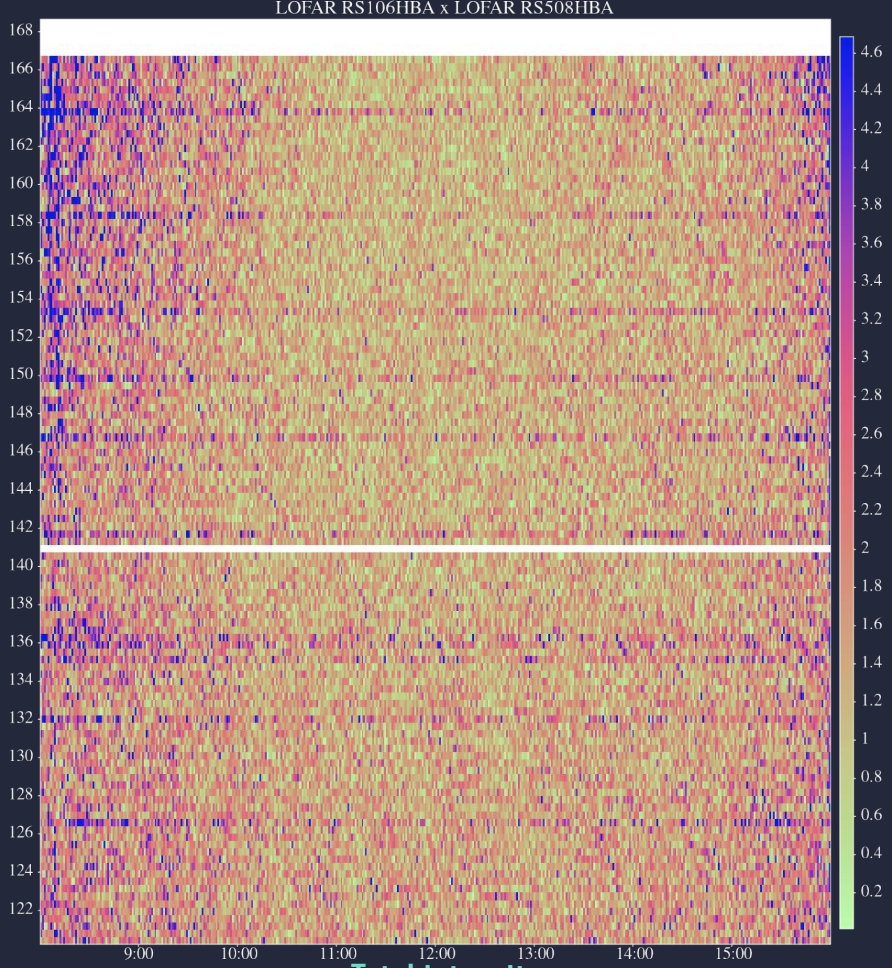
**Flagging of bad data**
RFI, low amplitudes, low elevation

**Final data averaging**
1 s, 12 kHz channels for typical HBA data
4-5 TB for a typical 8-hour scan

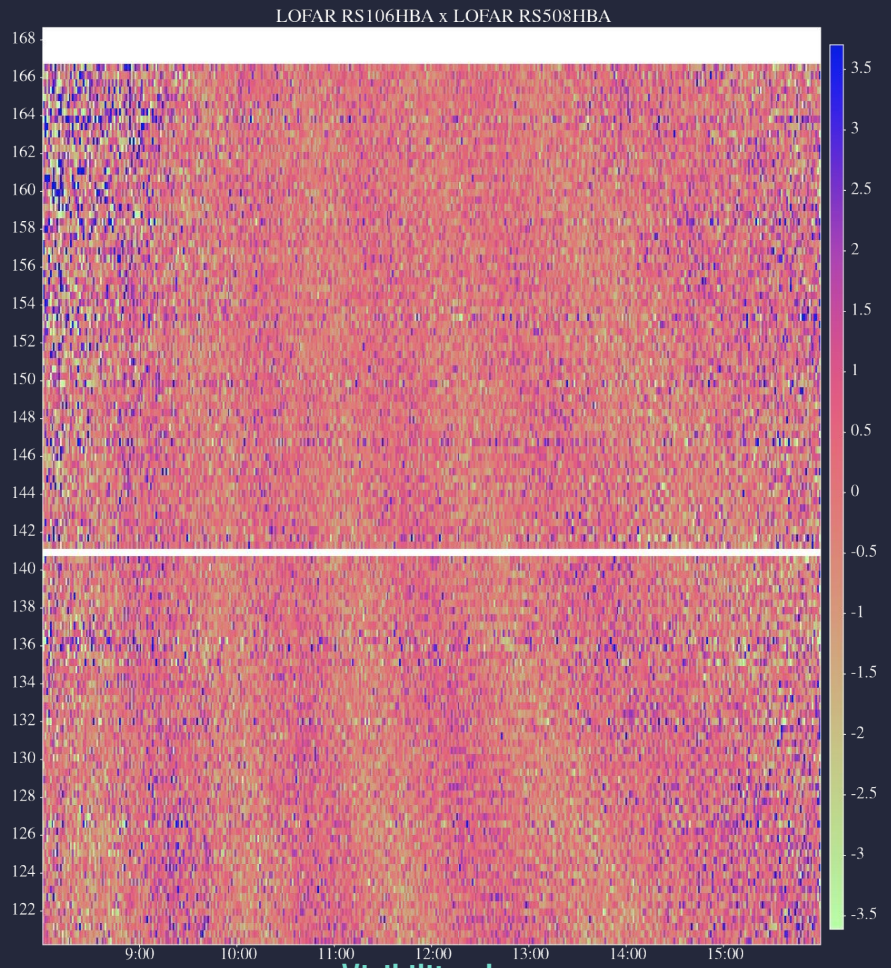


This is the “AveragingPipeline” data that you download from the Long Term Archive (LTA).
In other words: “raw data” is not always so raw.

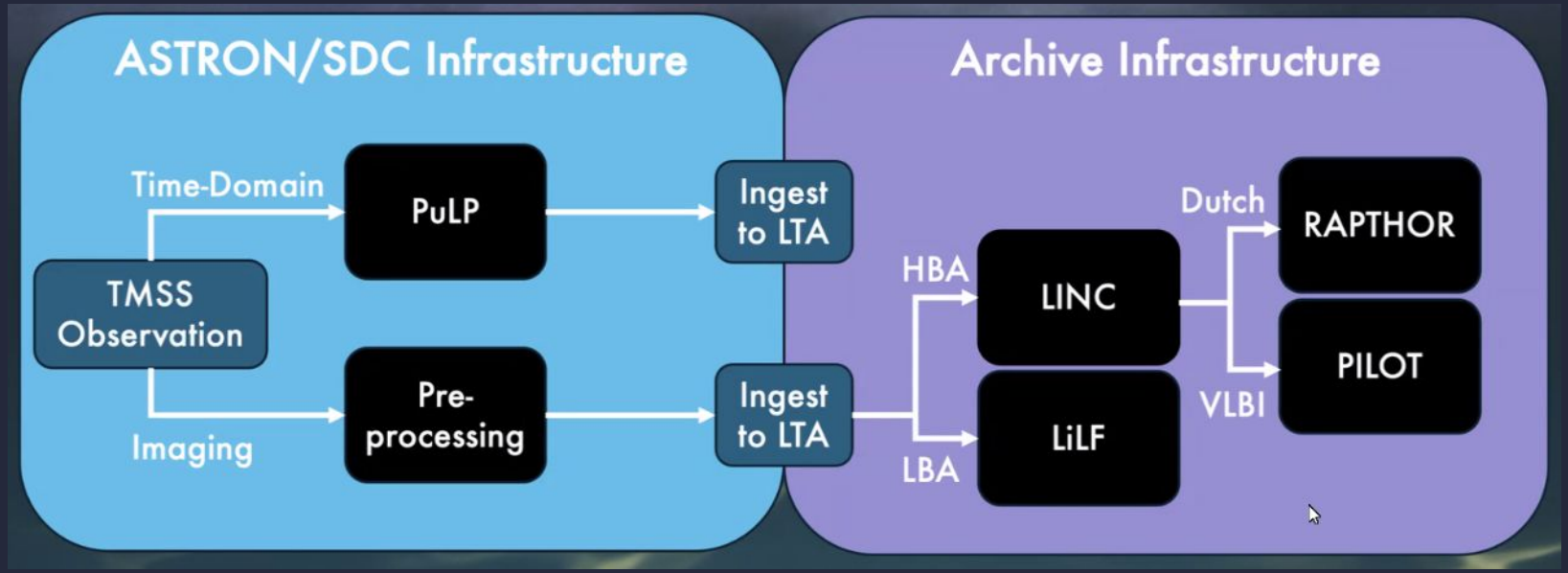





Total intensity




Visibility phase


ASTRON/SDC Infrastructure
Archive Infrastructure
Time-Domain
PuLP
TMSS Observation
Imaging
Pre-processing
Ingest to LTA
Ingest to LTA
HBA
LBA
LINC
LiLF
Dutch
VLBI
RAPTHOR
PILOT

N. Peters

# Contents

Telescope to archive

Correlation
Preprocessing

Archive to image - general concept

Instrumental corrections
Direction-independent calibration
Direction-dependent calibration

# Instrumental corrections

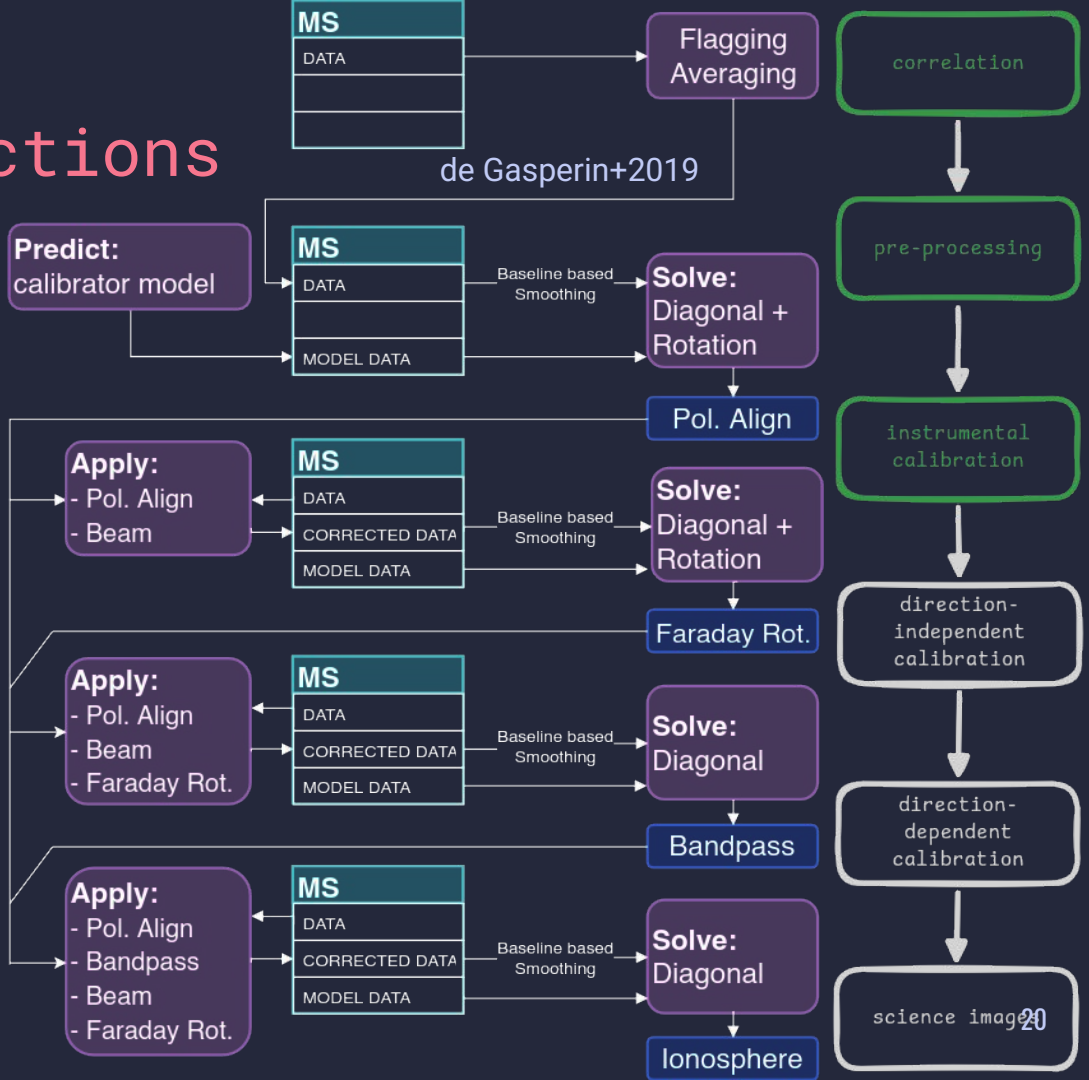
**Short (10-15 min; HBA) or simultaneous (LBA) observations of a flux density calibrator to derive**

Polarisation alignment corrections (XX-YY phase)

Station bandpasses

Station clock drifts

LINC for HBA, LiLF for LBA



# Direction-independent calibration

**Corrects for the bulk effects using an initial model (point or otherwise) and selfcal**

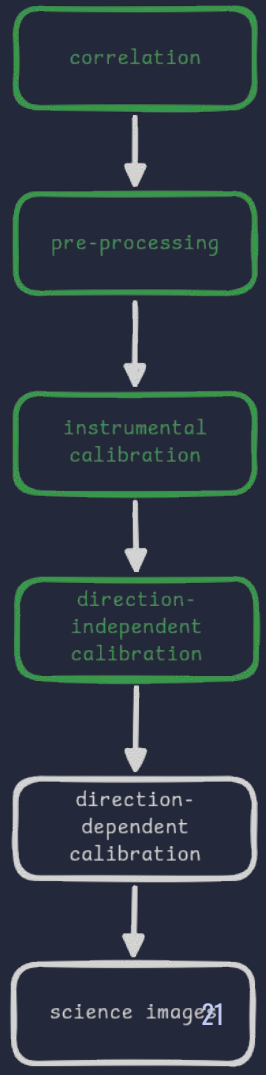
Fast ionospheric wrapping
Residual clock drifts
Primary beam at the pointing centre

LINC target HBA NL: simple phase correction against a model

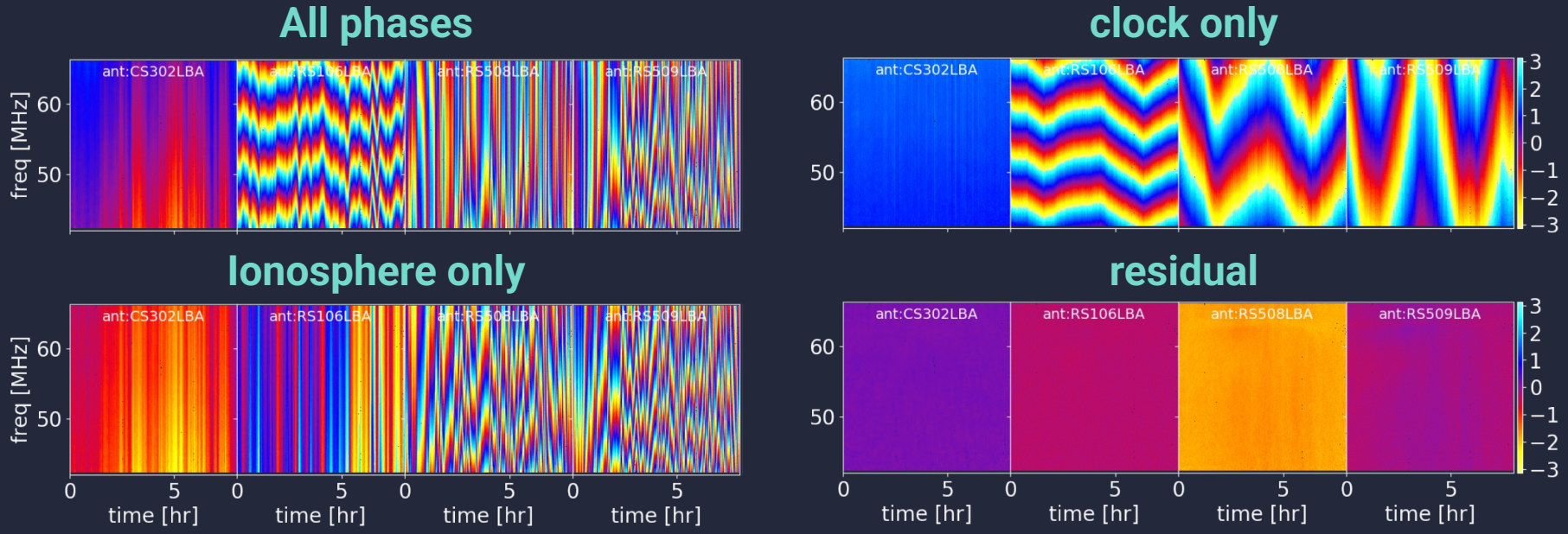
ddf-pipeline: refined DI solve accounting for DD corruptions

LiLF: refined DI solve after DD subtract

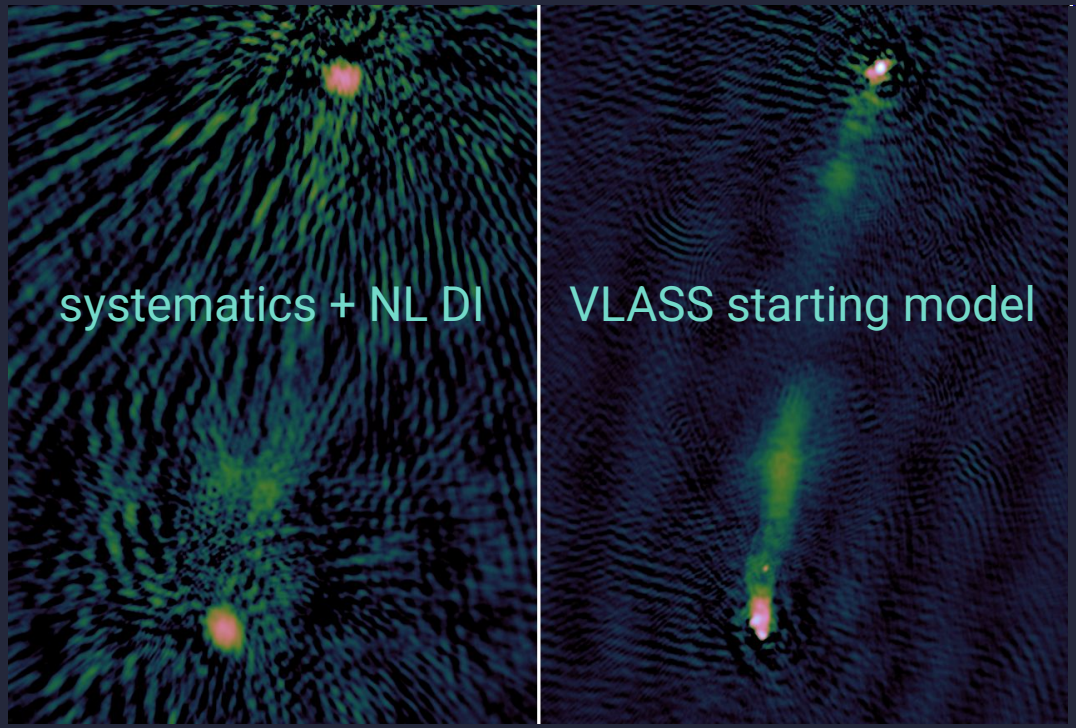
PILOT: full self calibration on reference source



# Direction-independent calibration

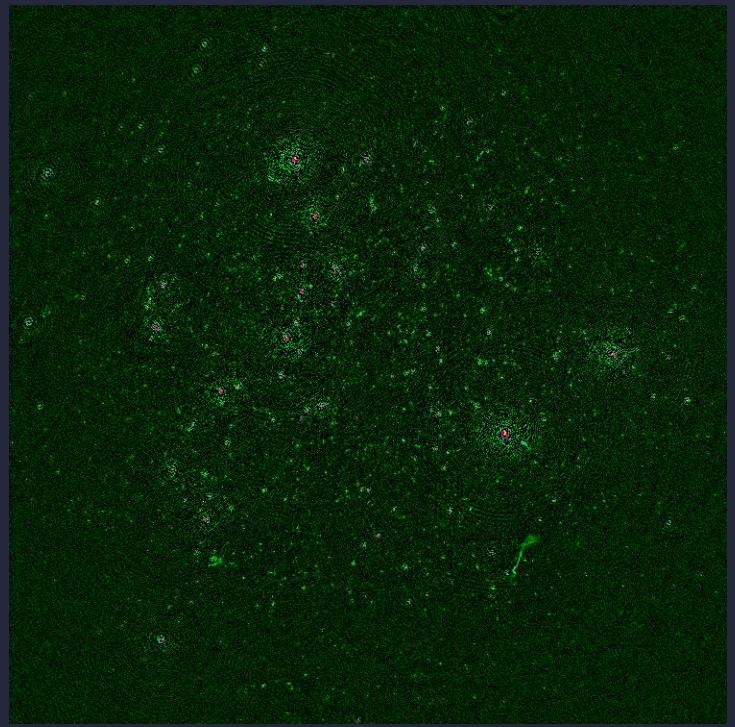


# Direction-independent calibration



# Direction-independent calibration

**Direction-dependent effects still clearly visible as spatial variations in ionosphere and station beams are not accounted for**



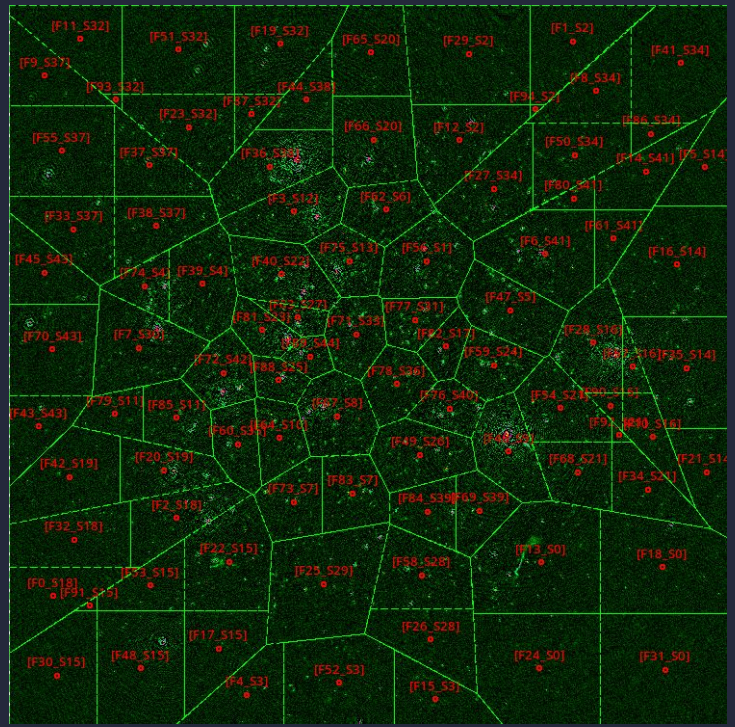
# Direction-independent calibration

**Direction-dependent effects still clearly visible as spatial variations in ionosphere and station beams are not accounted for**

Select regions with bright sources and high signal as “direction-dependent calibrators”

Form “facets” with Voronoi tessellation

SNR away from the centre will drop due to primary beam attenuation, so facets grow

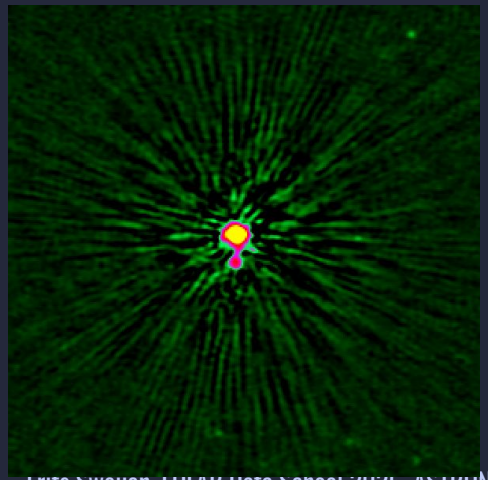


# Direction-dependent calibration

**Correct for the spatial variations across the field of view**

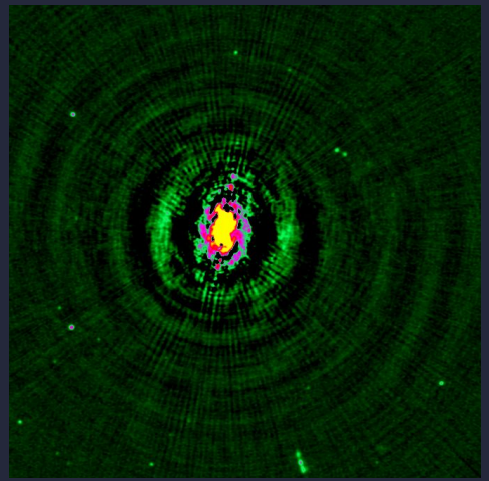
Differential ionosphere (phase)

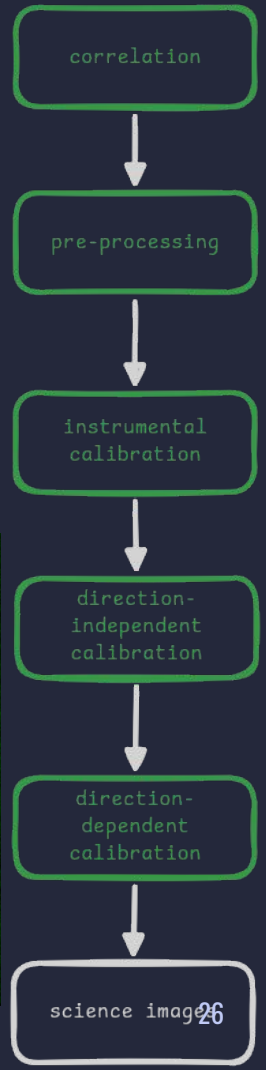
Primary beam variations (amplitude) due to e.g. imperfect beam model



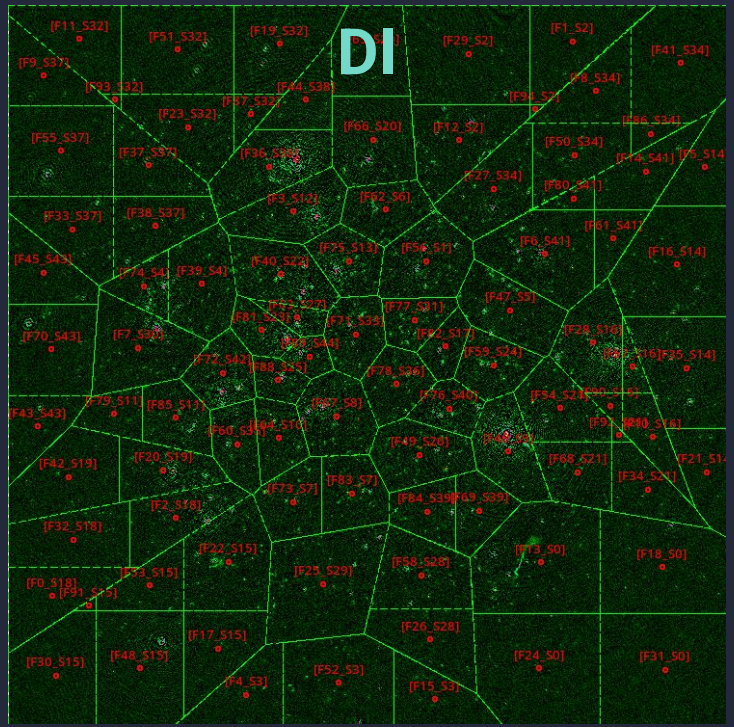
← phase error: spoke-like and antisymmetric

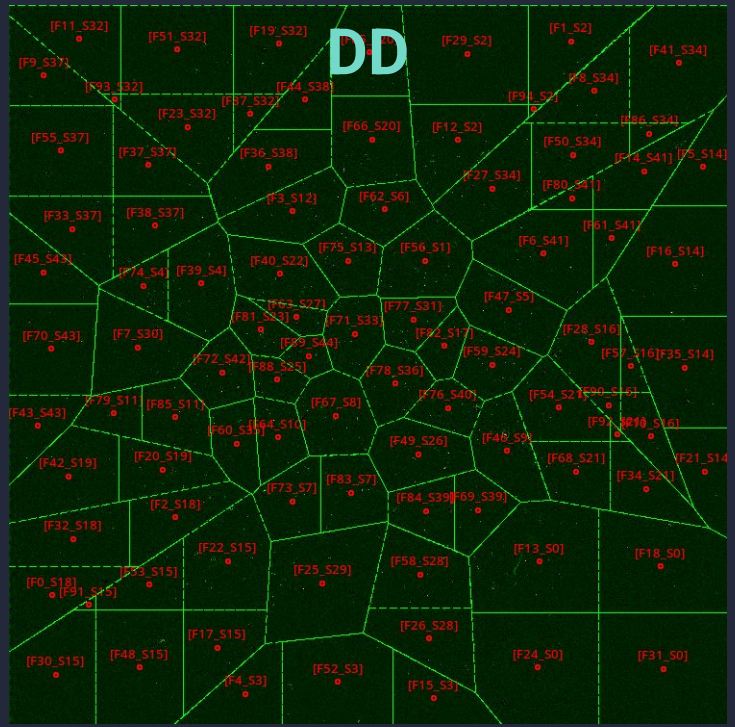
amplitude error: typically → ring-like and symmetric



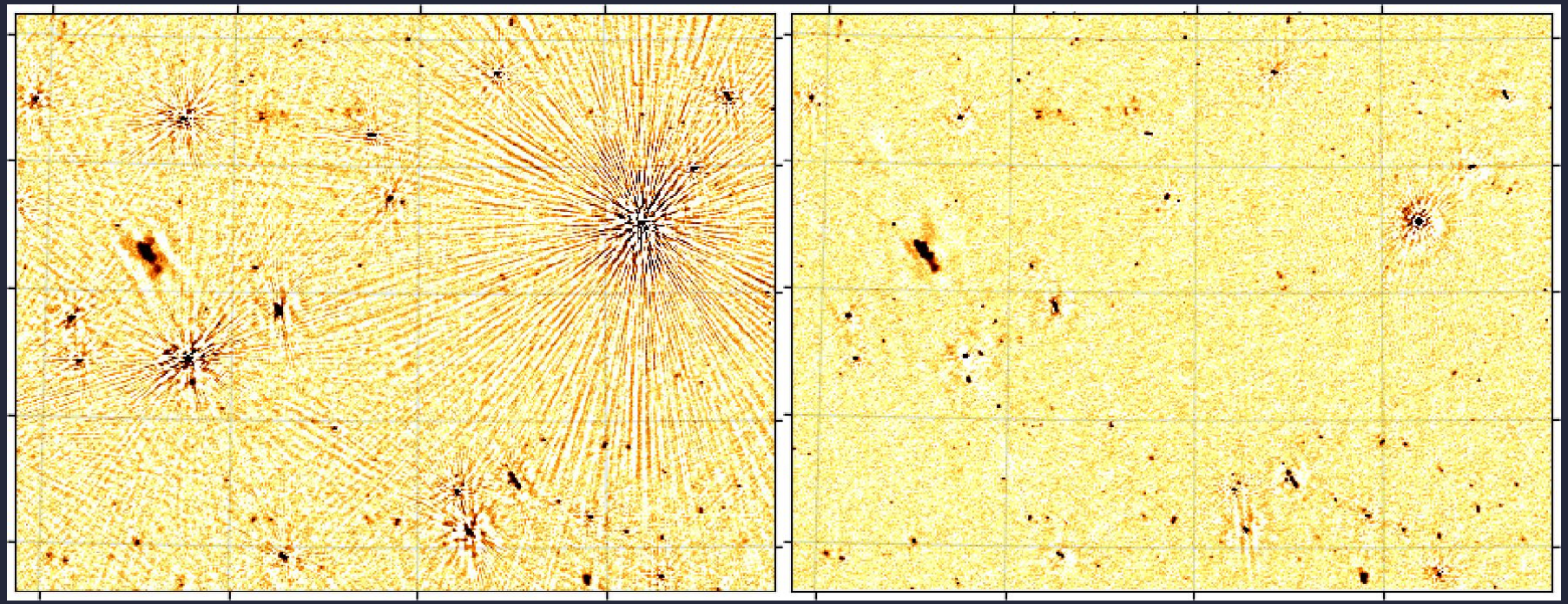


# Direction-dependent calibration



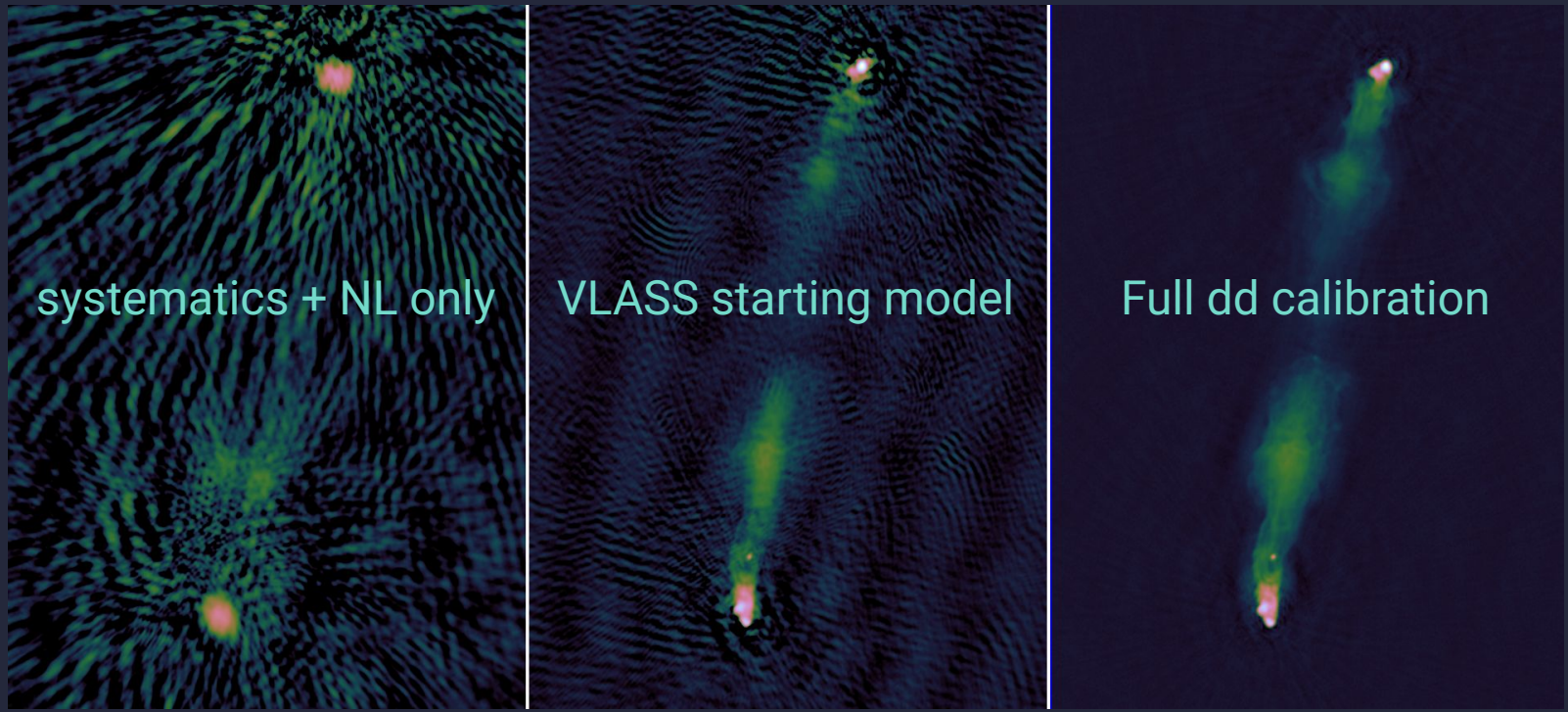



# Direction-dependent calibration



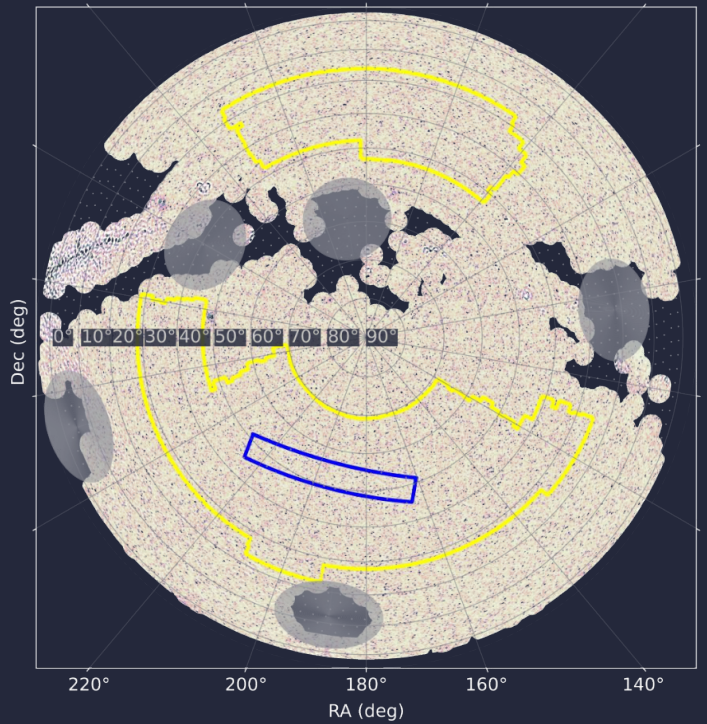
de Gasperin+2023

# Direction-dependent calibration



# Science images

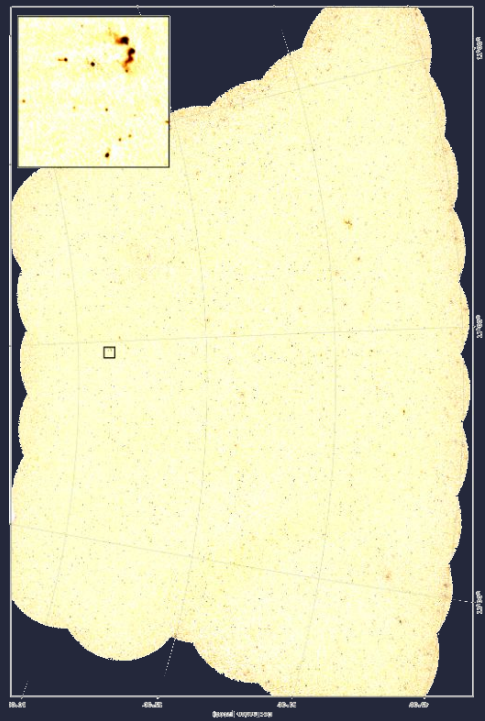
Shimwell+2026



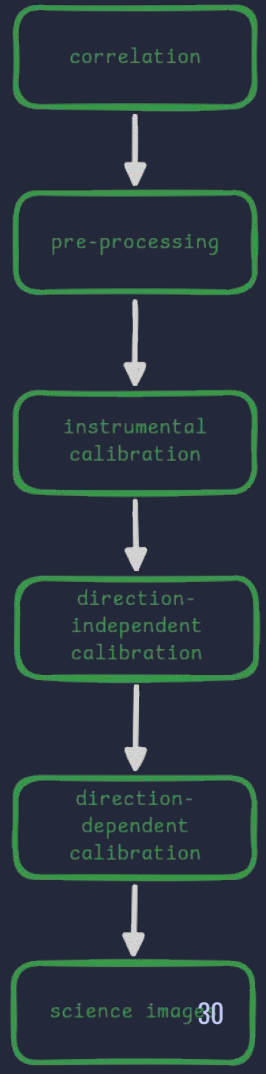

LoTSS DR3 - ~93%
of northern sky @ 144 MHz

download
process

de Gasperin+2023



LoLSS DR1 - ~3.2%
of northern sky @ 54 MHz





# Compute costs

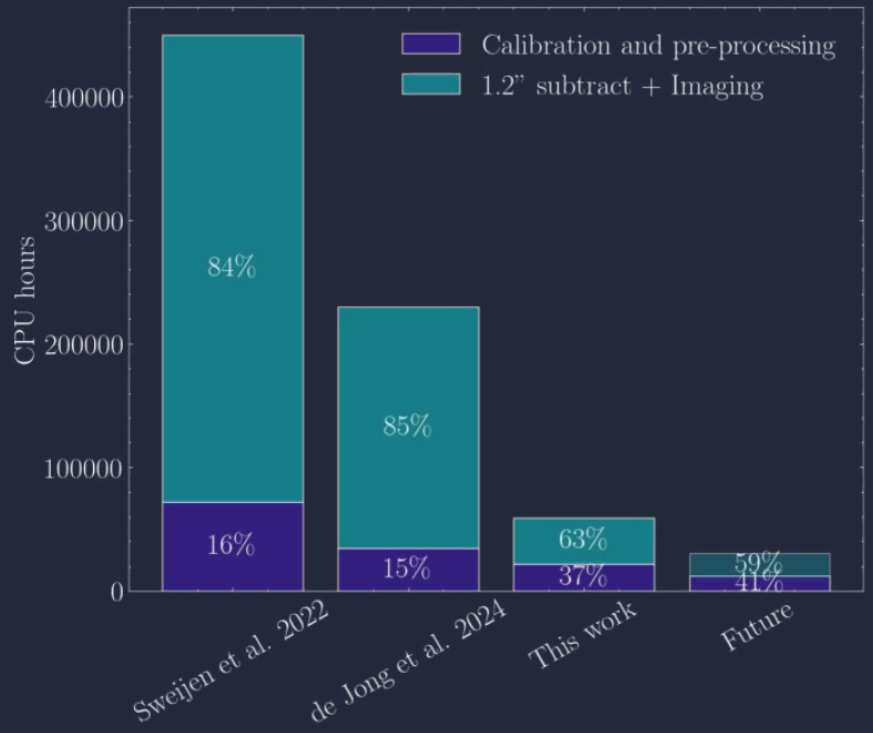
6” widefield image: ~5,000 CH end-to-end
-> 20M core hours for LoTSS

0.3” widefield imaging is now becoming affordable:
~60,000 CH end-to-end
-> 180M core hours for iLoTSS

Processing happens on (inter)national or university supercomputer infrastructure

**Surveys tend to be a multi-year endeavour!**

de Jong+2025



# Summary




Multiple pipelines for various stages of the data

Conceptually straightforward, but still demanding due to data volumes

With LOFAR2 hopefully more science-ready products into the LTA :)