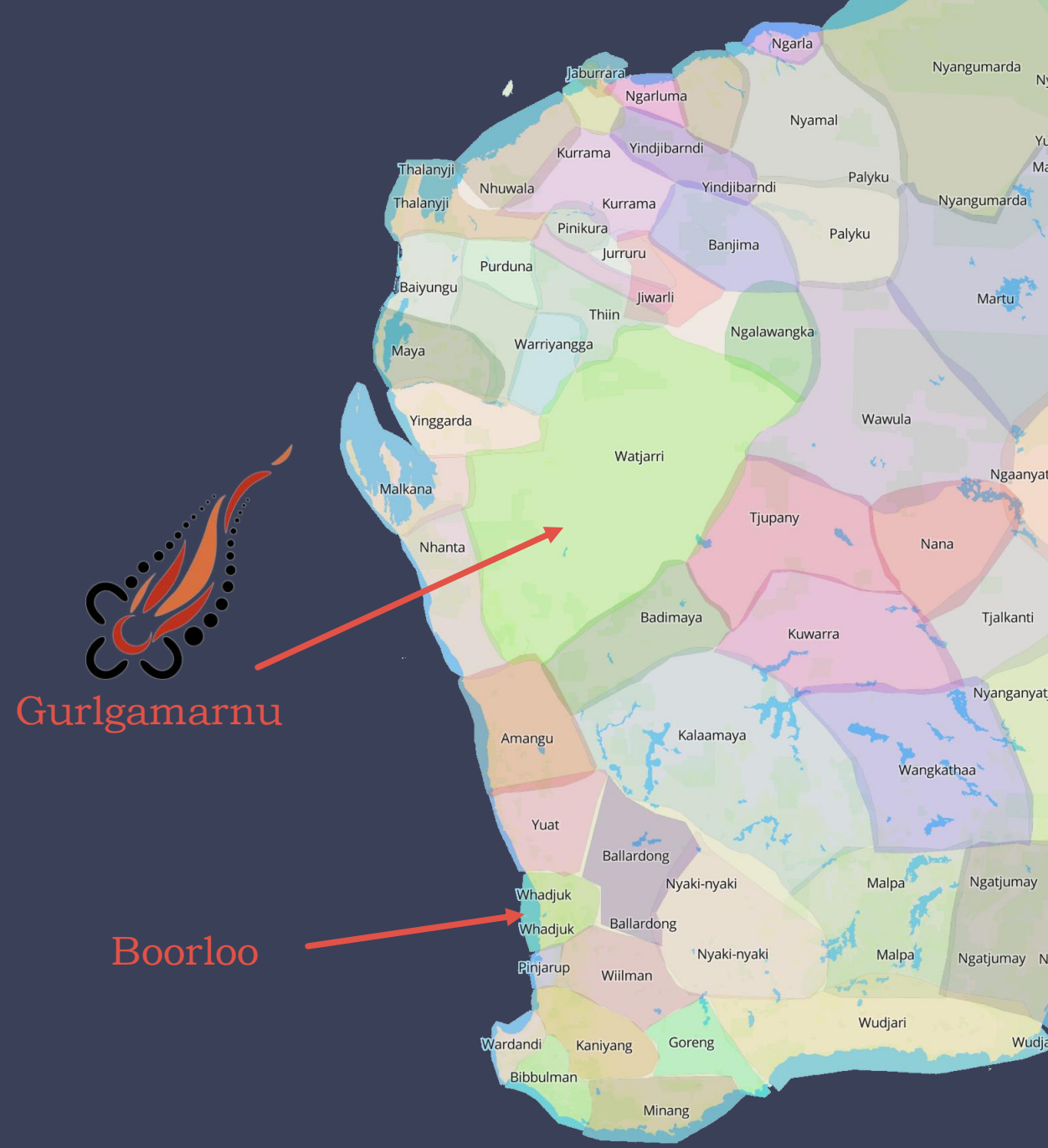
MX DEV NULL (THEY/THEM)

MWA PROJECT MEETING

2022-12-07

Acknowledgement

- Gurlgamarnu (Murchison Widefield Array, “the ear that listens to the sky.”) and Inyarrimanha Ilgari Bundara (CSIRO Murchison Radio-astronomy Observatory, “sharing sky and stars”) are on the lands of the Wajarri-Yamaji people
- Boorloo (Perth) is on the lands of the Whadjuk Noongar people
- Thank you for sharing your sky with us



source: native-land.ca

Motivation

- MWA is a sensitive instrument in a hostile environment
- MWA EoR Pipeline team has automated the detection of systematics that impact EoR science
- Our techniques are applicable to other science groups, software stacks
- Better awareness of systematics → better science!

source: Andrew McPhail, J: drive

The systematics presented in this talk do not represent typical MWA data.

MWA EoR Pipeline Team

Prof. Cathryn Trott

Pipeline design,
Power spectra (CHIPS)

Dr Ridhima Nunhokee

Quality analysis metrics
(mwa_qa)

Dr Chris Jordan

Calibration, Ionospheric
Subtraction (Hyperdrive)

Mx Dev Null

Nextflow pipeline,
HPC whisperer

Signal Chain Overview

Potential Systematics
Mitigation techniques - M&C
Mitigation techniques – MWA EoR Pipeline

local fauna

RFI

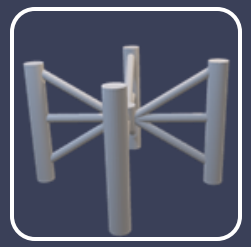
Bad pointings / beam shape

Bad receiver state / freqs / gains

Cable kinks

Bad fibers

Bad metadata



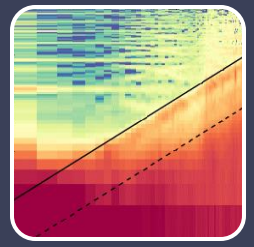
Dipoles, Tile Beamformer

Receiver

Correlator, Archive, Databases

Preprocessing
- Geometric Corrections
- PFB gains
- RFI flagging

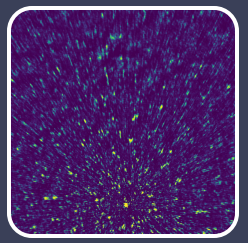
Calibration
- Direction Independent
- Ionospheric subtraction



Power Spectra

Delay Spectra

Window / Wedge power



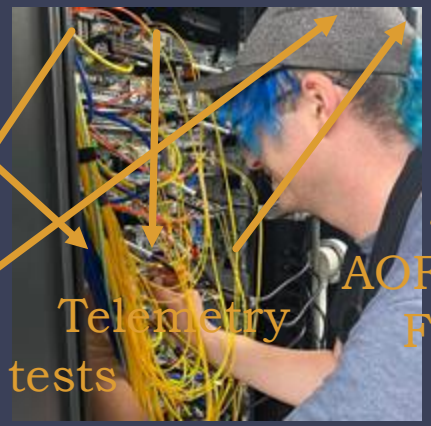
Imaging

Polarimetric RMS Ratios

All-sky images

Physical inspection

Daily Dipole tests



Telemetry

AOFlagger Flags

Calibrator DI solutions

SSINS Flags

Autocorrelation RMS

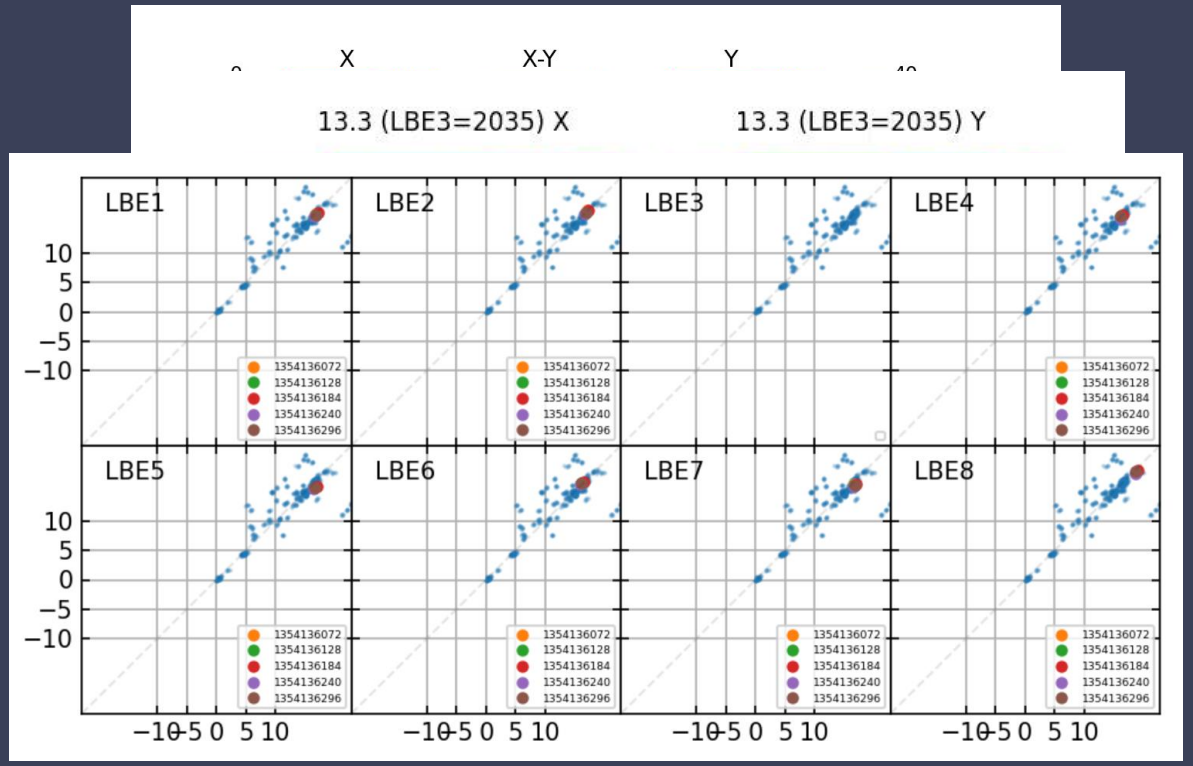
Observation DI solutions

Gain Variance

Convergence Variance

Web Services

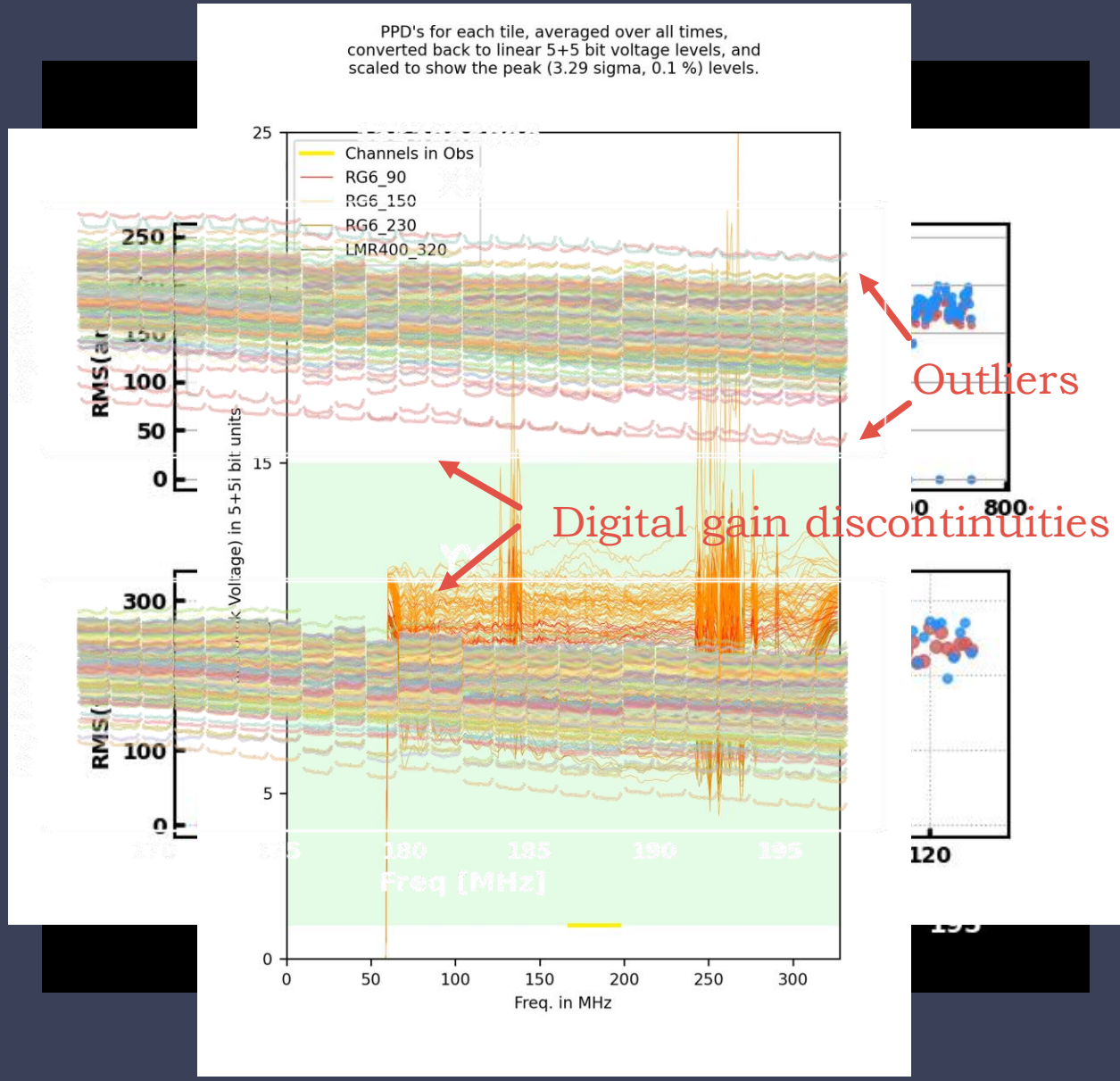
- PPD: Integrated pre-PFB power for each tile, polarization on all channels
- 12 Hour Waterfall Plots
  telemetry-static.mwa128t.org/powerplots
  - Median PPD over all tiles
  - PPD of each tile relative to median
  - Great for RFI spotting
- Daily dipole tests
  telemetry-static.mwa128t.org/dipoles
  - PPDs one dipole at a time
  - Great for spotting bad dipoles
- Calibration fits
  ws.mwatelescope.org/calib/summary
  - Gain and delay fits for each tile
  - Great for spotting bad tiles



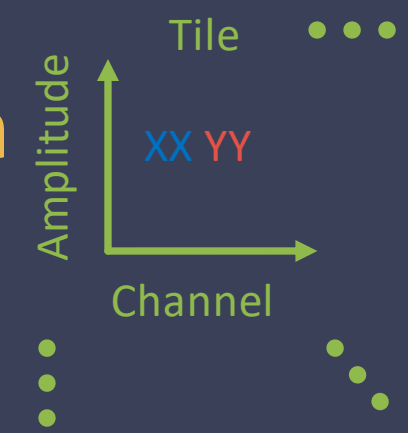
Due to its subjective nature, bad tile detection is not automated

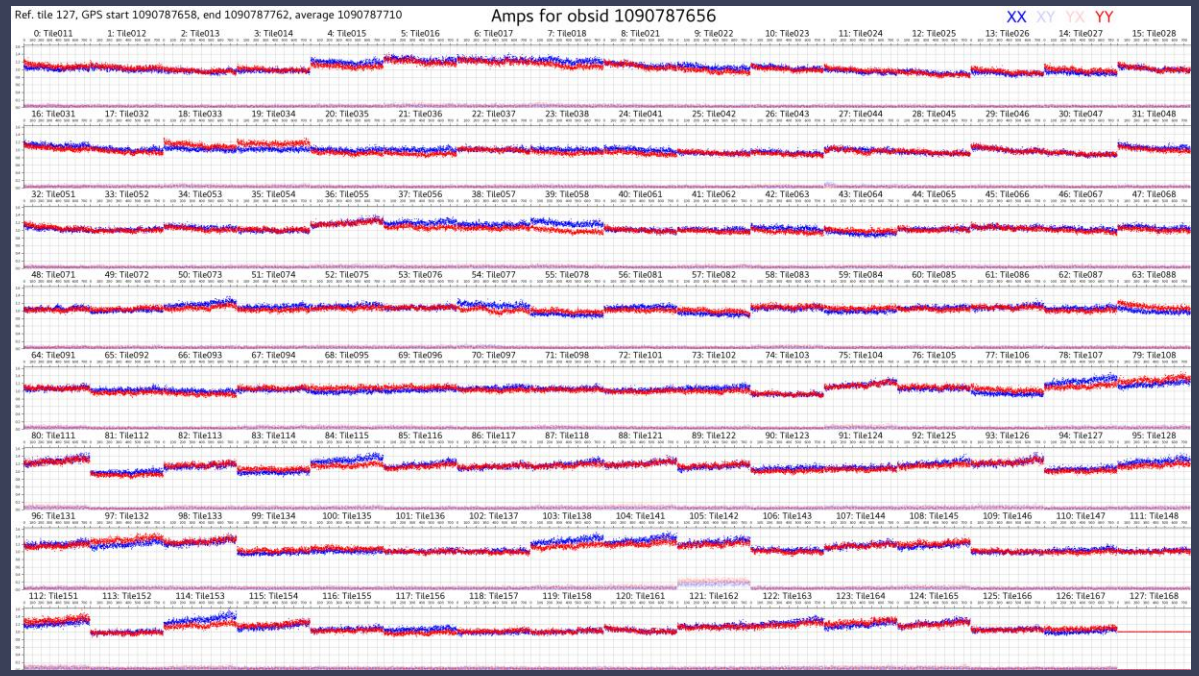
Diagnostics: Autocorrelations

- RMS analysis of autocorrelations
  - Exclude tiles flagged in metafits, edge channels, center channels, quack time
  - Take RMS of amplitude across frequency, time
- Flagging loud or quiet tiles can improve results
- Digital gain discontinuities are
- We flag tiles outside ±3σ RMS



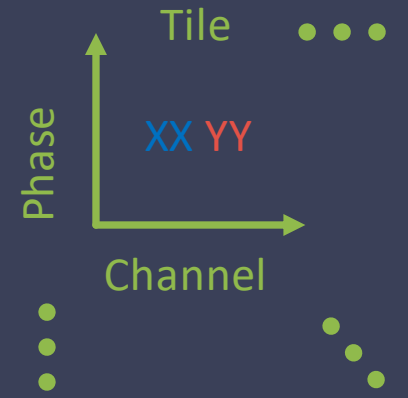
Diagnostics: Calibration Solutions

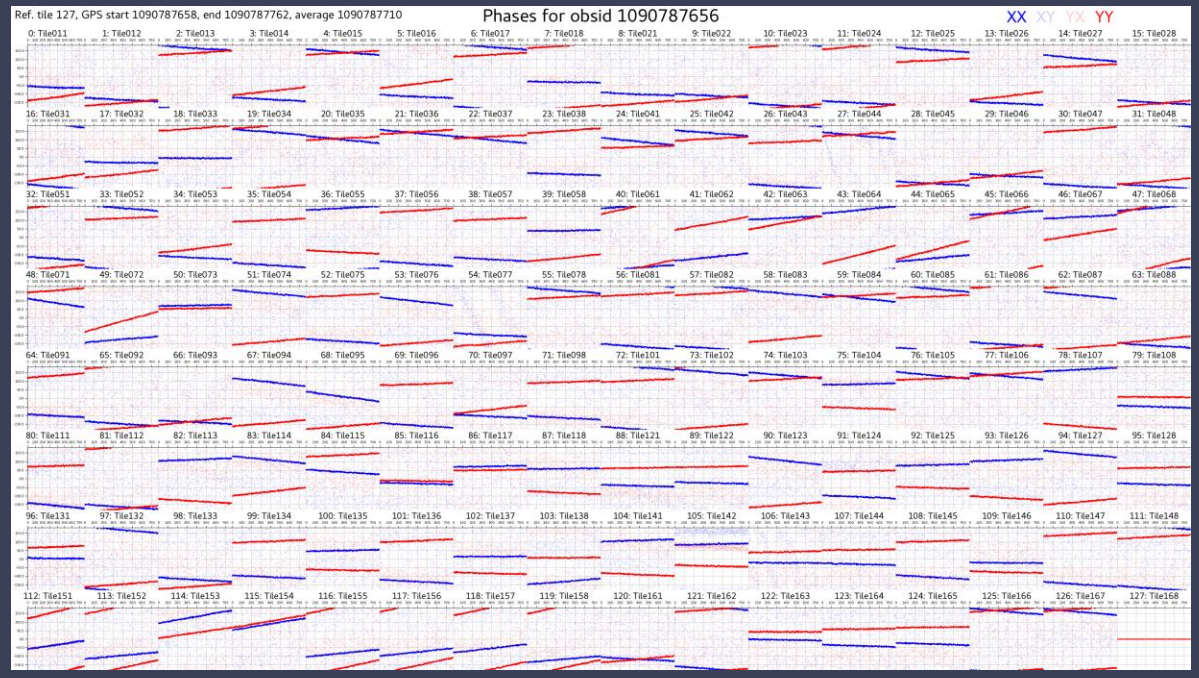




Nominal calibration solutions

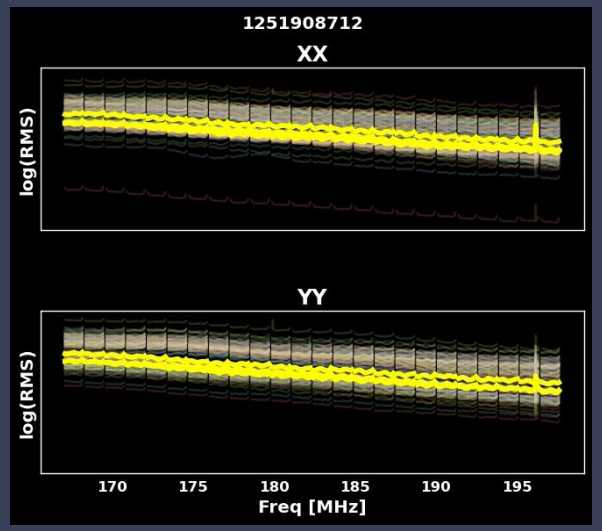
- Direction Independent calibration produces gains and phases for each polarization in each tile
- Good solutions are linear across channel number, amps are usually fuzzier



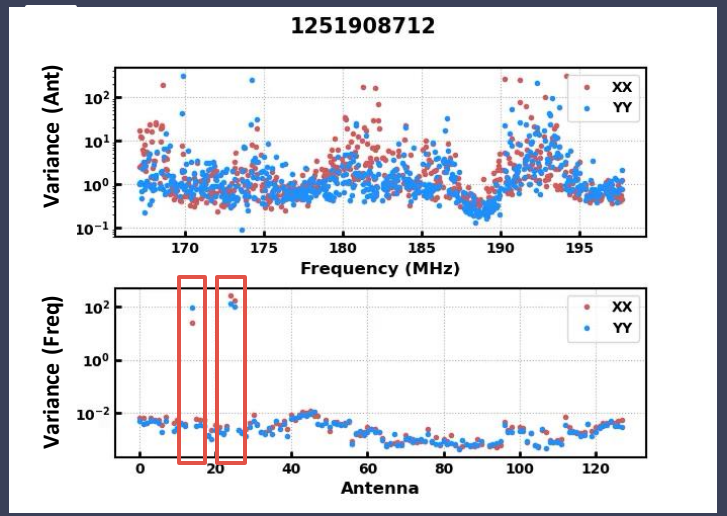


Calibration Solution Example

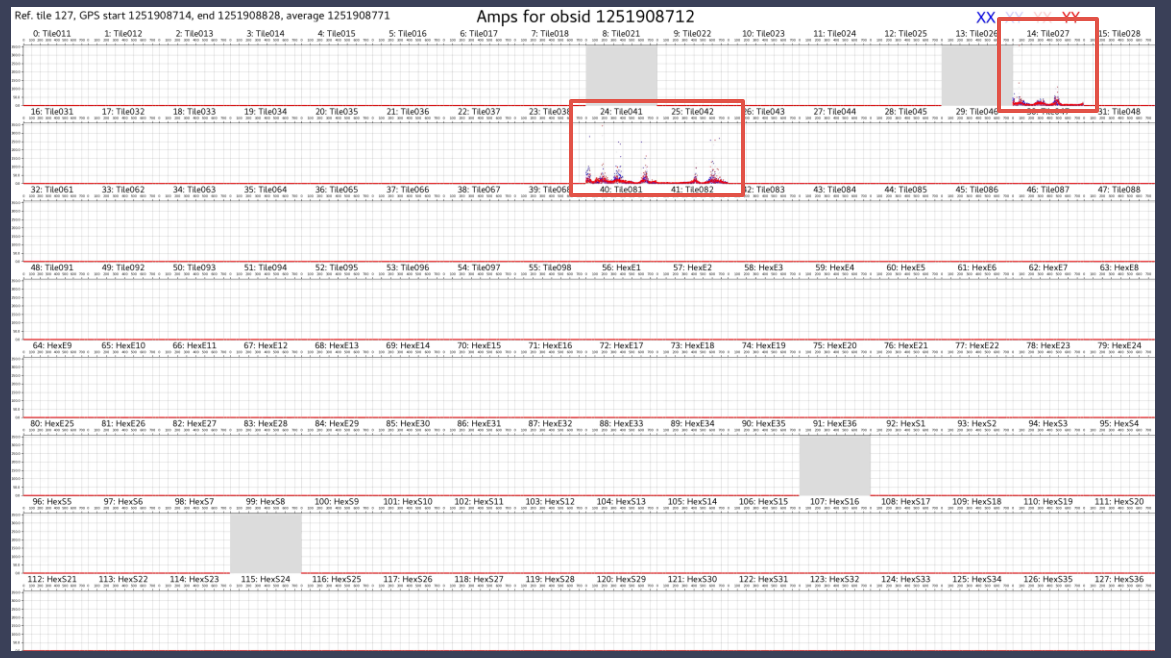
- Tile027, Tile041, Tile042 are not calibrating well
- But their Auto RMS is fine
- Look at the variance of the solutions across channel for each antenna



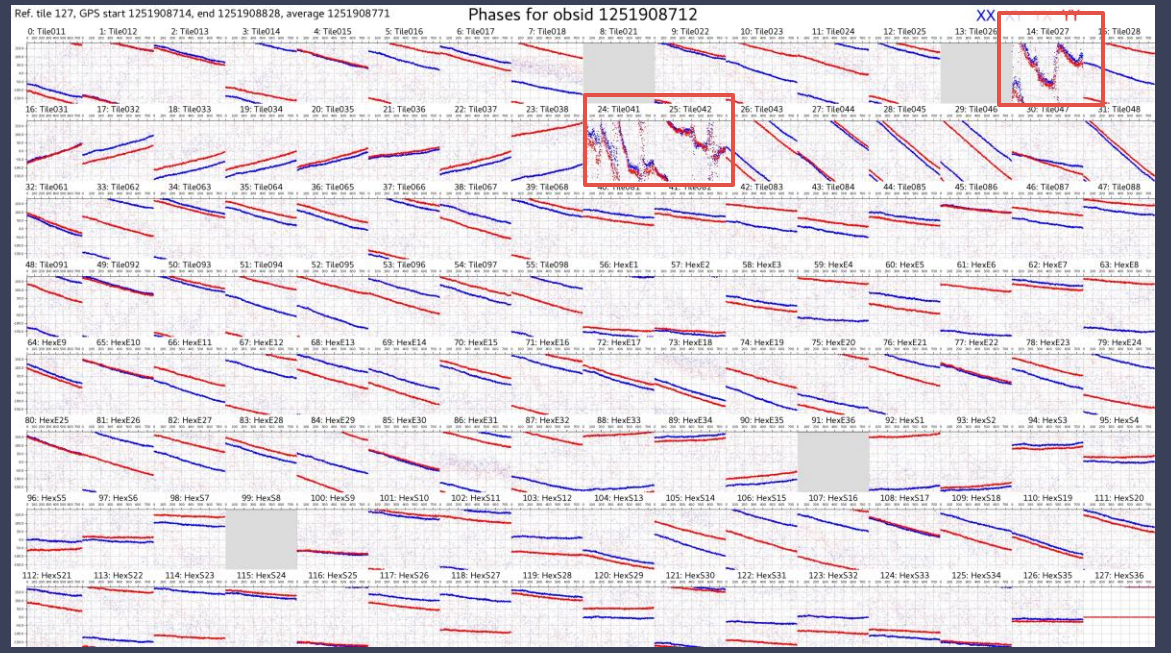
Auto RMS



Variance

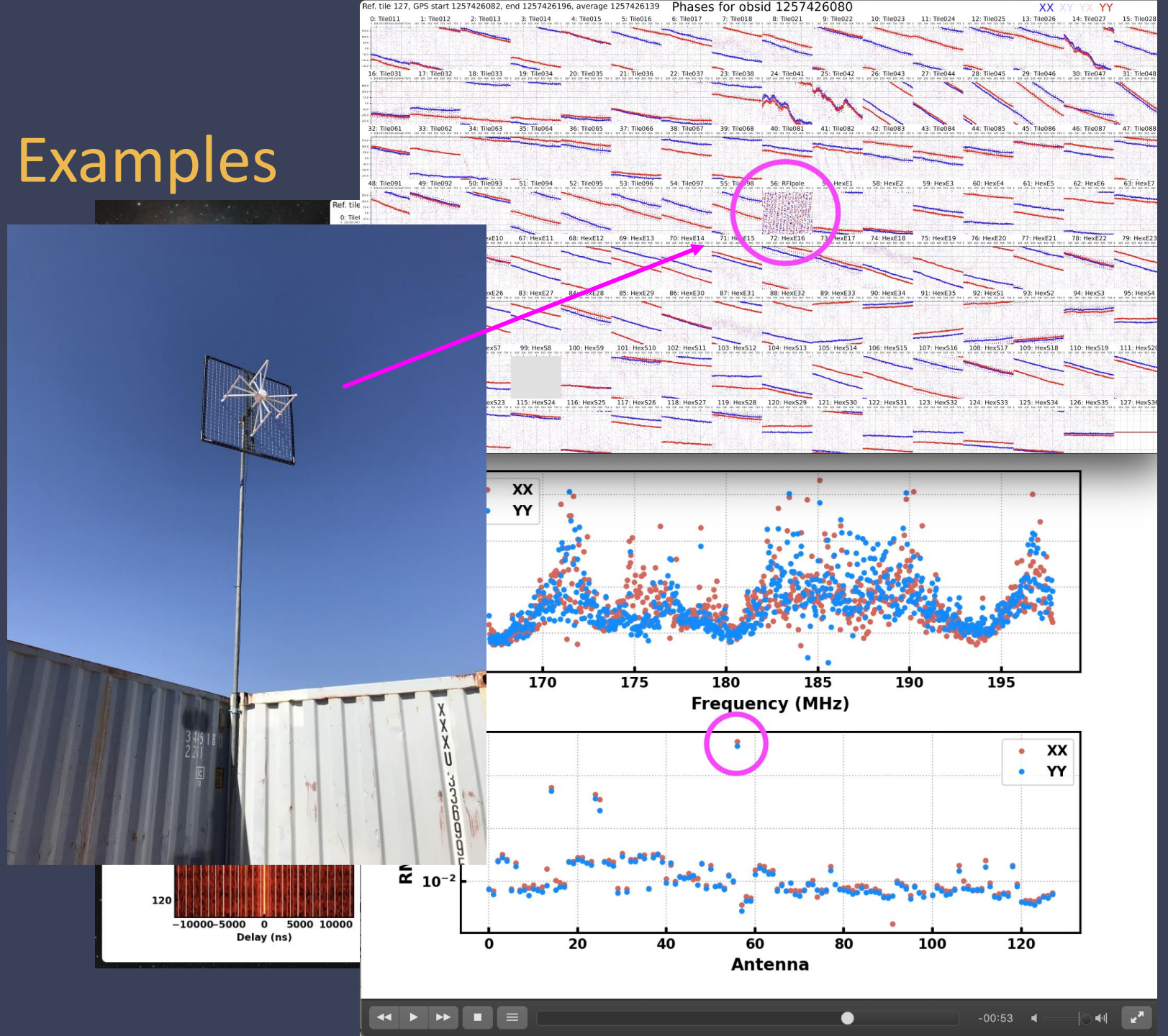


Calibration solutions



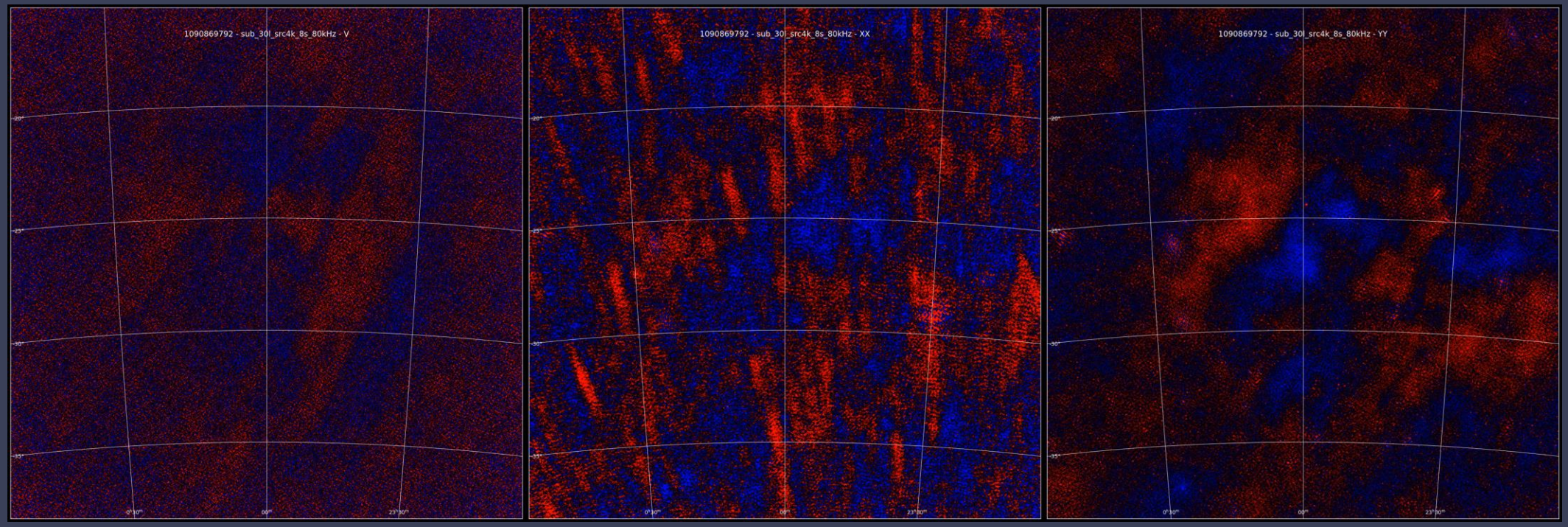
Calibration Solution Examples

- Diagnosed a fault in a fiber from receiver 4
- Found an RFIPole test observation



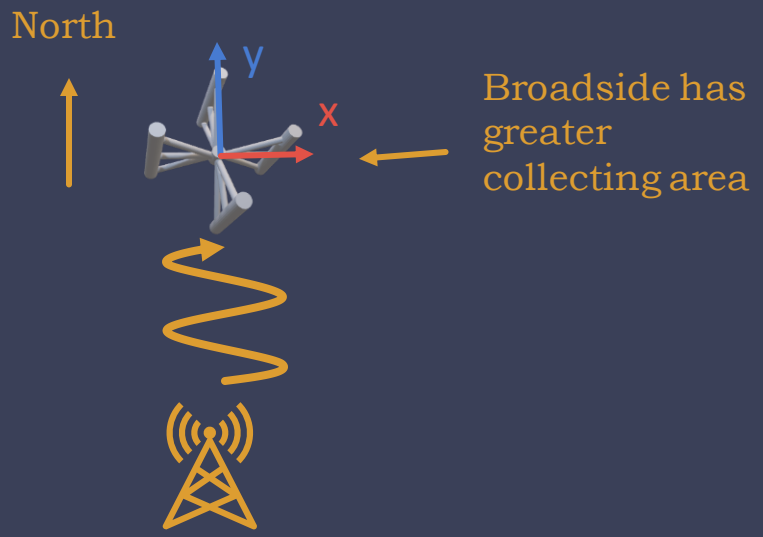
Diagnostics: POLARIMETRIC Imaging

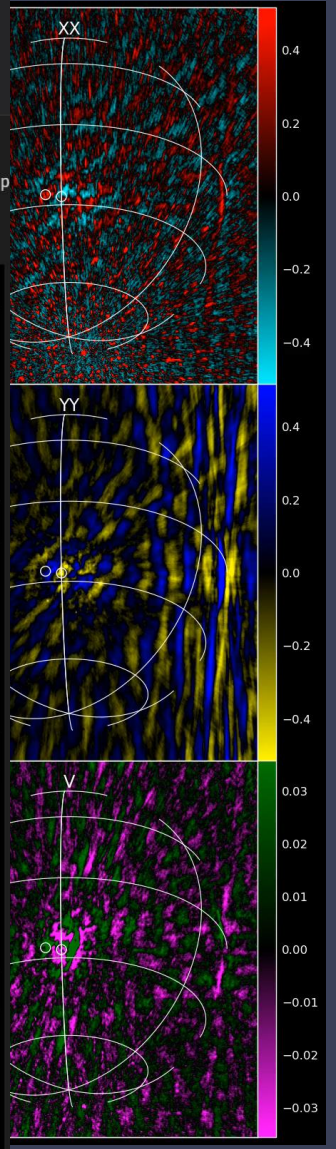
- Subtract model from calibrated visibilities
- Make polarimetric images
- Polarimetric power ratios → leakage, data quality issues



Diagnostics: All Sky Imaging

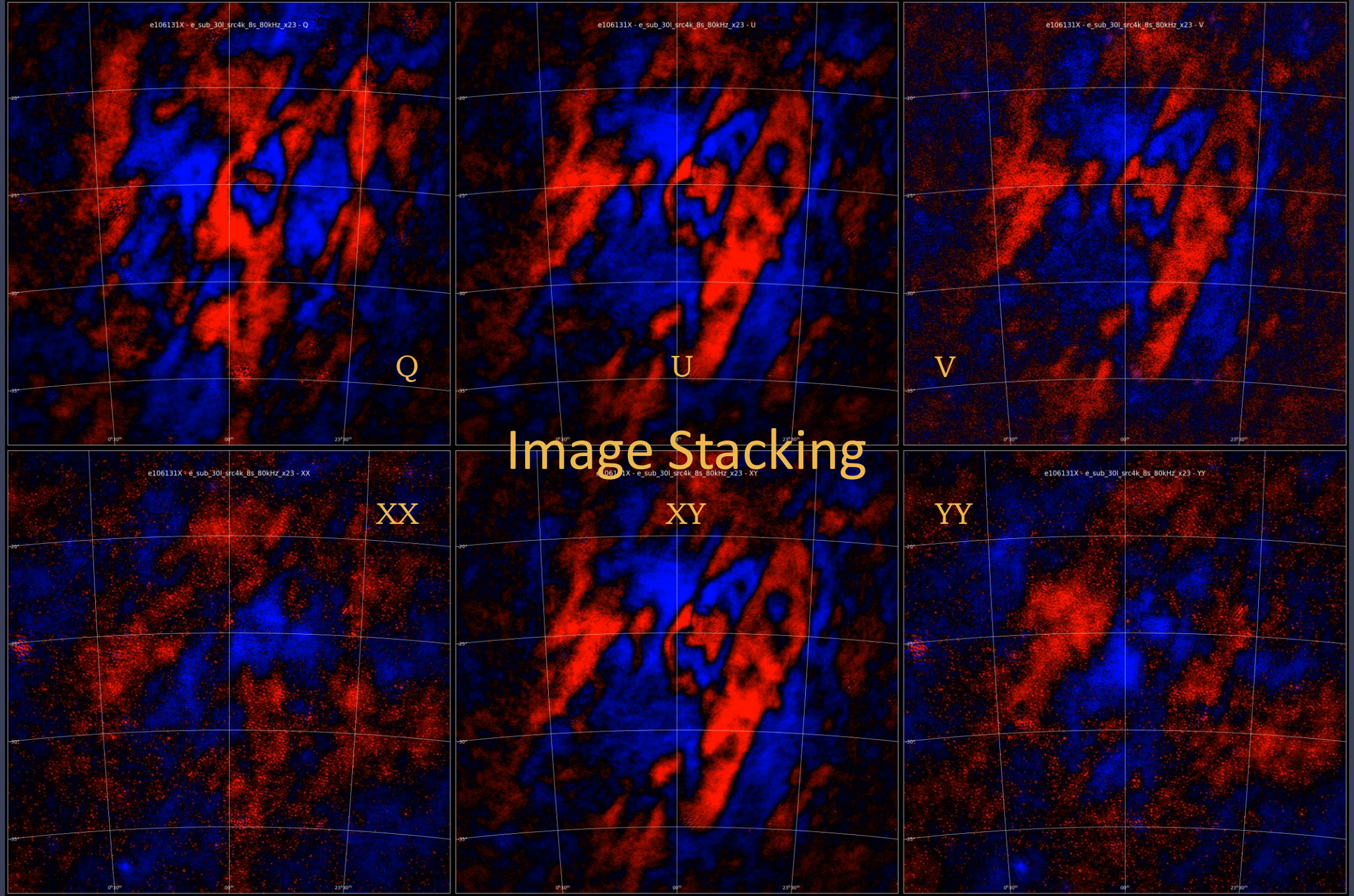
- Sources on the horizon are detected by broadside polarizations
- RFI can sometimes be visible on the horizon





(YY,V)

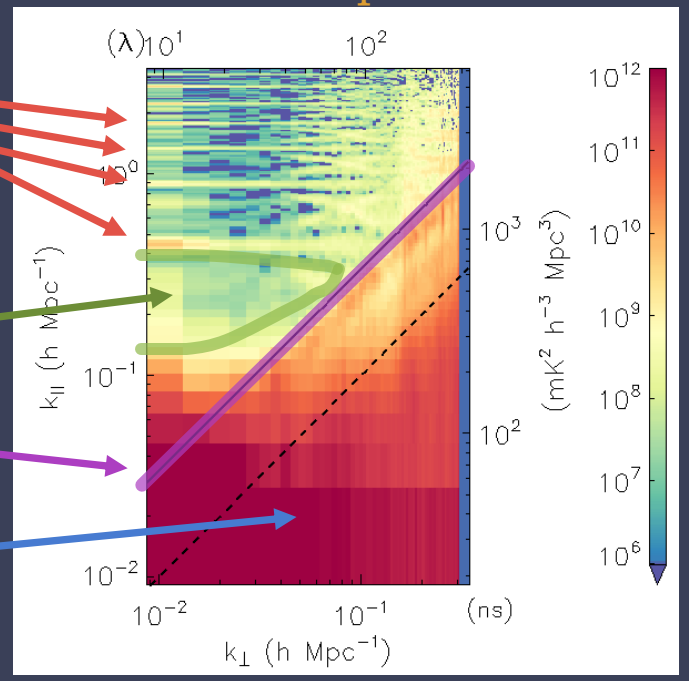
es (XX)


e106131X - e_sub_30l_src4k_8s_80kHz_x23 - Q
e106131X - e_sub_30l_src4k_8s_80kHz_x23 - U
e106131X - e_sub_30l_src4k_8s_80kHz_x23 - V
Q
U
V
Image Stacking
e106131X - e_sub_30l_src4k_8s_80kHz_x23 - XX
e106131X - e_sub_30l_src4k_8s_80kHz_x23 - XY
e106131X - e_sub_30l_src4k_8s_80kHz_x23 - YY
XX
XY
YY

Power Spectra

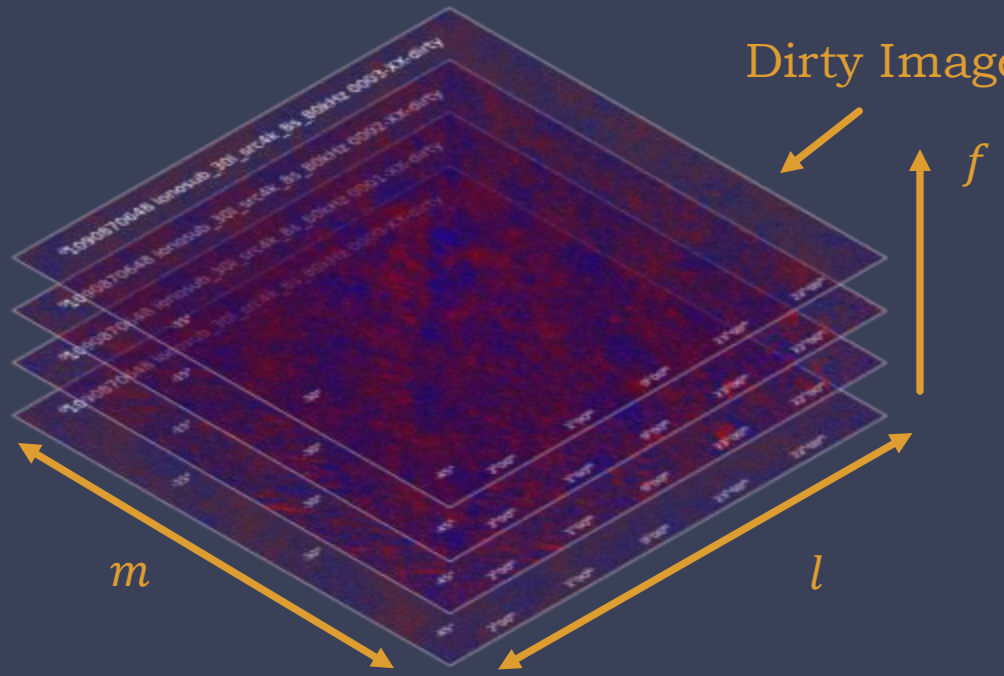
- Spectral Fourier transform of gridded visibilities
- Need several observations for good sensitivity

Power Spectrum



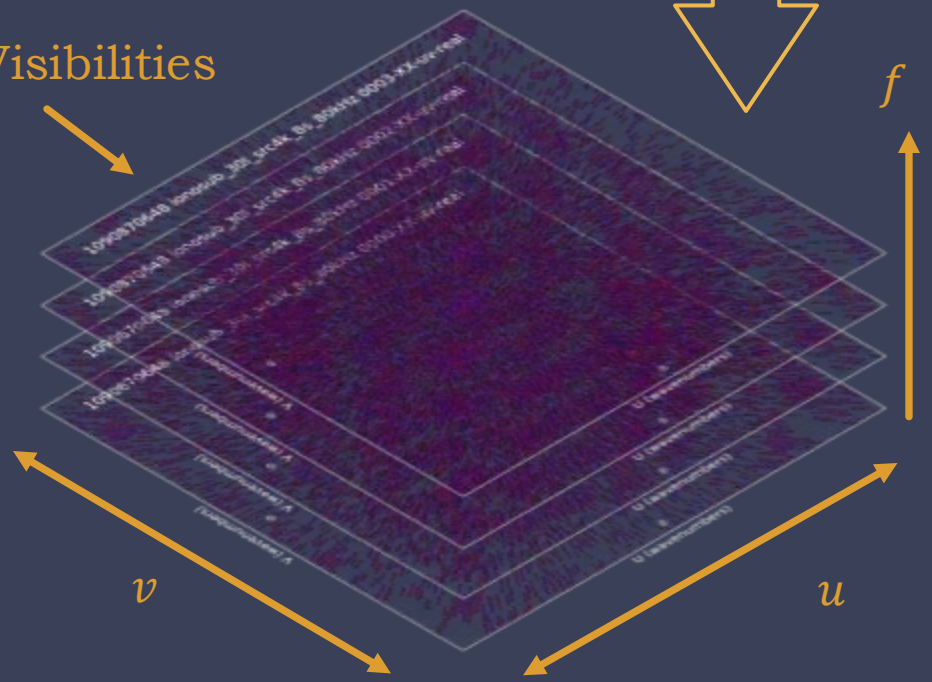
$\mathcal{F}$ spectral

Dirty Images



$\mathcal{F}$ spatial

Gridded Visibilities



Diagnostics: Delay Spectra

- Spectral Fourier transform of visibilities “gridded” by baseline length
- Faster to compute than power spectra, can use single observation.
- Easy to spot RFI with spectral structure in window

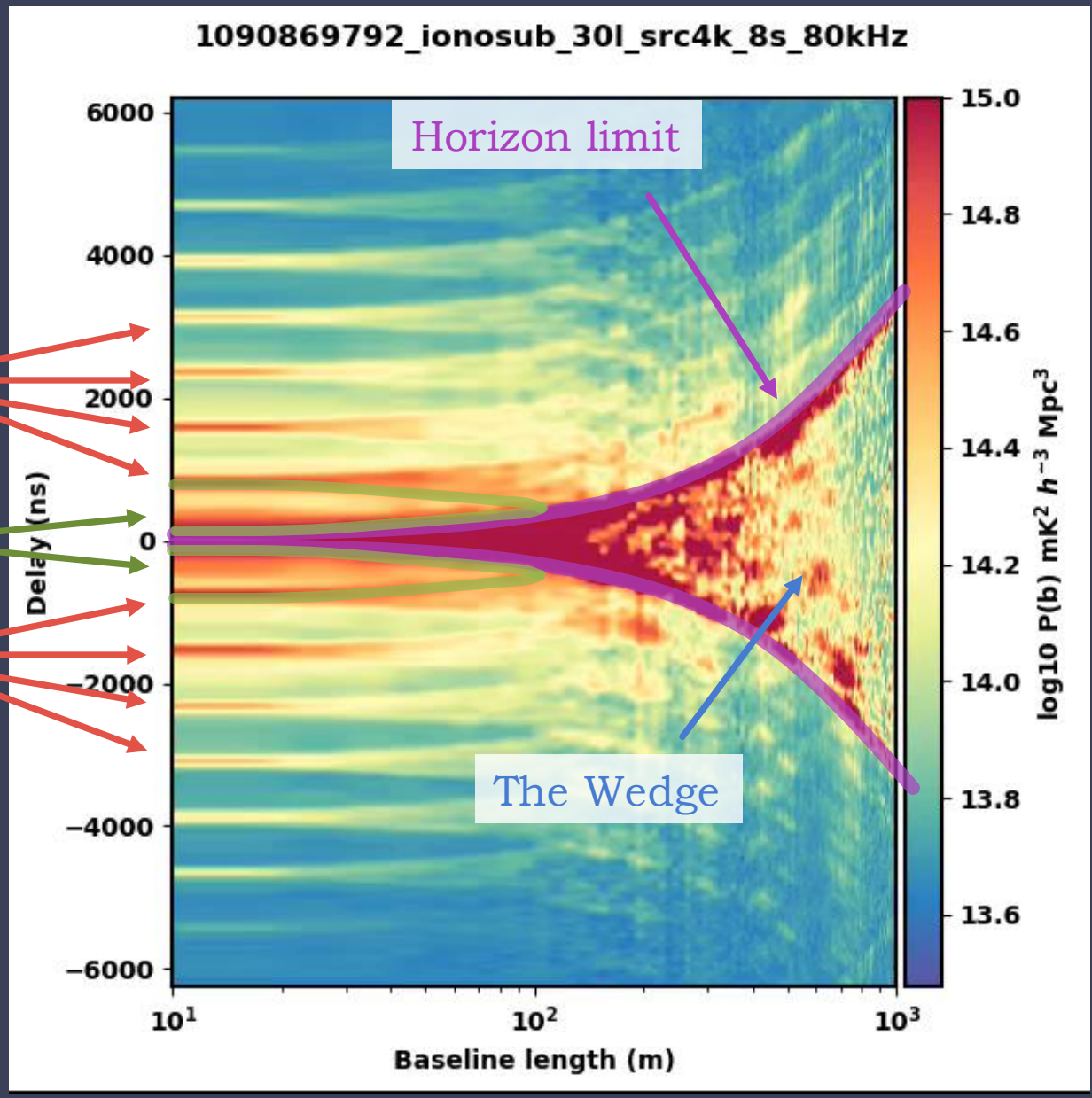
Dr Jack Line

Delay spectra code

Coarse channel harmonics

First harmonic window

Coarse channel harmonics



Pipeline Design Philosophy

Automate everything

Treat observations like cattle, not pets

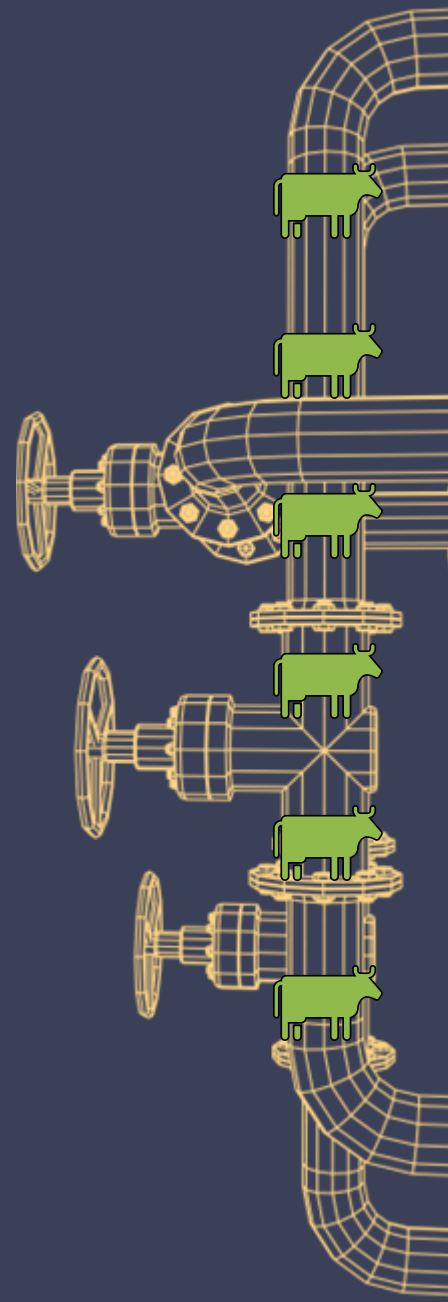
- Manual flagging is not scalable

Modular re-usable components > Monoliths

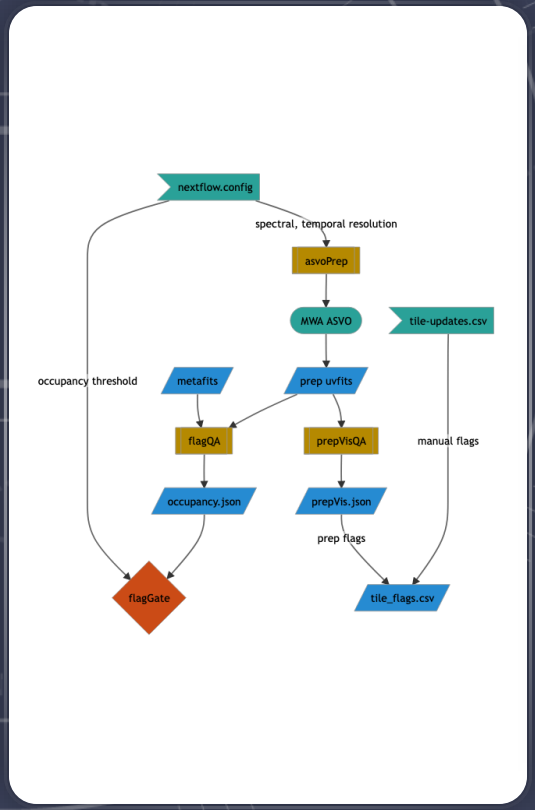
- Granular resourcing – better scheduling
- More granular error handling

Parametrize > Hardcode

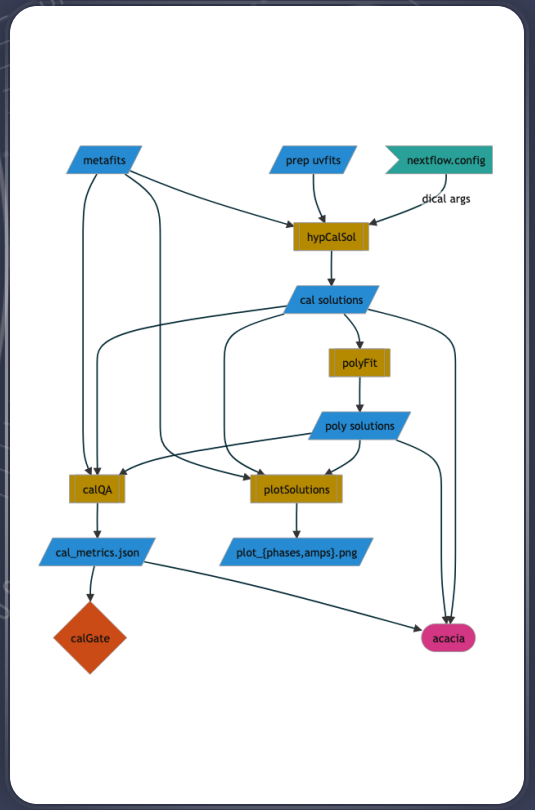
- Make no assumptions about the dimensions of your data
- Future observations will have 9180-32640 baselines



Preprocessing



Calibration



Visibility analysis

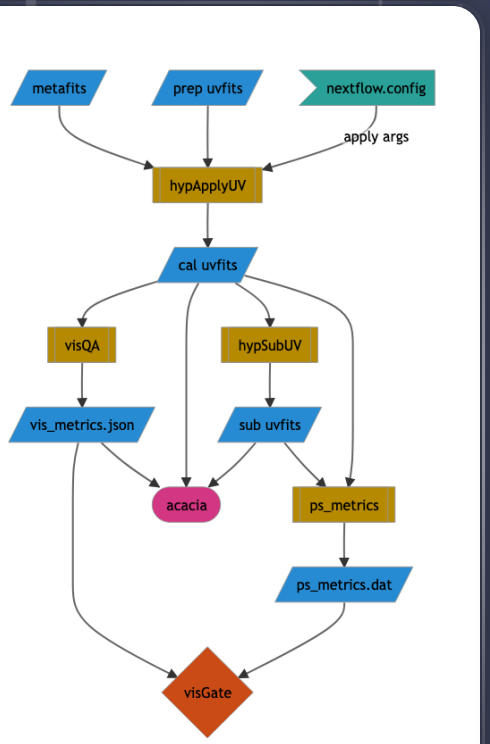
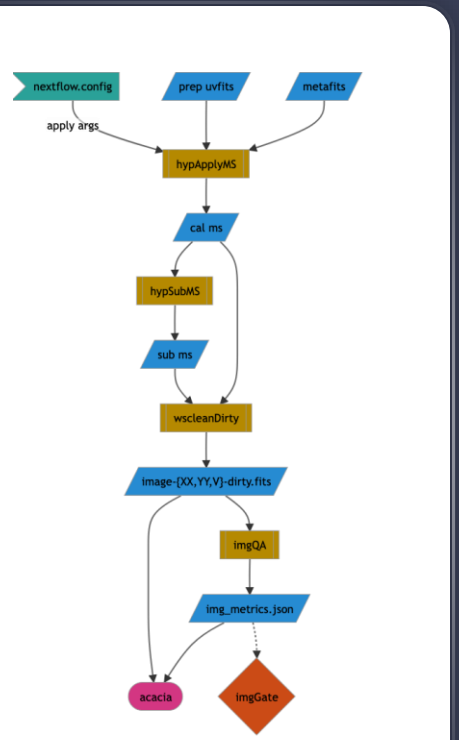


Image analysis



Pipeline Components

Gates
Processing component
Configuration
Cached intermediate product
Object store

Technologies

Nextflow:

- Expressive, flexible, scalable pipeline definition language
- Execution is abstracted, runs on your laptop
- HPC job scheduling, RAM-disk, caching results
- Java ecosystem without Java

Docker images on Singularity:

- Dependency management, HPC execution
- Idempotence

Thank You

Questions?

Slides available on Indico

Nextflow pipeline code:
github.com/MWATelescope/MWAEoR-Pipeline

dev.null@curtin.edu.au