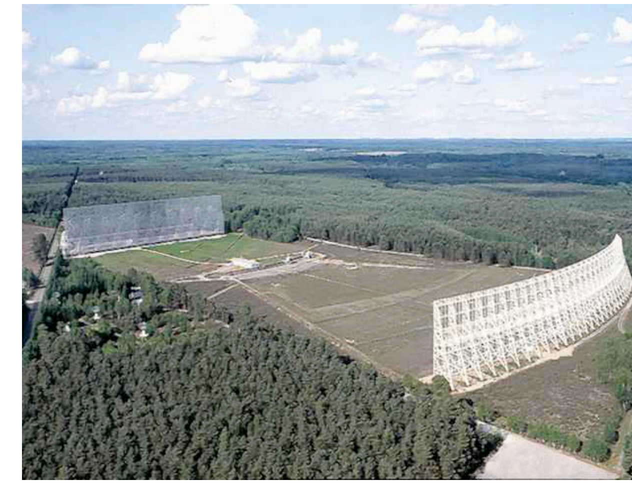
University of Amsterdam

Master Astronomy and Astrophysics - 5214RAAS6Y



Radio Astronomy

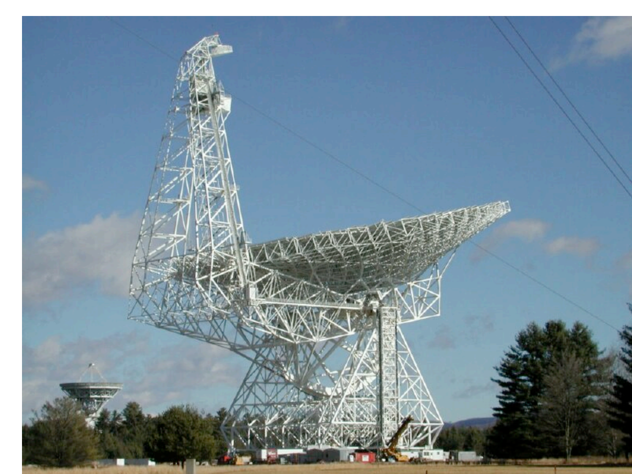
Lecture 7

The Techniques of Radio Interferometry II: Calibration

Lecturer: Michael Wise (wise@astron.nl)

April 22nd, 2013



Outline

- Definition of Calibration
- Visibilities, *uv* Coverage, Gains, Phases
- Real Data, Data Examination, Data Editing
- Formalism, Ideal vs. Real Measurements
- Calibration Strategies and Effectiveness

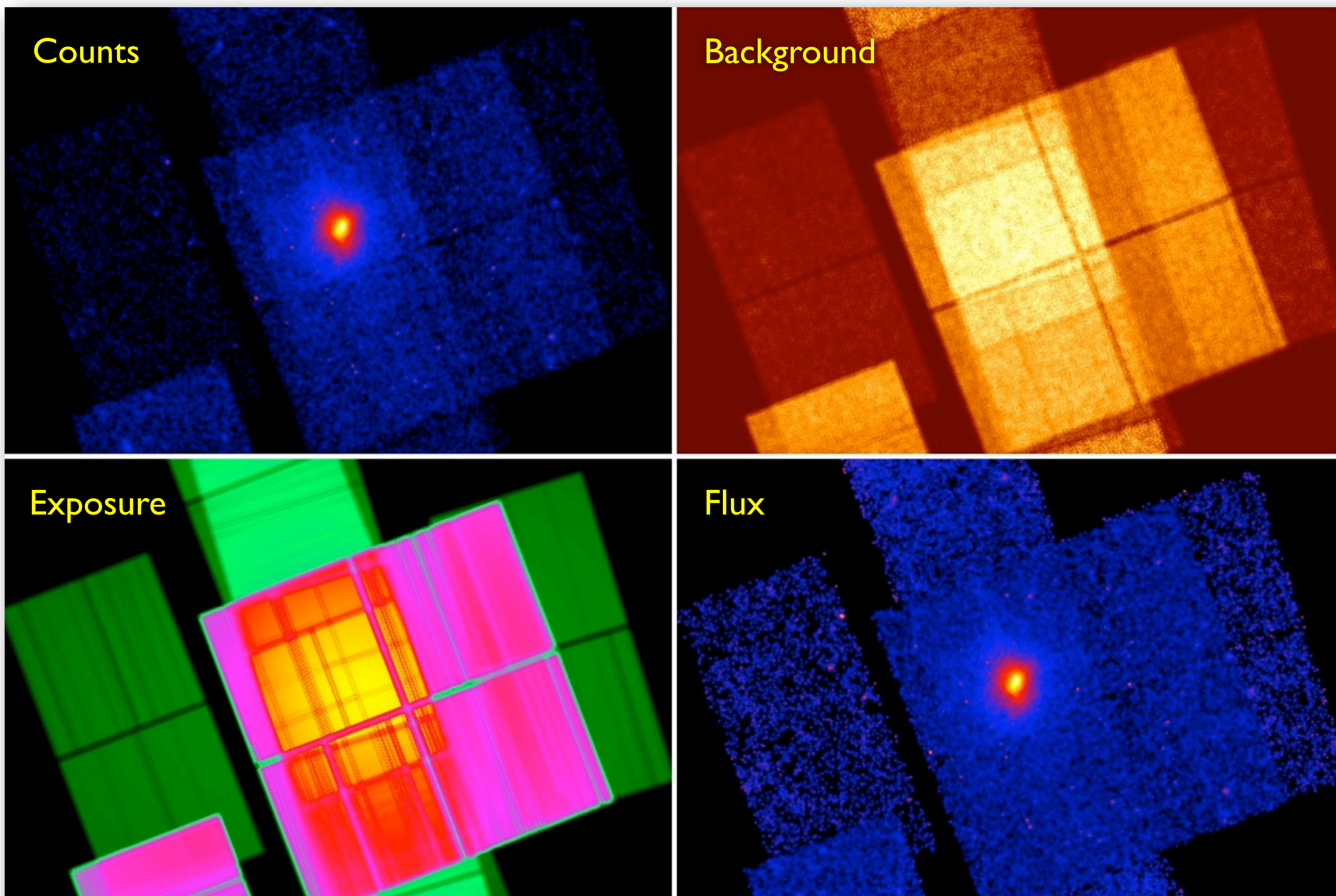
Definition of Calibration

What is Calibration?

to calibrate

"to correlate the readings of an instrument with those of a standard in order to check the instrument's accuracy."

- Oxford English Dictionary



Example of flat-fielding a Chandra X-ray imaging dataset

⇒ *Separate things you care about from the things you don't!*

⇒ *Source properties from instrument and environment.*

Why Calibrate?

- Radio telescopes are not perfect (e.g., surface accuracy, receiver noise, polarization purity, stability, etc.)
- Need to accommodate engineering (e.g., frequency conversion, digital electronics, etc.)
- Hardware or control software occasionally fails or behaves unpredictably
- Scheduling/observation errors sometimes occur (e.g., wrong source positions)
- Atmospheric conditions not ideal (not limited to "bad" weather, especially important at low frequencies)
- Radio Frequency Interference (RFI)
- Contamination from other sources (especially at low frequencies)

Determining instrumental properties (calibration)
⇒ Prerequisite to determining source properties

Types of Calibration

- **A priori "calibrations"**
  - Information provided by the observatory
  - Antenna positions, earth orientation and rate, clocks
  - Antenna pointing, voltage pattern, gain curve
  - Calibrator coordinates, flux densities, polarization properties
- **Cross-calibration**
  - Observe strong nearby sources against which calibration can be solved, and transfer solutions to target observations
  - Choose appropriate calibrators, usually point sources because we can easily predict their visibilities (Amplitude ~ constant, Phase ~ 0)
  - Choose appropriate timescales for calibration
- **Self-calibration**
  - Correct for antenna based phase and amplitude errors together with imaging
  - Iterative, non-linear relaxation process
  - Requires sufficient signal-to-noise at each solution interval
  - Dangerous with small N arrays, complex sources, low signal-to-noise

Astronomical Calibrations

- **Flux Density Calibration**
  - Radio astronomy flux density scale set according to several "constant" radio sources, and planets/moons
  - Use resolved models where appropriate
- **Astrometry**
  - Most calibrators come from astrometric catalogs; sky coordinate accuracy of target images tied to that of the calibrators
  - Beware of resolved and evolving structures, and phase transfer biases due to troposphere (especially for VLBI)
- **Linear Polarization Position Angle**
  - Usual flux density calibrators also have significant stable linear polarization position angle for registration
- Relative calibration solutions (and dynamic range) insensitive to errors in these "scaling" parameters

Visibility Data

Review of Visibilities

- We DEFINE a complex function, the complex visibility, V, from the two independent (real) correlator outputs R_C and R_S:

$$V = R_C - iR_S = Ae^{-i\phi}$$

where

$$A = \sqrt{R_C^2 - R_S^2} \qquad \phi = \tan^{-1}\left(\frac{R_S}{R_C}\right)$$

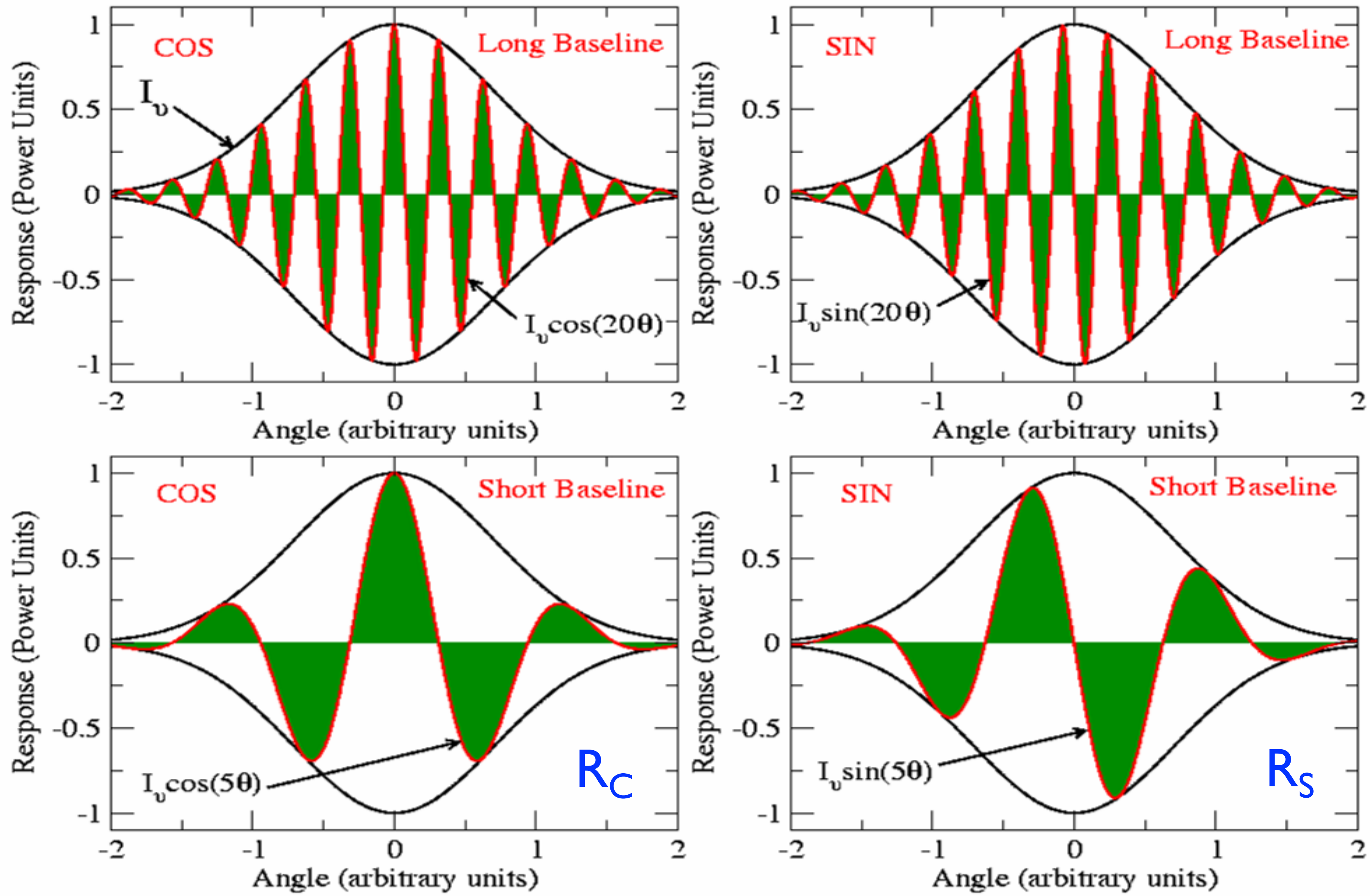
- This gives us a beautiful and useful relationship between the source brightness, and the response of an interferometer:

$$V_\nu(\vec{b}) = R_C - iR_S = \iint I_\nu(\vec{s})e^{-2\pi i\nu\vec{b}\cdot\vec{s}/c}d\Omega$$

- Under some circumstances, this is a 2-D Fourier transform, giving us a well established way to recover *I*(**s**) from *V*(**b**).

Visualizing Visibilities

- The source brightness is Gaussian, shown in black
- The interferometer 'fringes' are in red
- The visibility is the integral of the product (net dark green area)



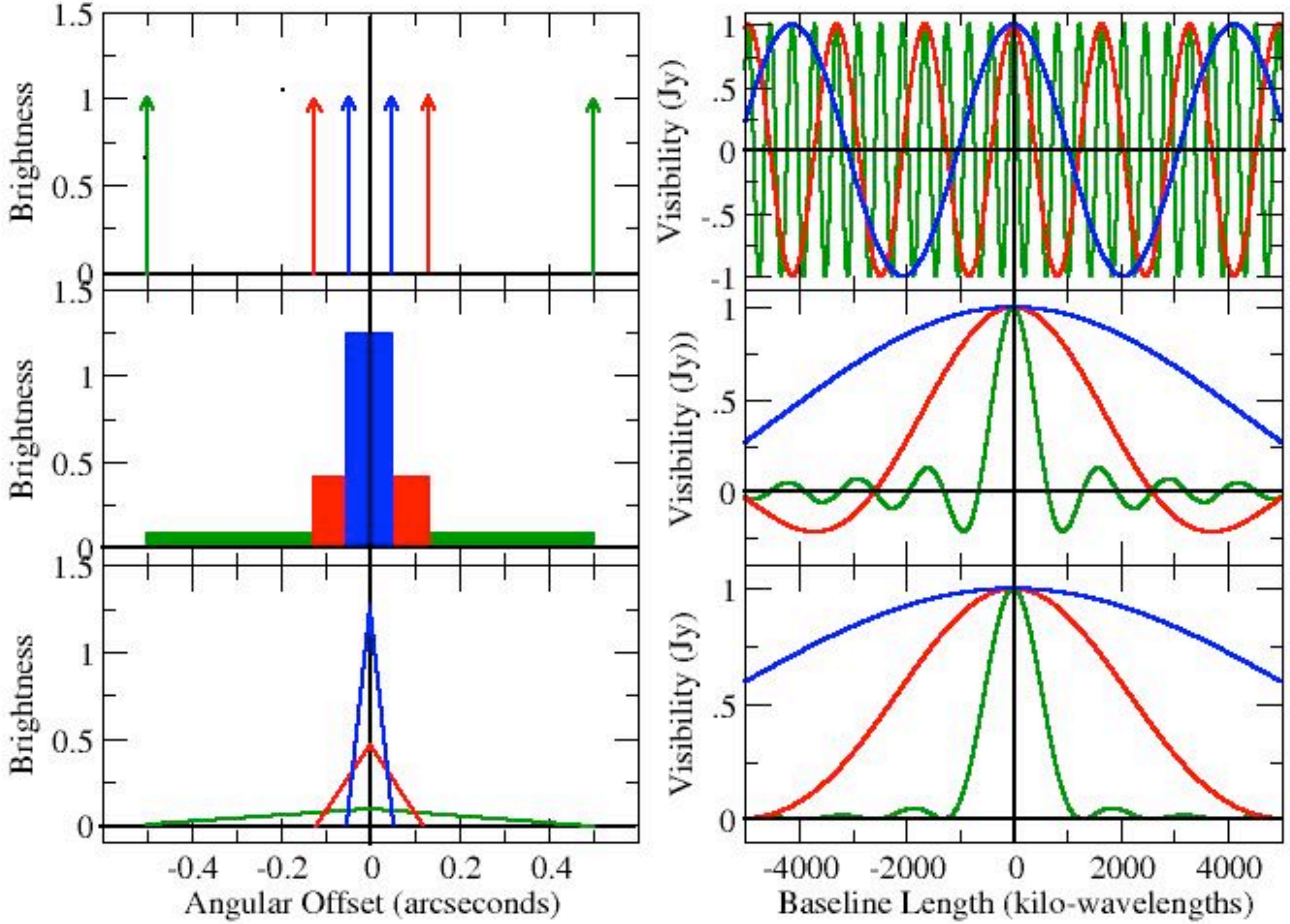
Simple Visibility Functions

Brightness Distribution | Visibility Function

Unresolved Doubles

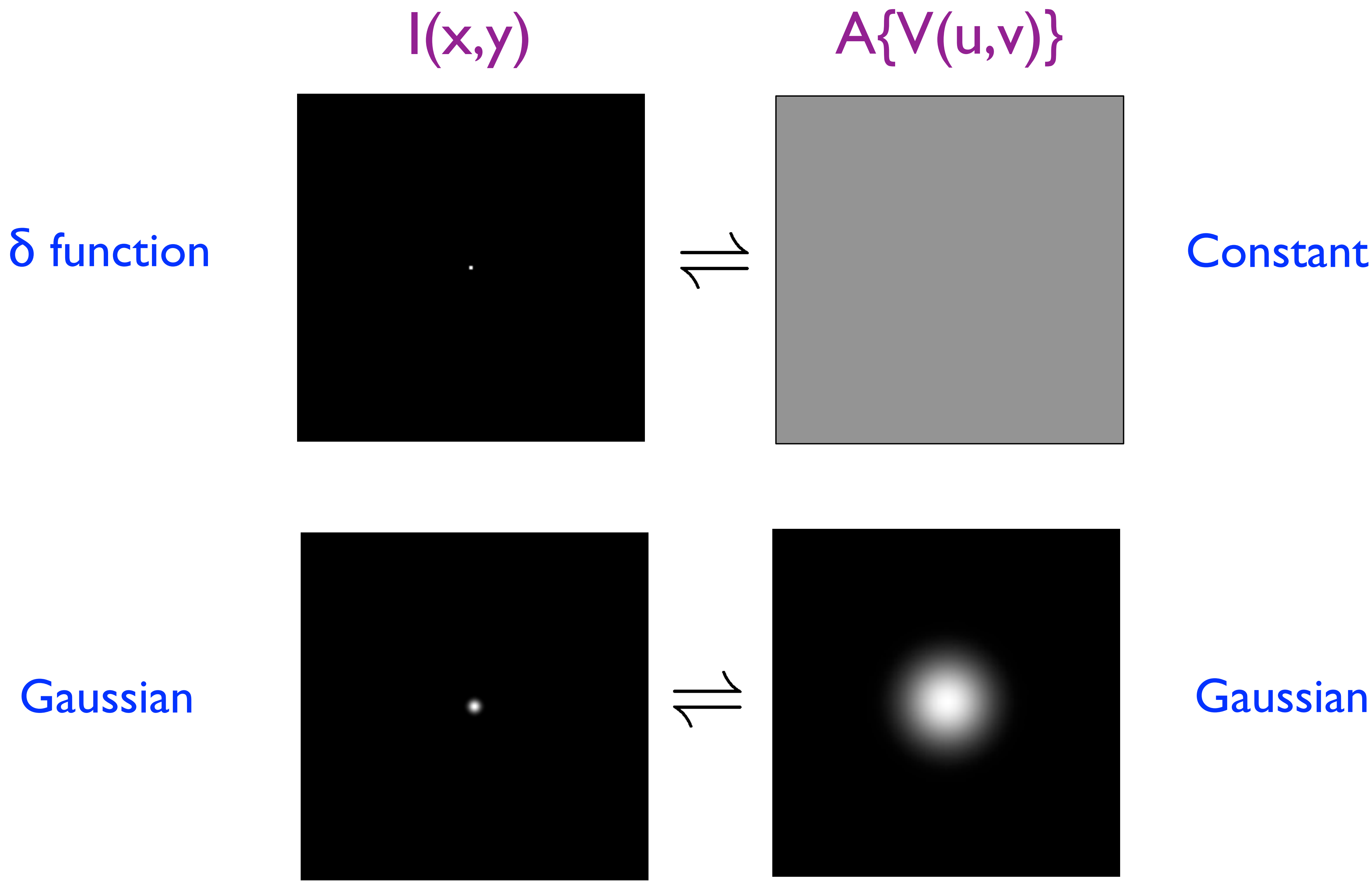
Resolved Uniform

Centrally Peaked



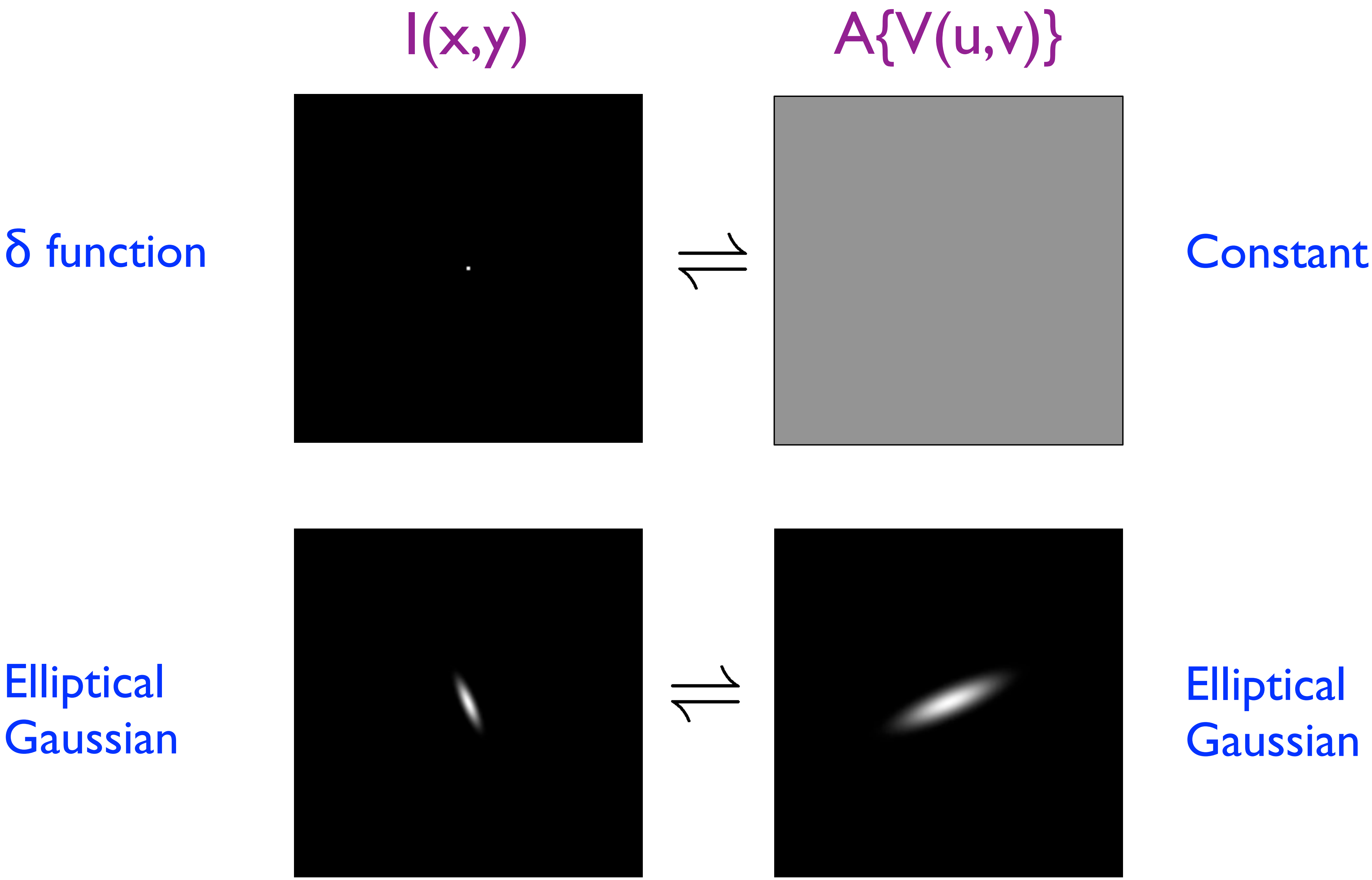
Fourier Transform of *I(x,y)*

- V(u,v) is a complex quantity expressed as (real, imaginary) or (amplitude, phase)
- Narrow features transform into wide features (and vice-versa)



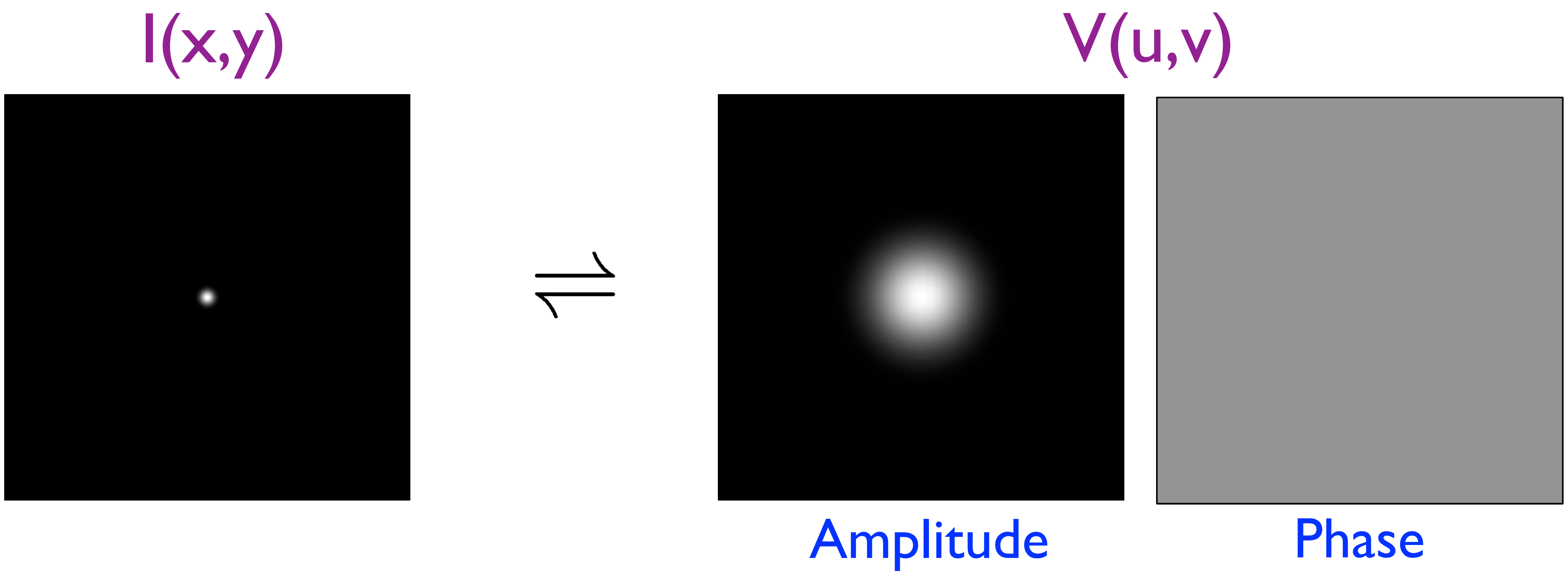
Fourier Transform of *I(x,y)*

- V(u,v) is a complex quantity expressed as (real, imaginary) or (amplitude, phase)
- Narrow features transform into wide features (and vice-versa)



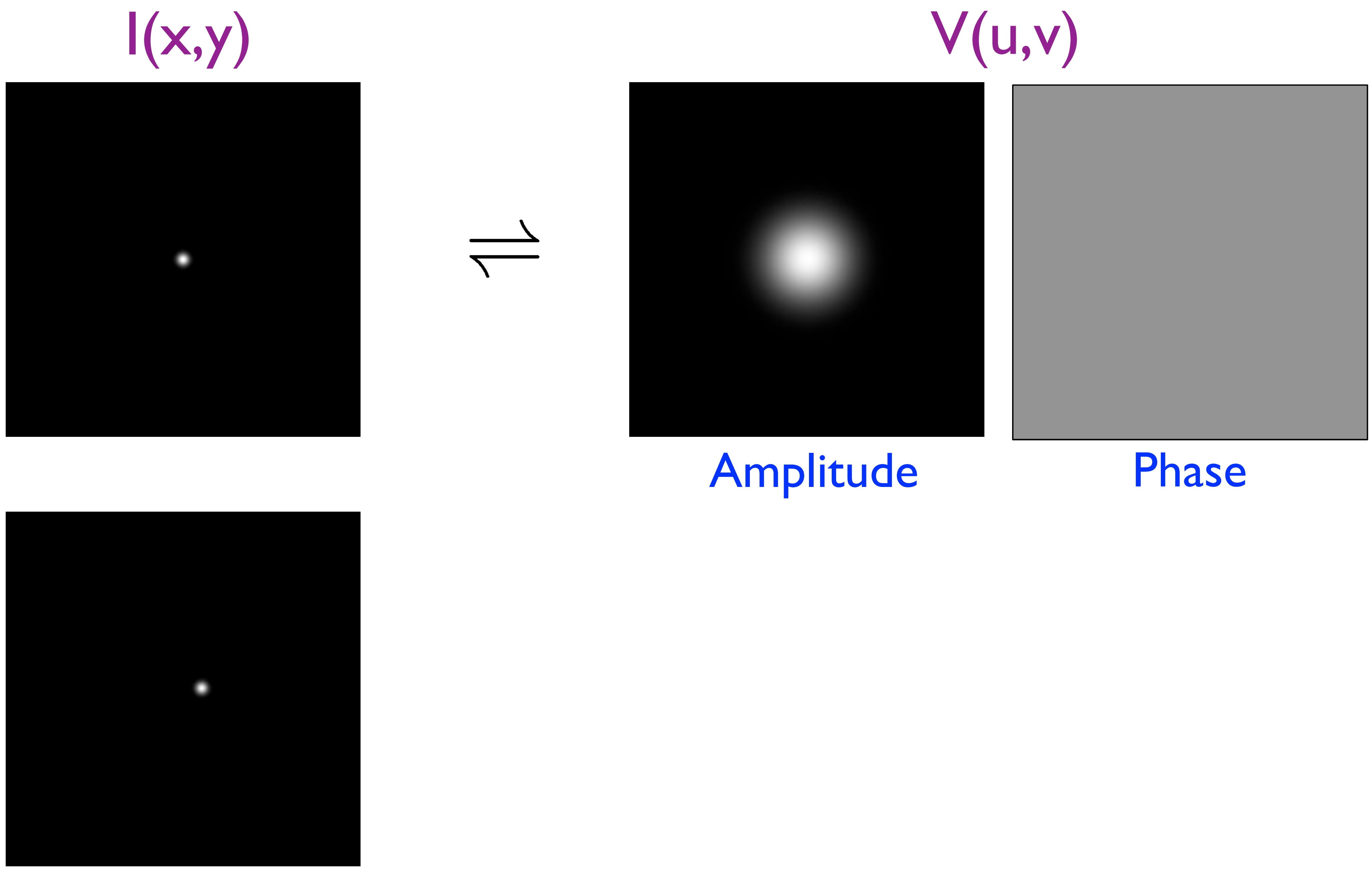
Amplitude and Phase

- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



Amplitude and Phase

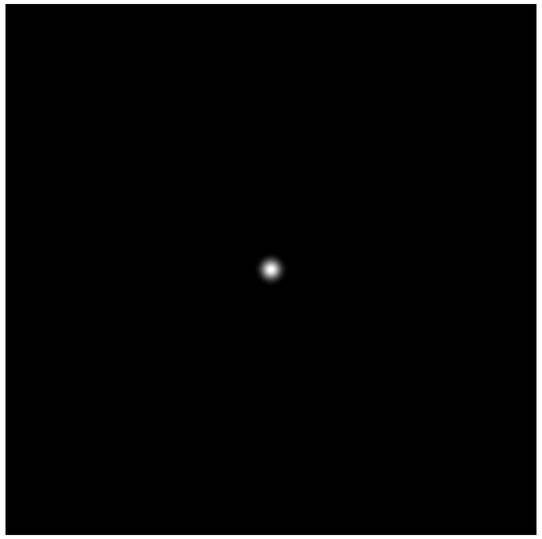
- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



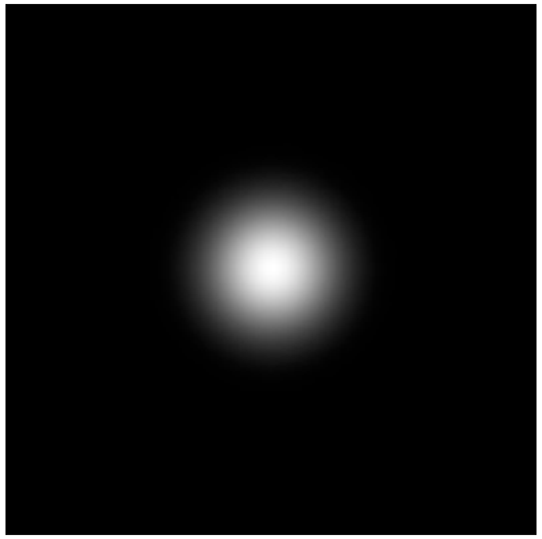
Amplitude and Phase

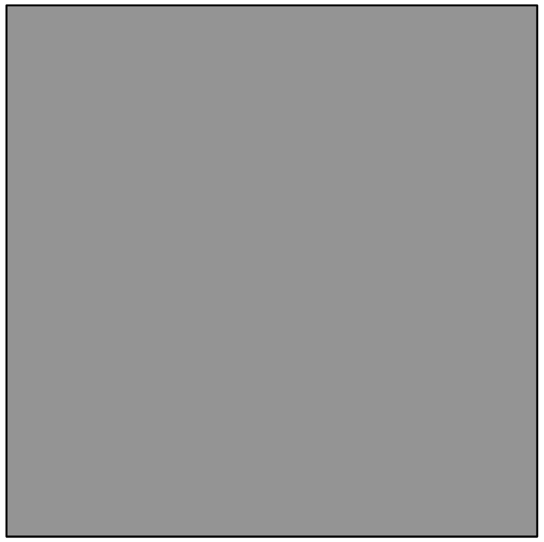
- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located

I(x,y) V(u,v)

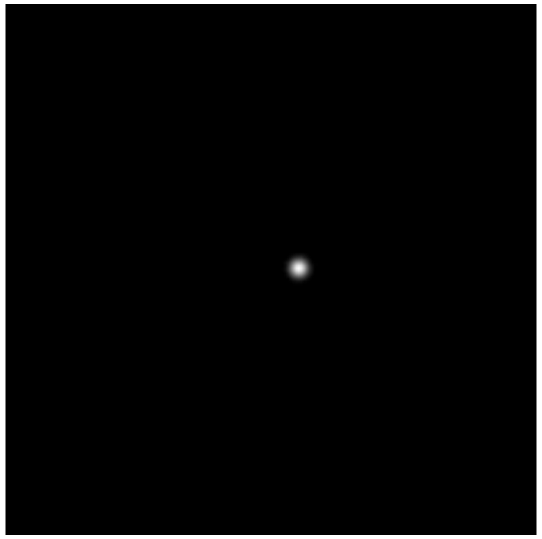


⇌





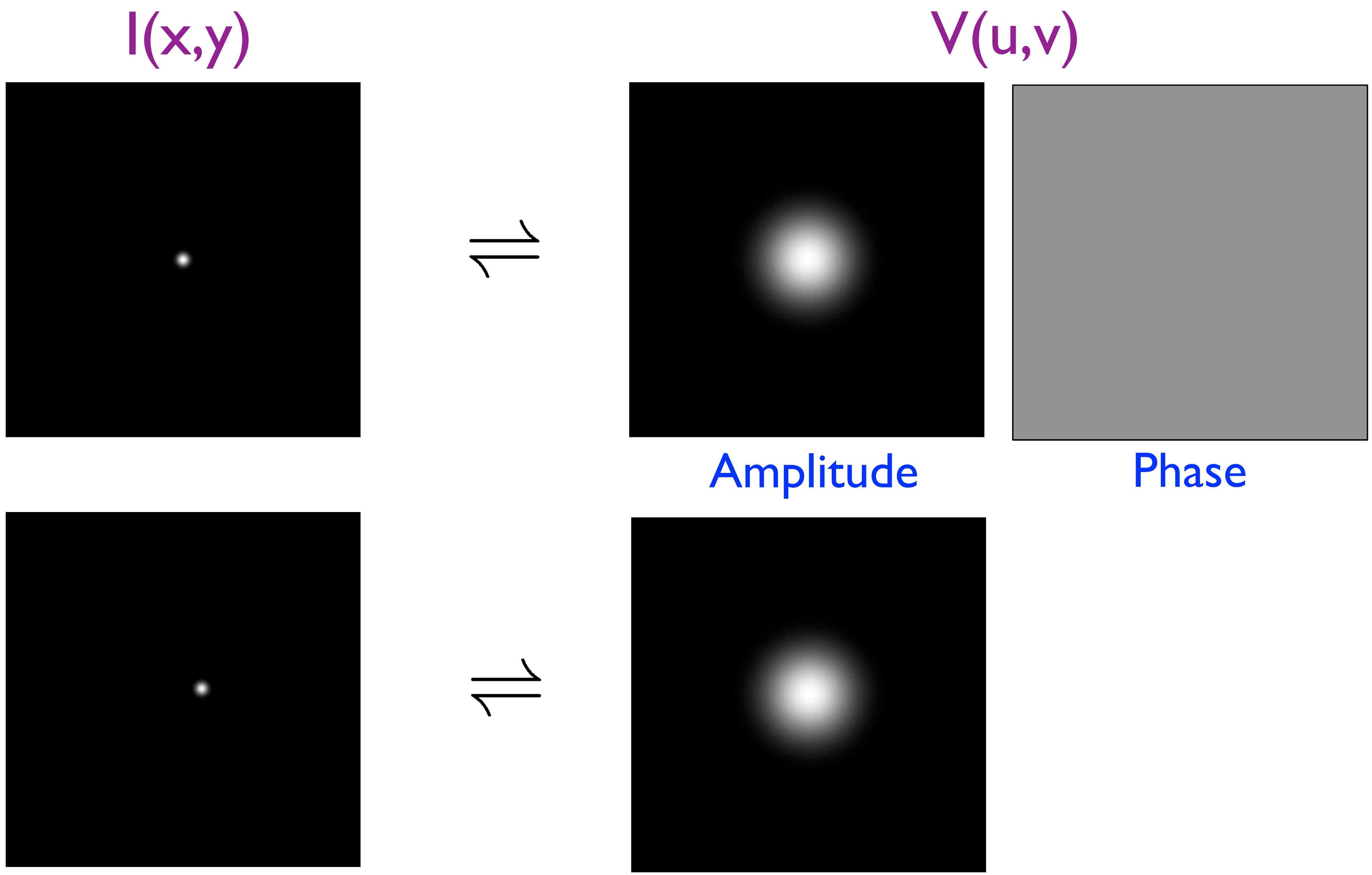
Amplitude Phase



⇌

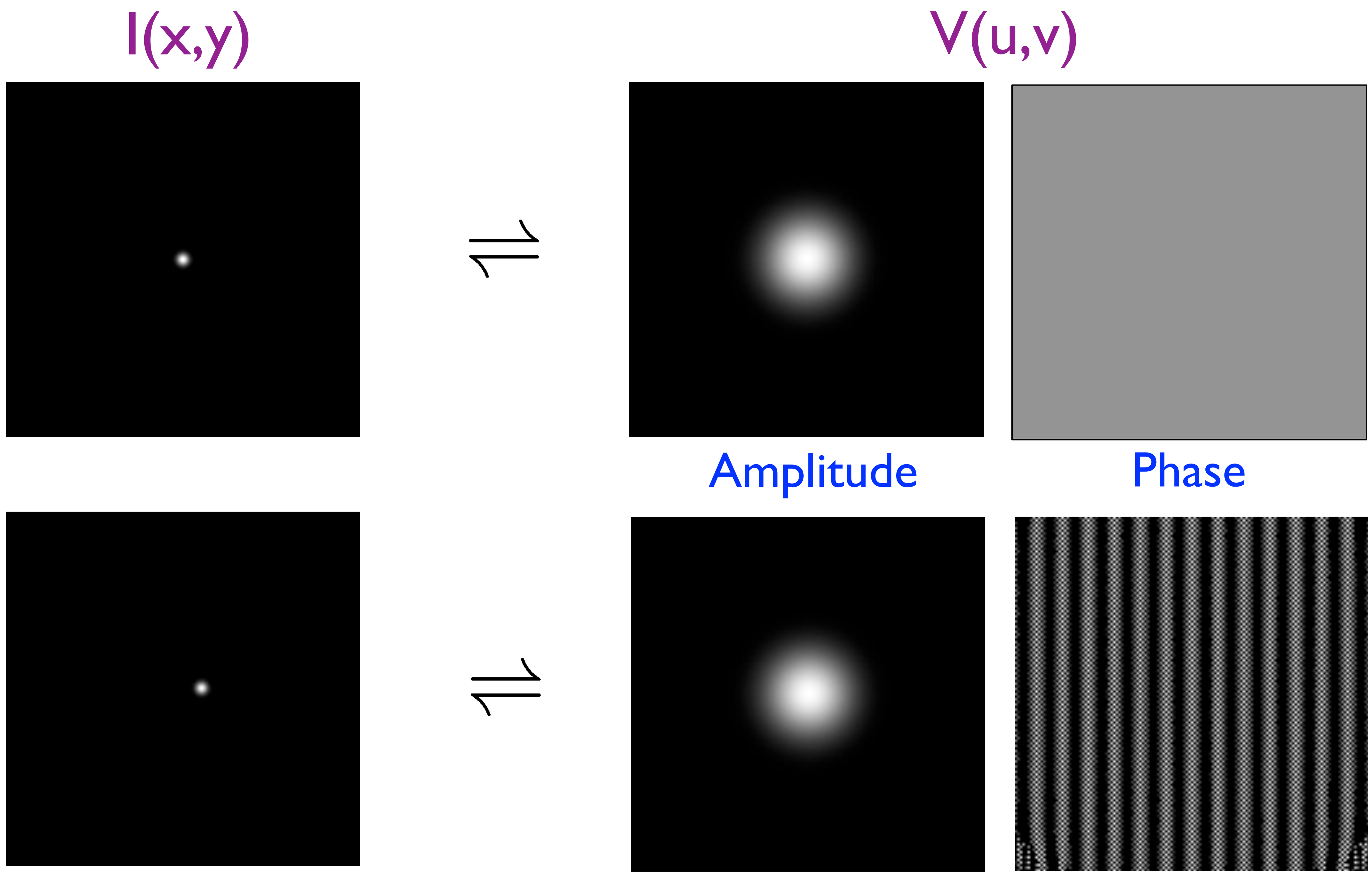
Amplitude and Phase

- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



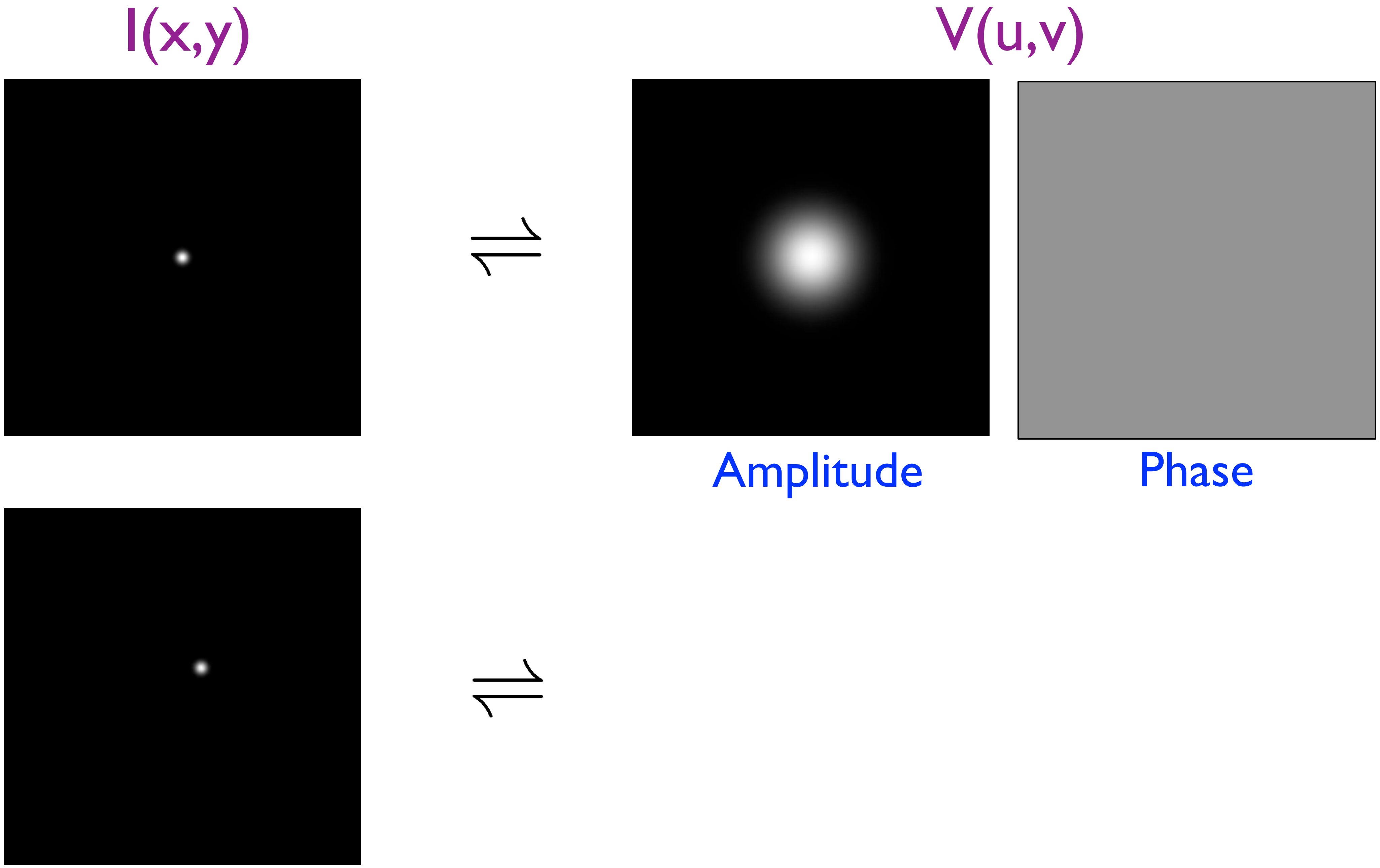
Amplitude and Phase

- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



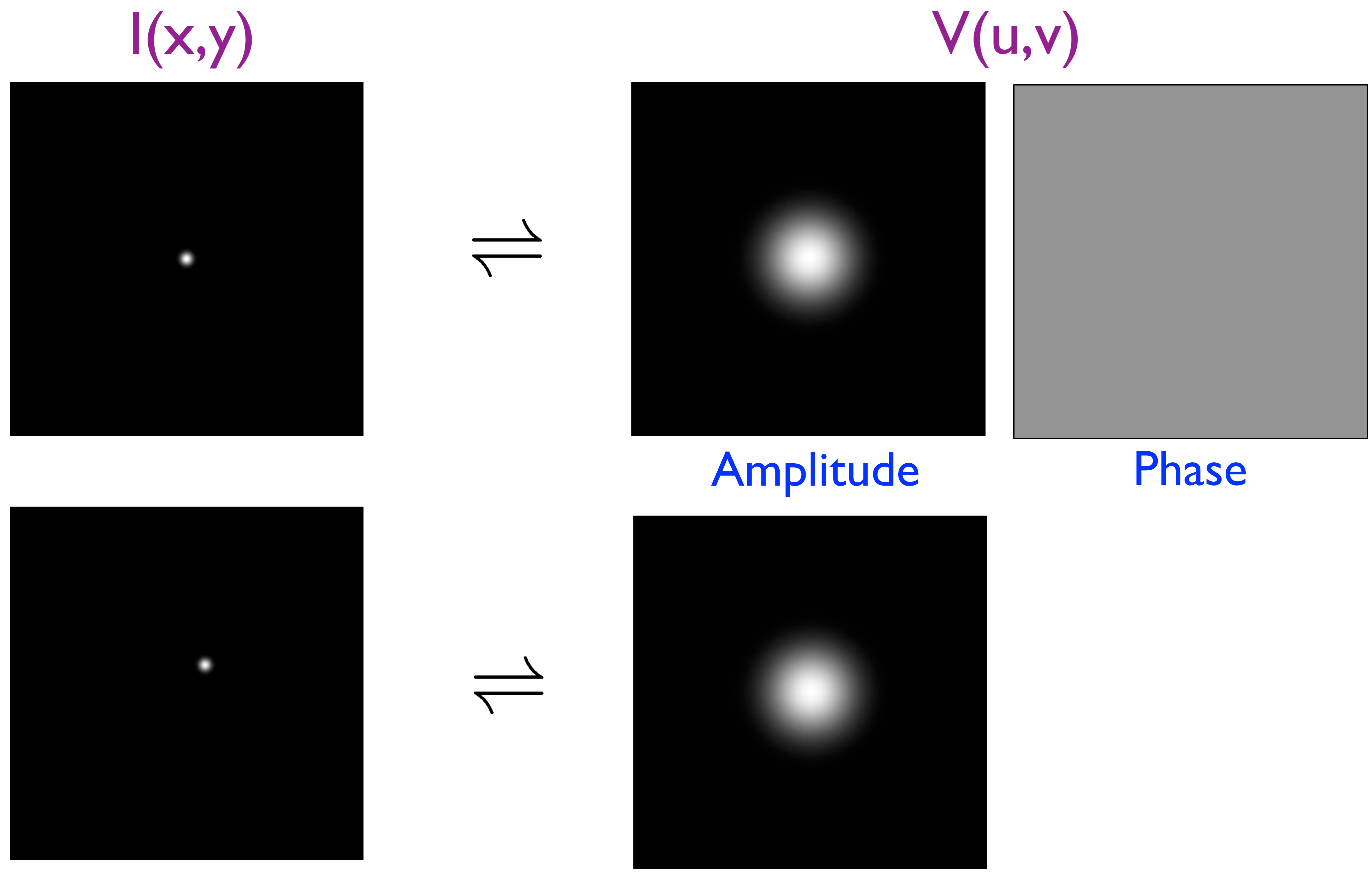
Amplitude and Phase

- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



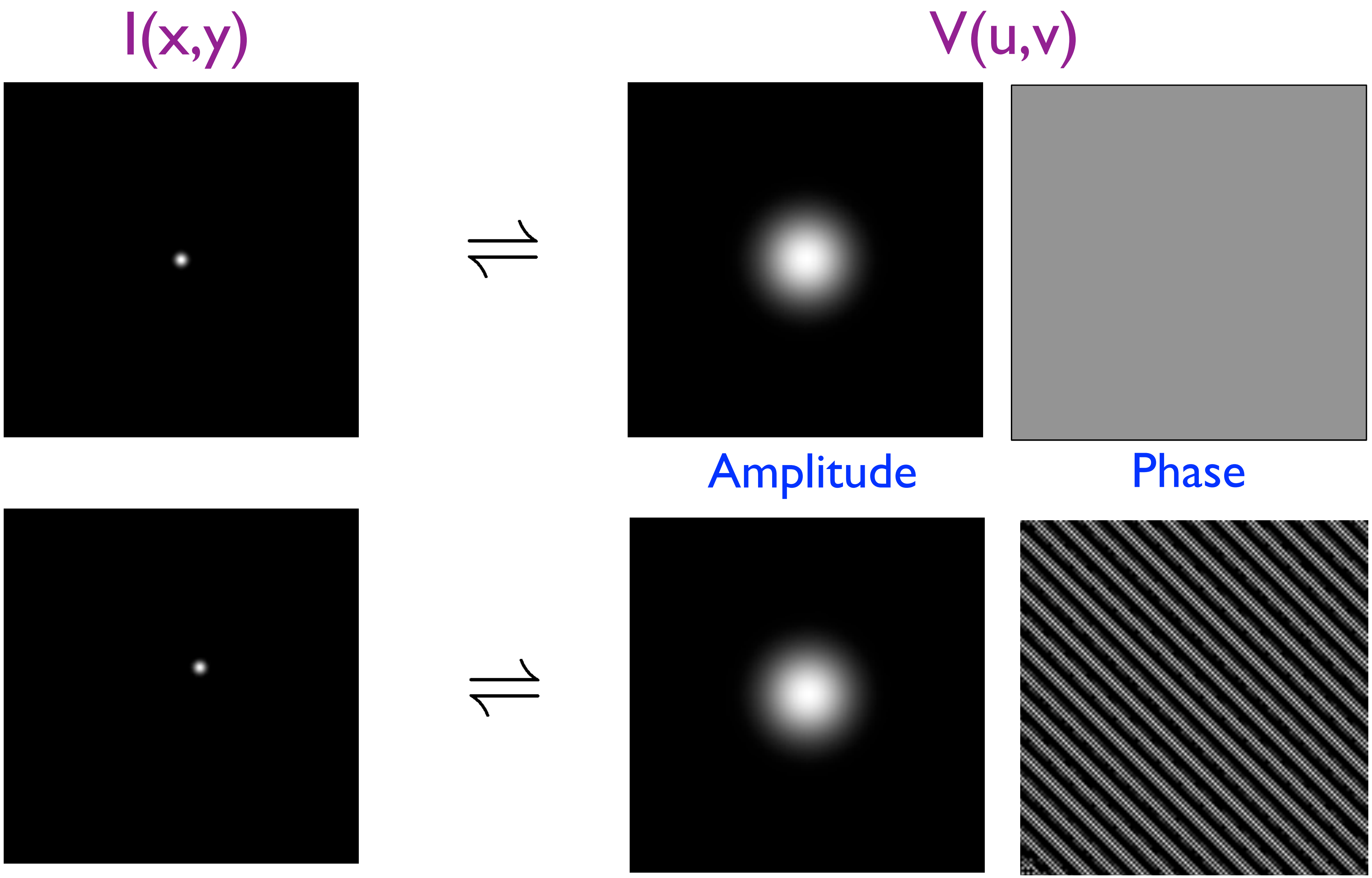
Amplitude and Phase

- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



Amplitude and Phase

- Amplitude tells "how much" of a certain spatial frequency
- Phase tells "where" this component is located



uv Coverage and Beams

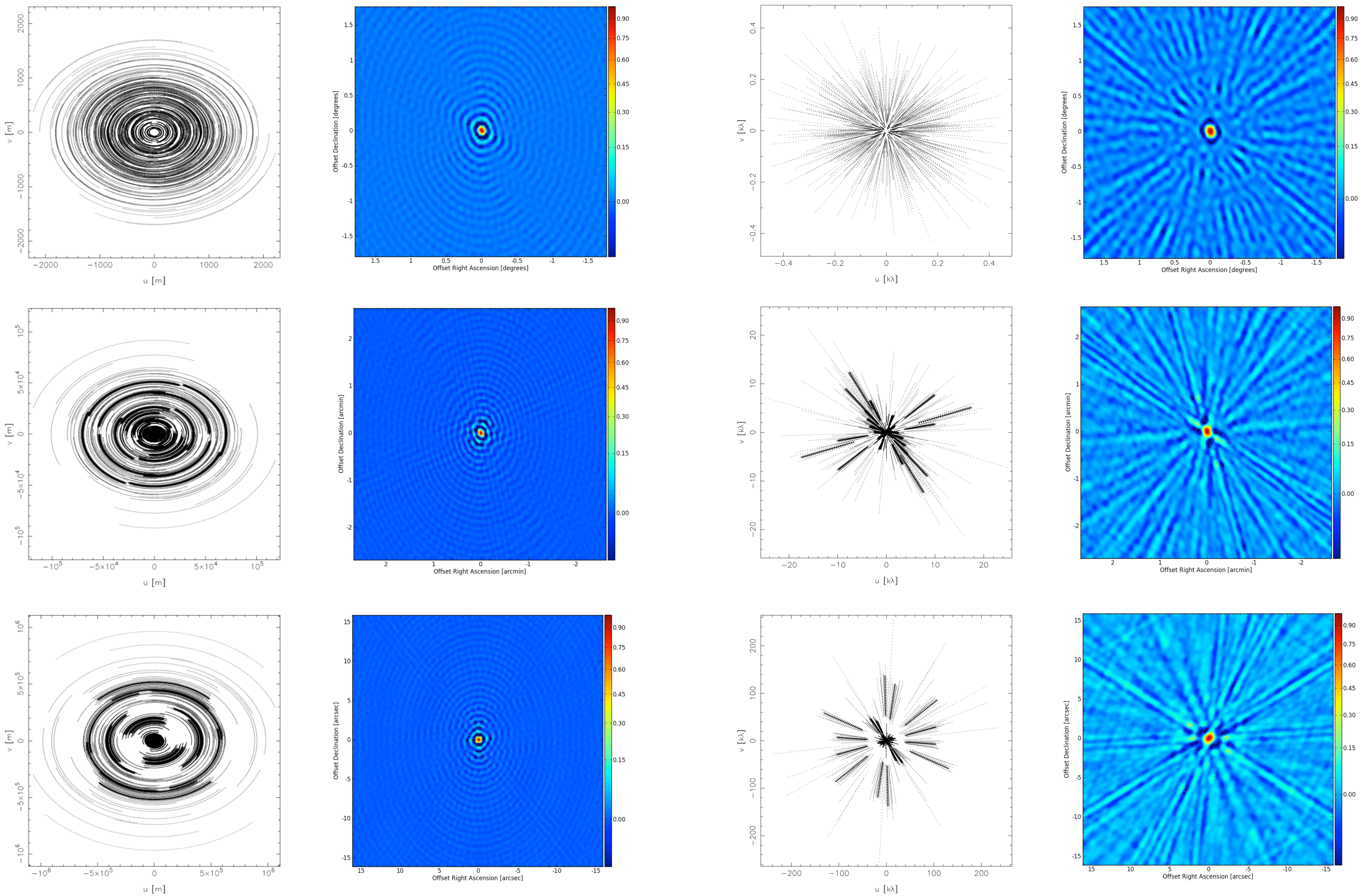
6 hr track | *Instantaneous*

LOFAR

10 km

100 km

1000 km



Data, Examination, Editing

What Data is Delivered?

- An *enormous* list of complex visibilities! *(Enormous!)*
  - At each timestamp (~1-10s intervals): N(N-1)/2 baselines
    - EVLA: 351baselines
    - VLBA: 45 baselines
    - ALMA: 1225-2016 baselines
    - LOFAR: 1128 (LBA), 2016 (HBA), 41328 (AARTFAAC)
  - For each baseline: 64-256 Spectral Windows ("spws", "subbands" or "IFs")
  - For each spectral window: tens to thousands of channels
  - For each channel: 1, 2, or 4 complex correlations (polarizations)
    - EVLA or VLBA: RR or LL or (RR,LL), or (RR,RL,LR,LL)
    - ALMA or LOFAR: XX or YY or (XX,YY) or (XX,XY,YX,YY)
  - With each correlation, a weight value and a flag (T/F)
  - Meta-info: Coordinates, antenna, field, frequency label info
- $N_{total} = N_t \times N_{bl} \times N_{spw} \times N_{chan} \times N_{corr}$ visibilities

$\Rightarrow$ *10s of GB to 10s of TBs of visibility data*

Data Contents

- Usually presented to astronomer as $V_{ij}(\nu,t)$
  - Cross (and auto) correlation spectra
  - Sampled at visibility dump time, integration time
- Metadata information needed for calibration and processing
  - IF labels, and polarizations
  - Time tags
  - frequency information, edge and increment
  - Antenna indexes
  - *uvw* coordinates
  - Telescope pointing and source labeling
- Format for transport: FITS, Measurement Set (MS), HDF5
  - Standard formats, but content not standardized
  - But calibration software depends critically on content

LOFAR Data Volumes

- 2688 dipoles (LBA), 200 MHz sampling, 2 polarizations, 12 bit digitization

  ⇒ 13 Tbits/s ~ 1.6 TB/s ~ 138 PB/day

- 1128 baselines, 242 sub-bands, 256 channels, 4 pol., 1 sec correlator dump-time

  ⇒ ~ 10 TB/hr ~ 240 TB/day ~ 0.1 EB/yr

- 10° x 10° FoV, 2.0″ resolution, 1.0″ pixels (HBA, NL baselines only)

  ⇒ ~ 10^9 pixels ~ 5 Gbytes / frequency

- 10° x 10° FoV, 0.2″ resolution, 0.1″ pixels (HBA, including longest baselines)

  ⇒ ~ 10^{11} pixels ~ 500 Gbytes / frequency

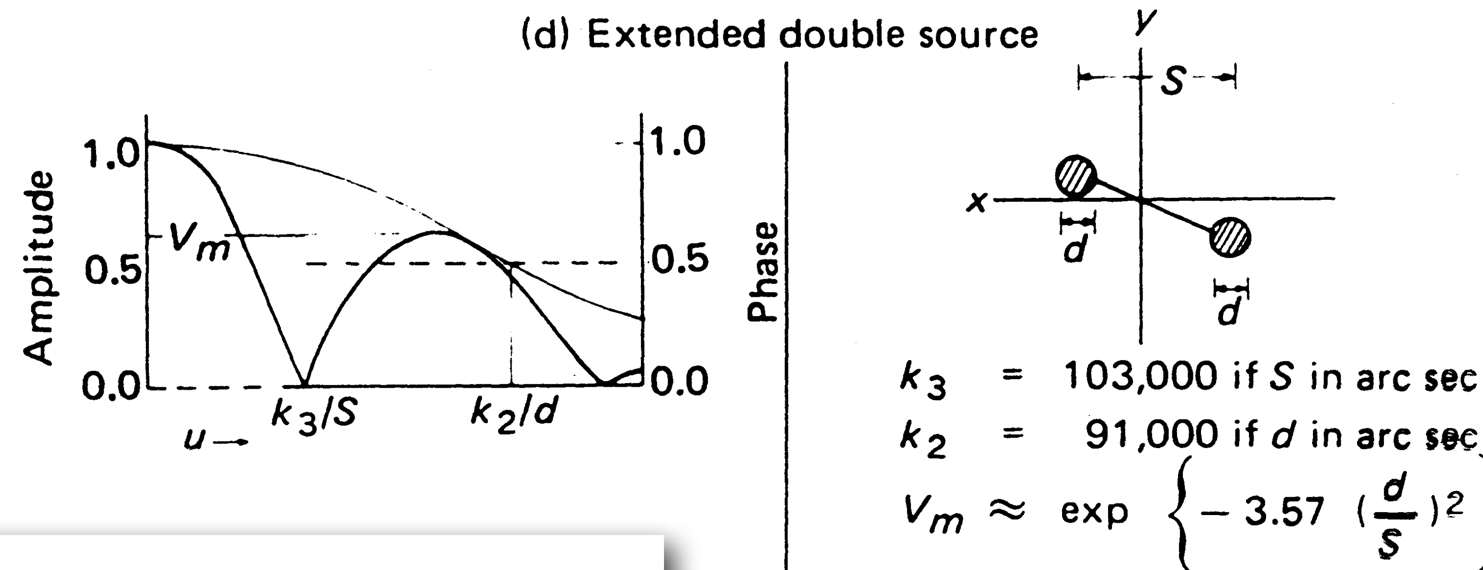
Storage limits give a ~1 week processing window

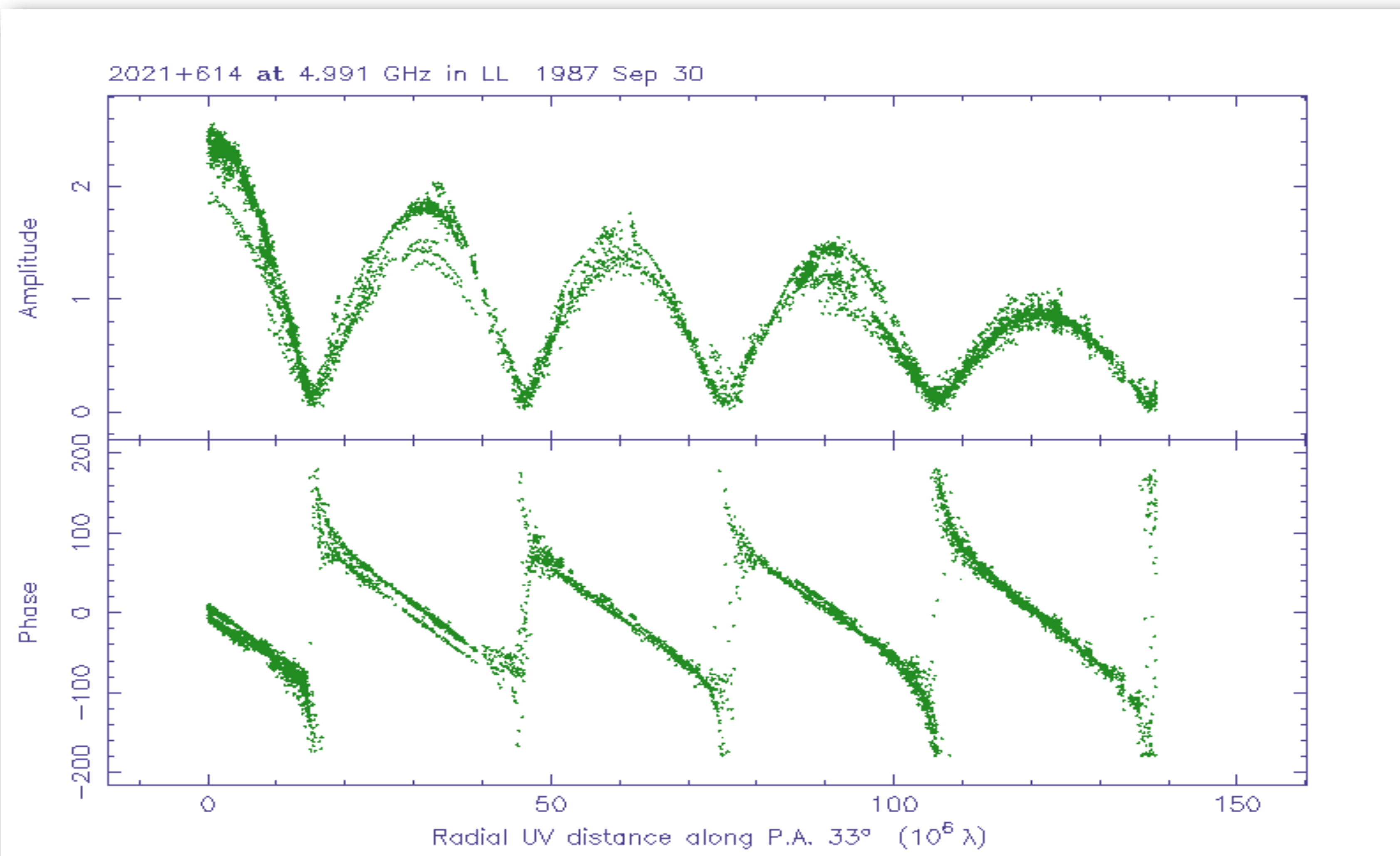
Inspecting Visibility Data

- **Useful visualizations**
  - Sampling of the (u,v) plane
  - Amplitude and phase vs. radius in the (u,v) plane
  - Amplitude and phase vs. time on each baseline
  - Amplitude variation across the (u,v) plane
  - Projection onto a particular orientation in the (u,v) plane
- **Advantages to inspecting *uv* data**
  - Insufficient (u,v)-plane coverage to make an image
  - Inadequate calibration
  - Quantitative analysis
  - Direct comparison of two data sets
  - Noise is uncorrelated in the (u,v) plane but correlated in the image
  - Systematic errors are usually localized in the (u,v) plane

Estimating Initial Model

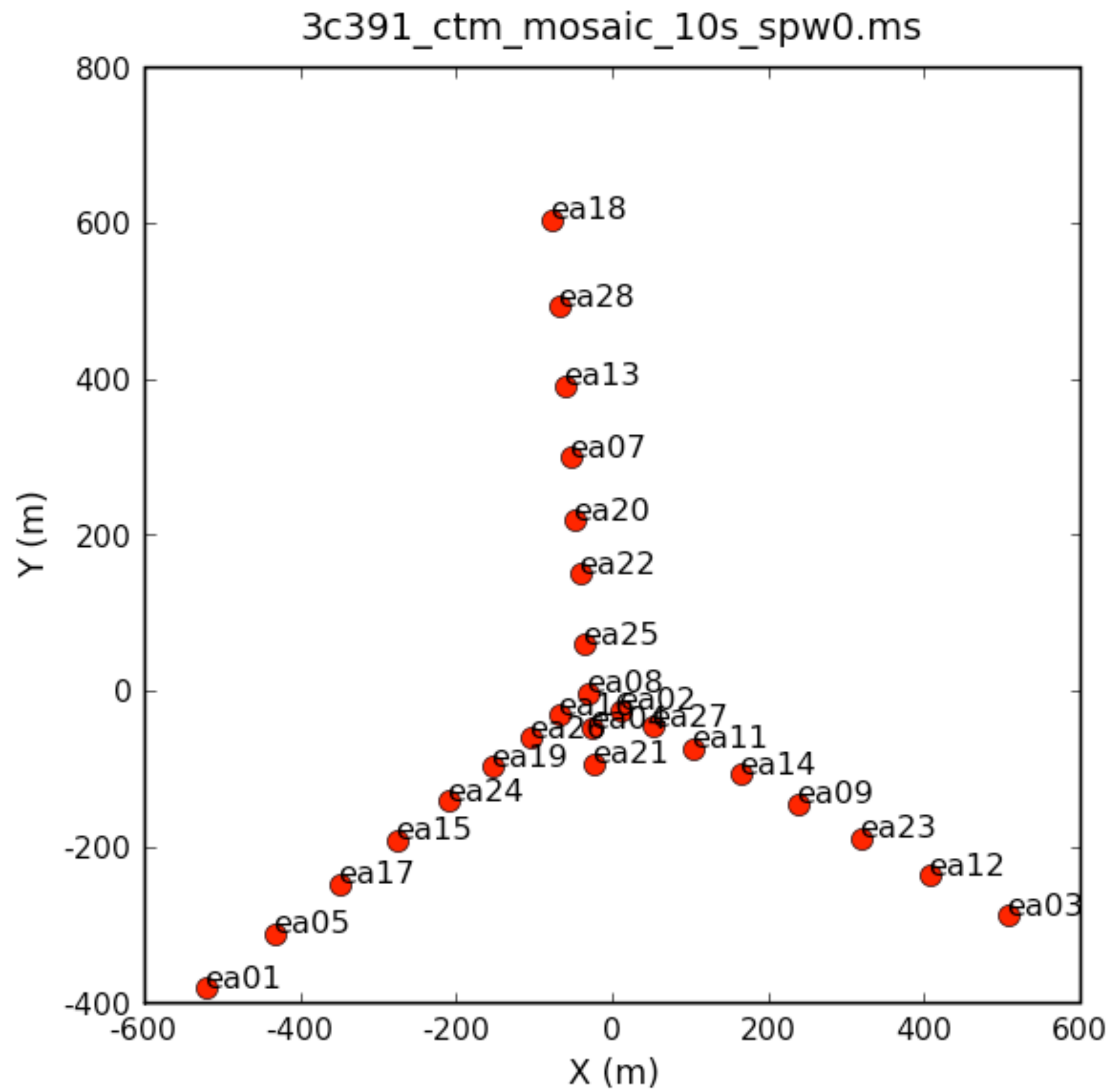
Can derive an initial model by inspection!





Typical Dataset (VLA)

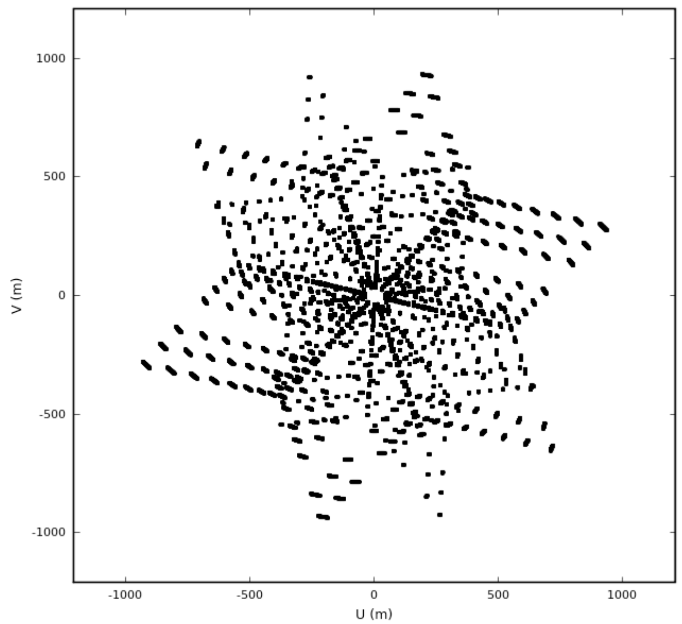
- Array:
  - EVLA D-configuration (Apr 2010)
- Sources:
  - Science Target: 3C391 (7 mosaic pointings)
  - Near-target calibrator: J1822-0938 (~11 deg from target)
  - Flux Density calibrator: 3C286
  - Instrumental Polarization Calibrator: 3c84
- Signals:
  - RR,RL,LR,LL correlations
  - One spectral window centered at 4600 MHz
  - 128 MHz bandwidth, 64 channels

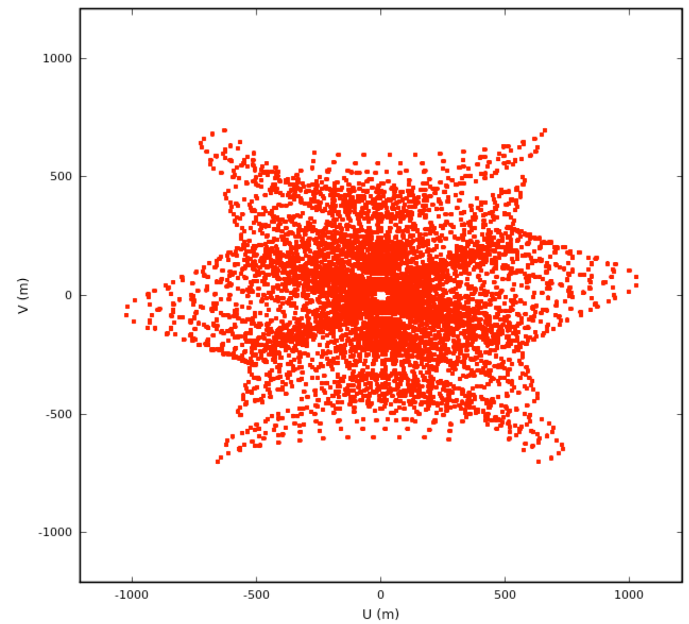


EVLA Antenna Designations

Observed *uv* Coverages

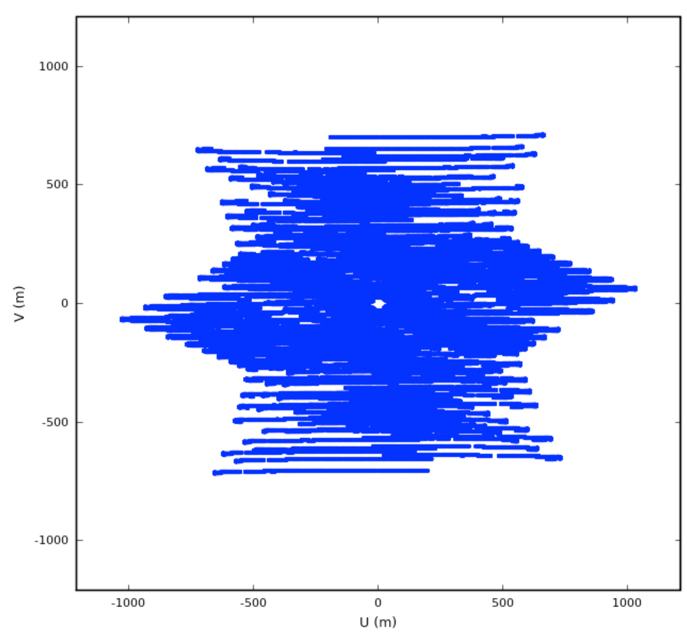
3C286
(Flux Density)

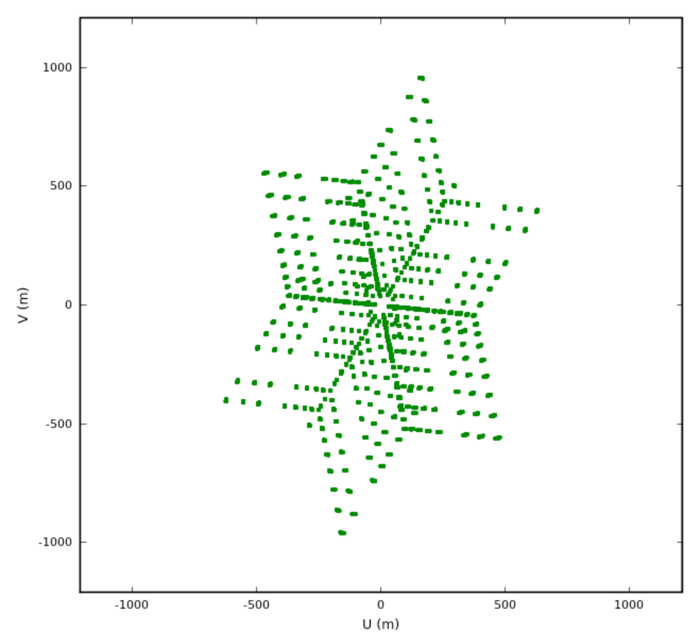




J1822-0938
(Gain Calibrator)

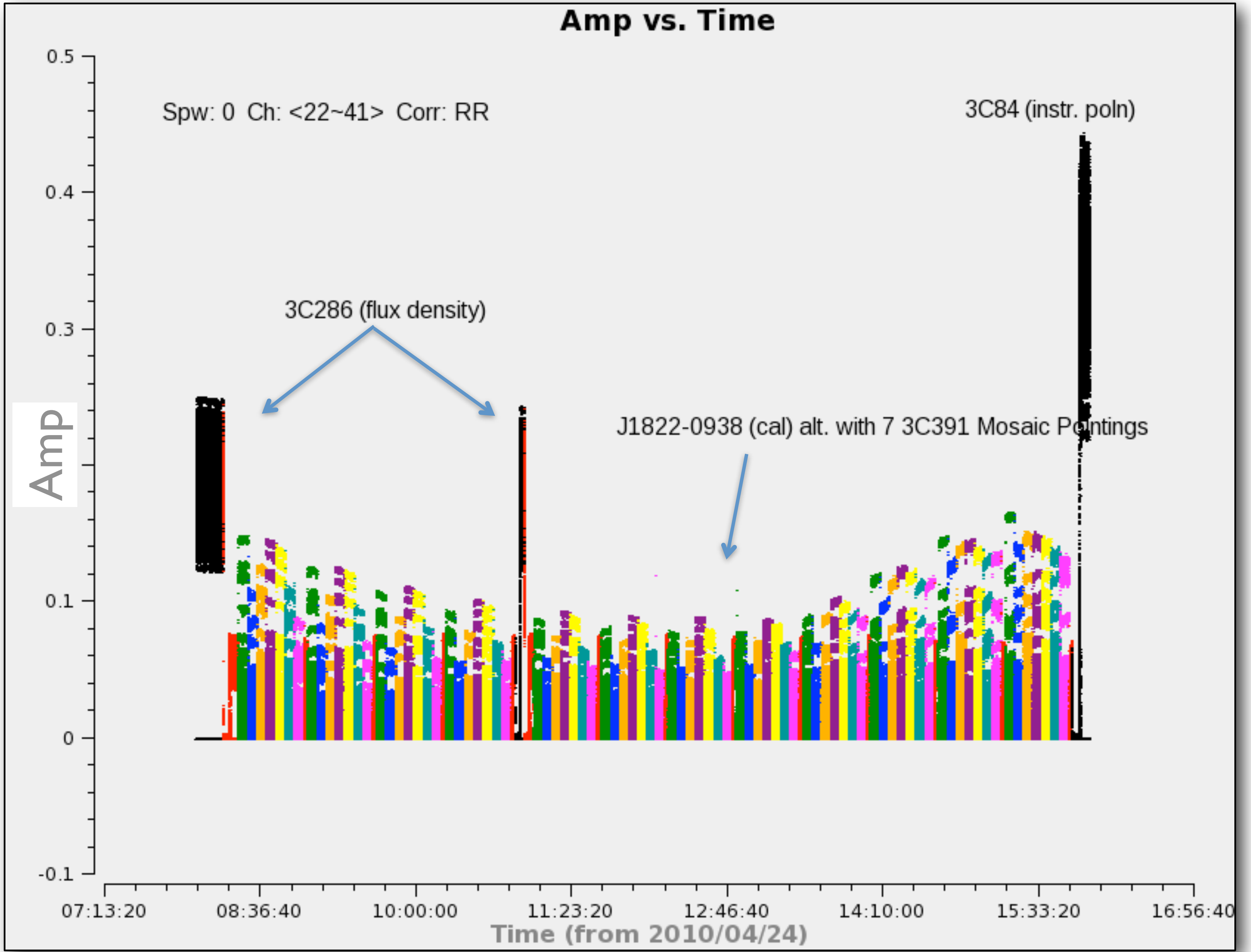
3C391
(Science Target)



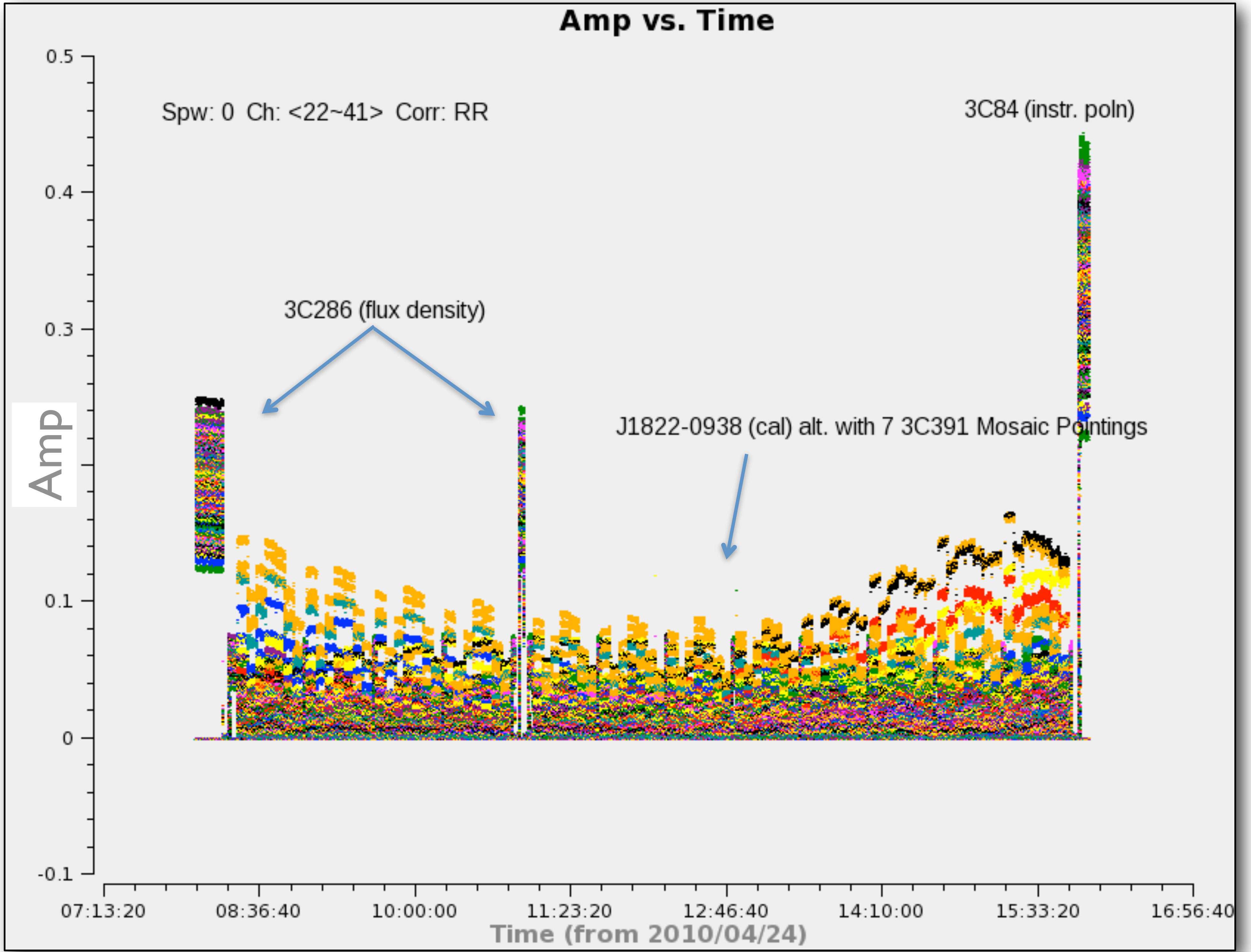


3C84
(Pol. Calibrator)

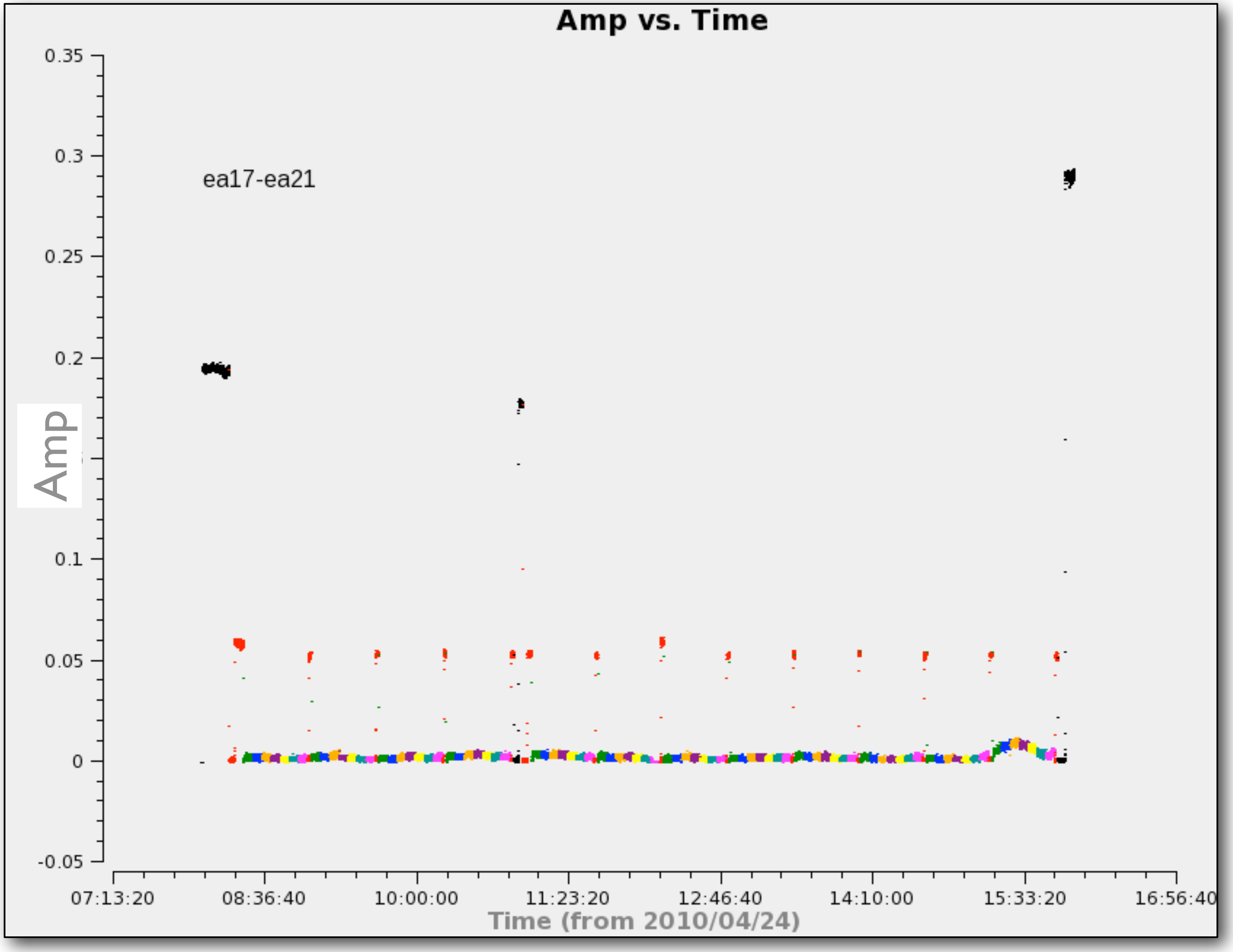
Visibilities (source colors)



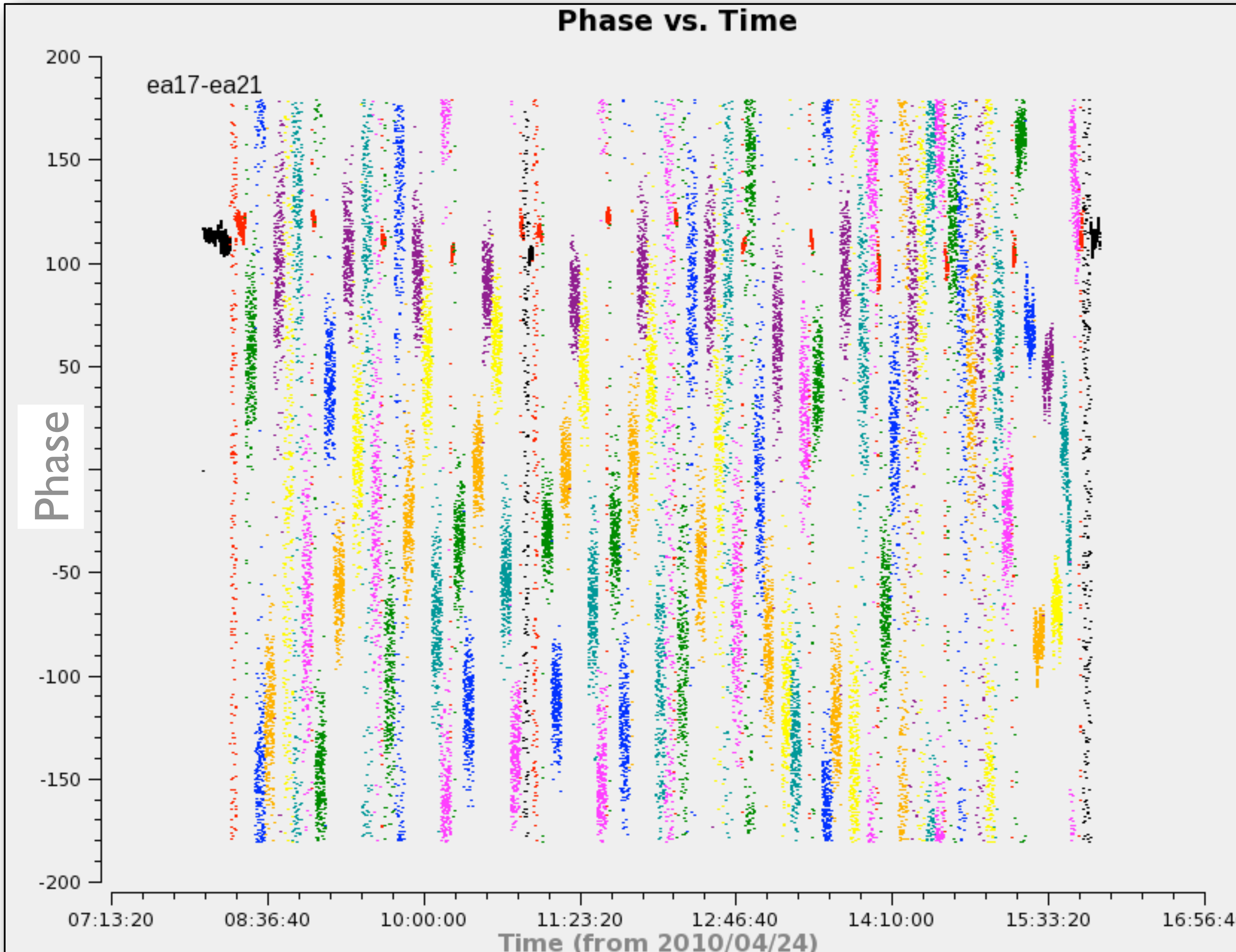
Visibilities (baseline colors)



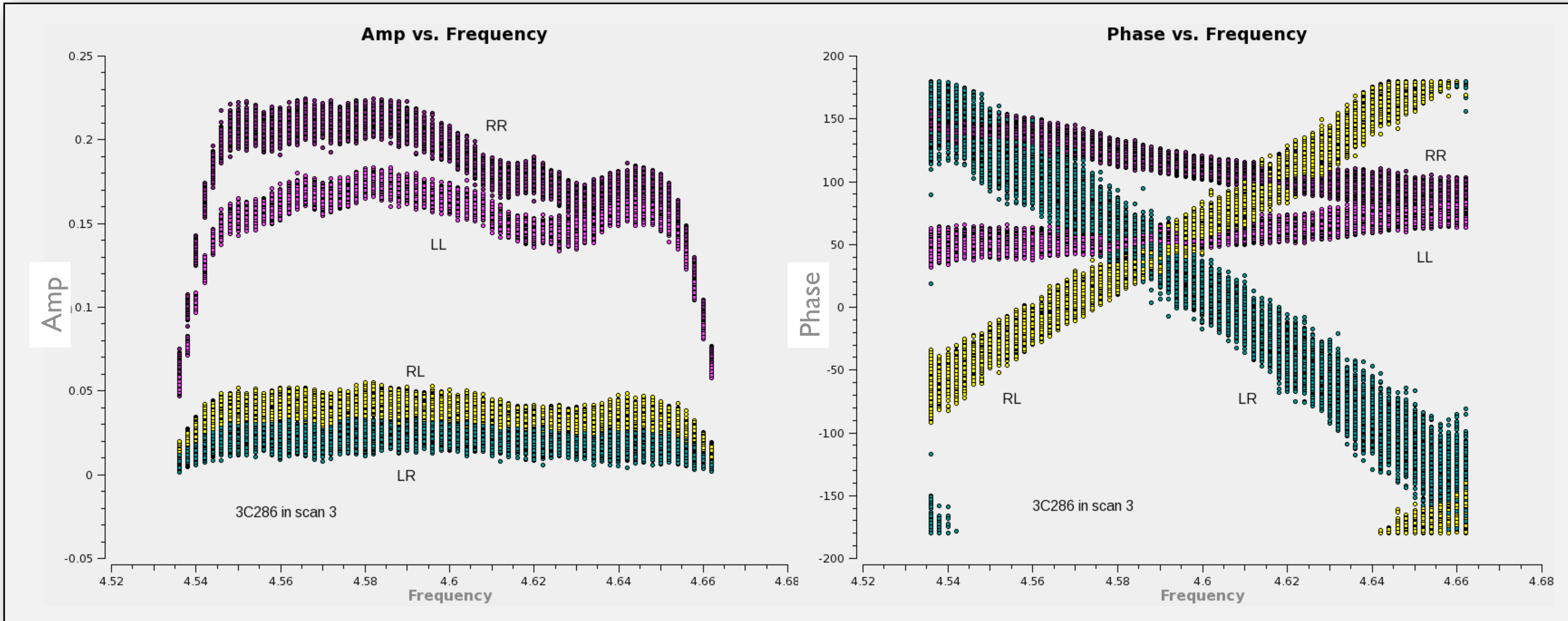
Single Baseline (Amplitude)



Single Baseline (Phase)

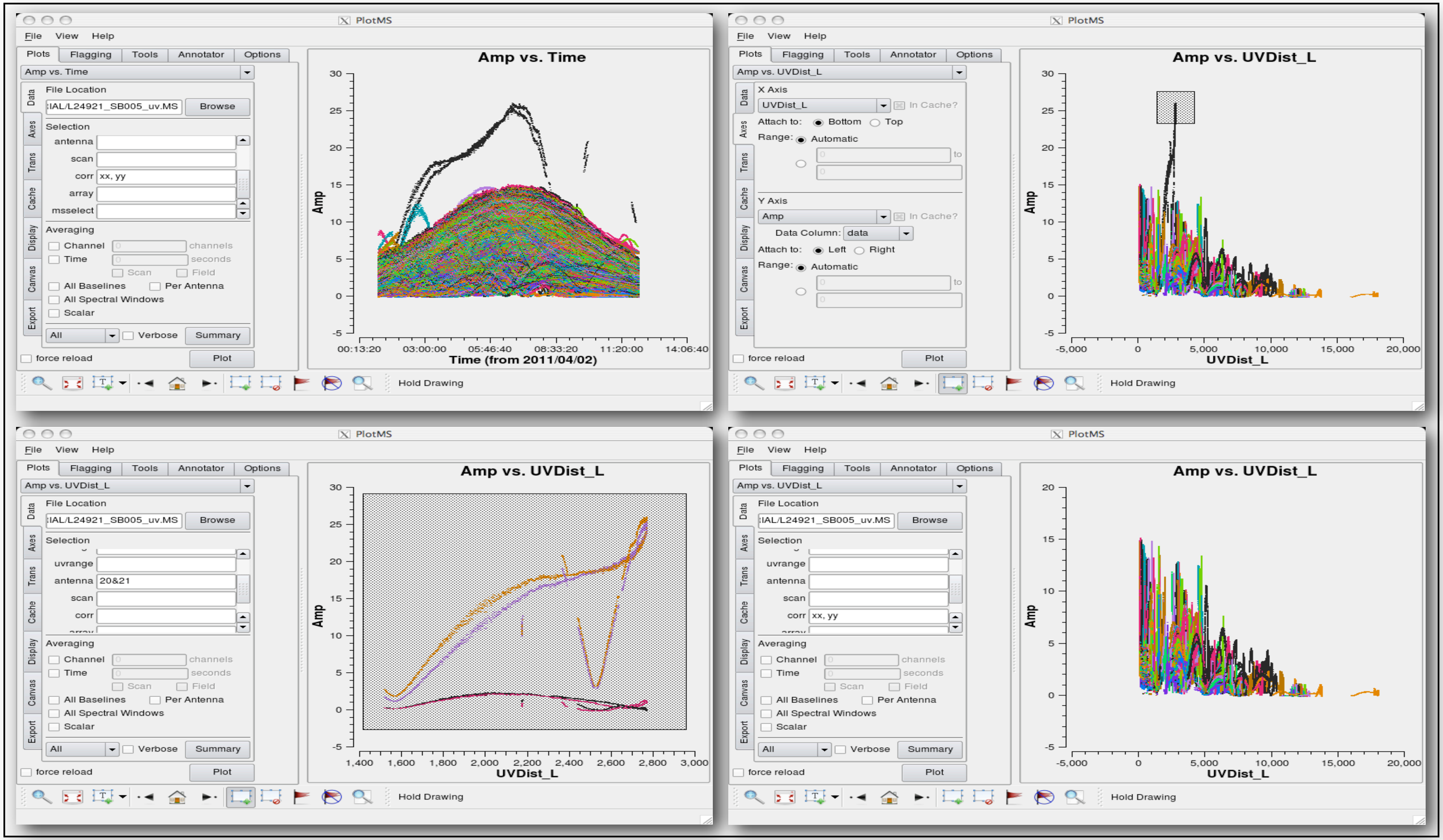


Single Baseline Spectra



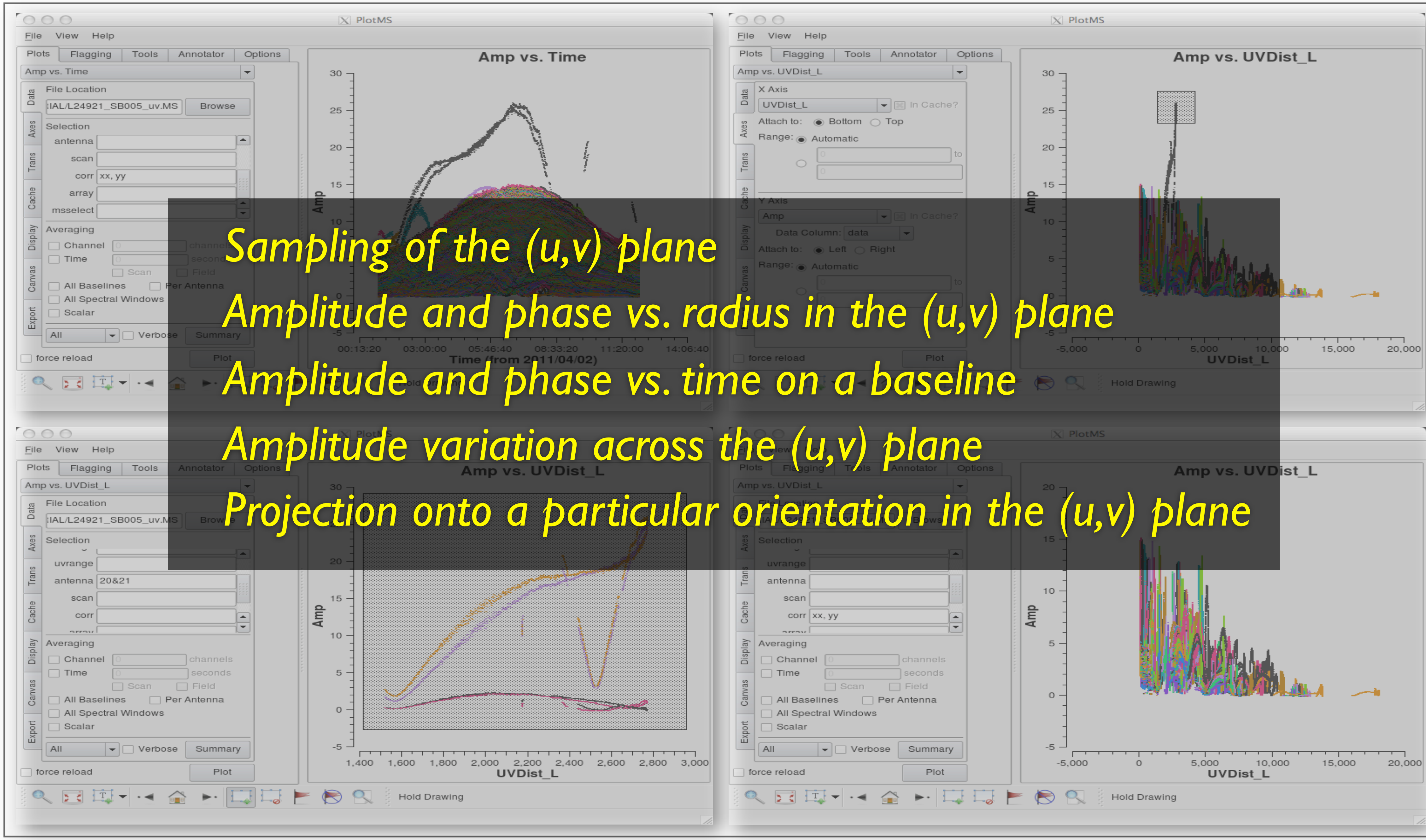
Baseline ea17-ea21   (all 4 polarizations)

Today's Practicum I



- Examine the visibility data from a LOFAR observation
- Use the interactive CASA tool "casaplotms"

Today's Practicum I



Sampling of the (u,v) plane

Amplitude and phase vs. radius in the (u,v) plane

Amplitude and phase vs. time on a baseline

Amplitude variation across the (u,v) plane

Projection onto a particular orientation in the (u,v) plane

- Examine the visibility data from a LOFAR observation
- Use the interactive CASA tool "casaplotms"

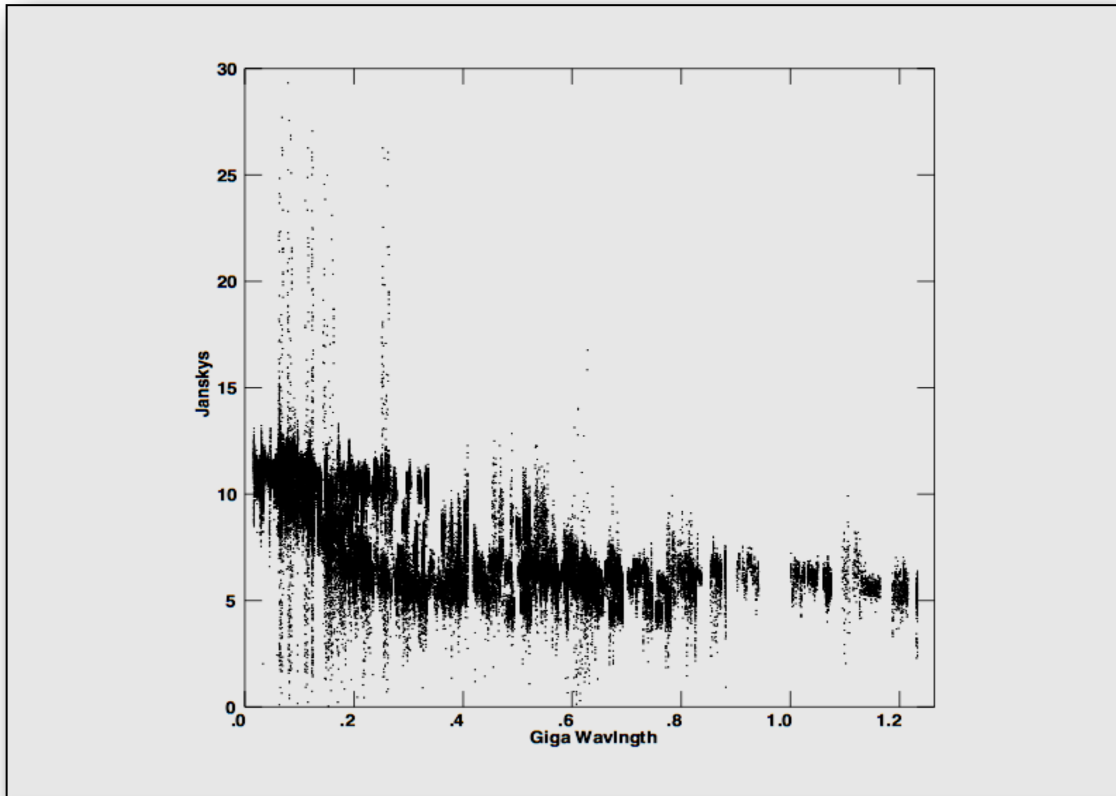
Intermission

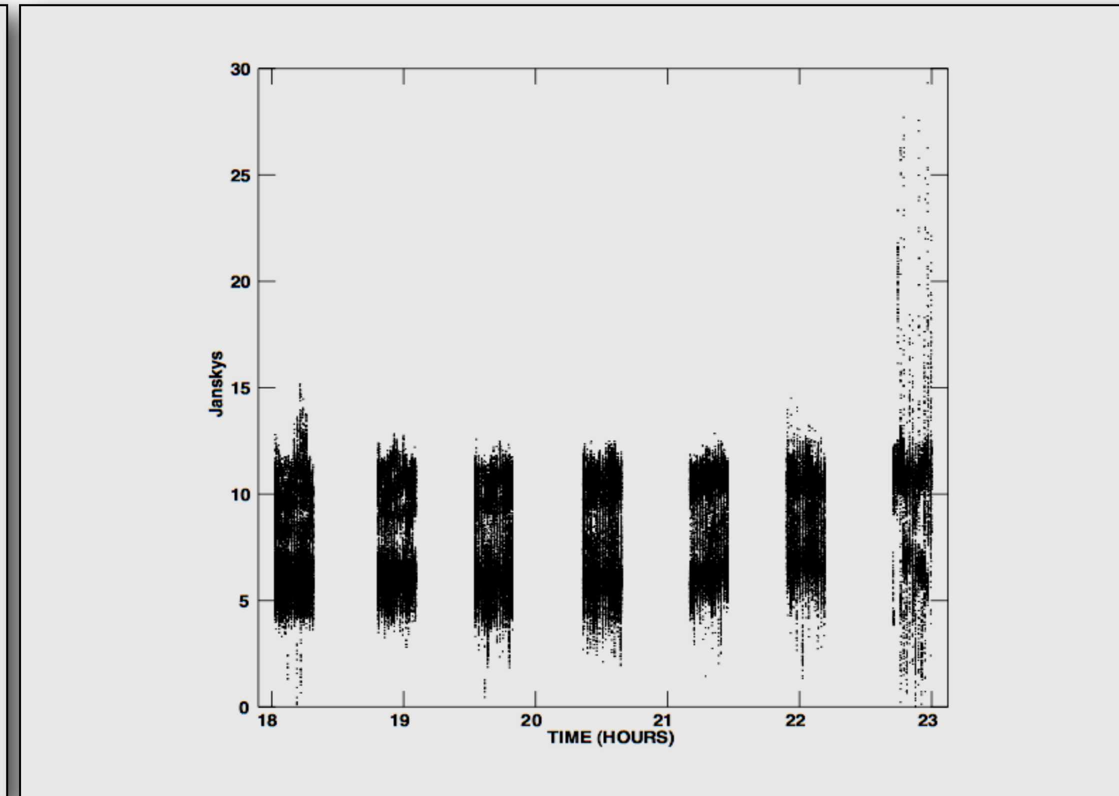
Data Editing

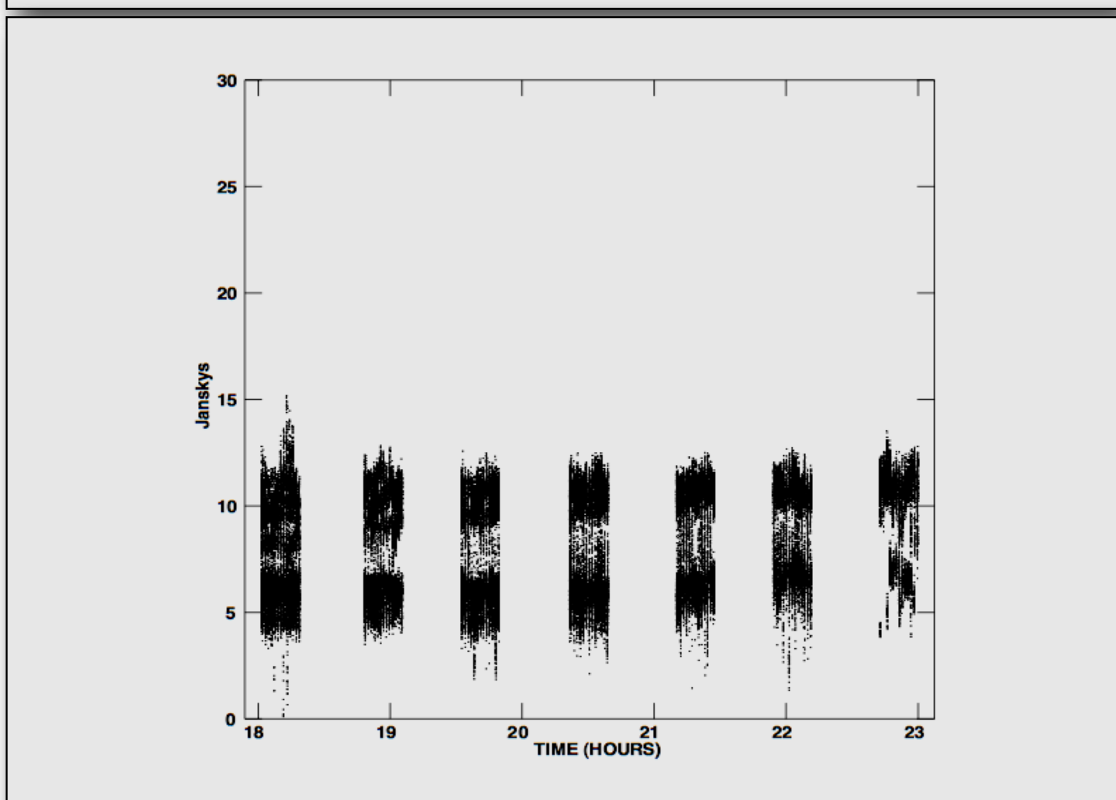
- Initial data examination and editing very important
- What to edit (much of this is automated):
  - Some real-time flagging occurred during observation
  - Any such bad data left over?
  - Any persistently "dead" antennas?
  - Periods of especially poor weather?
  - Amplitude and phase should be continuously varying $\Longrightarrow$ remove outliers
  - Any Radio Frequency Interference (RFI)?
- Caution:
  - Be careful editing noise-dominated data.
  - Be conservative $\Longrightarrow$ antennas or time-ranges which are bad on calibrators are probably bad on weak target sources $\Longrightarrow$ remove them
  - Distinguish between bad (hopeless) data and poorly-calibrated data. E.g., some antennas may have significantly different amplitude response which may not be fatal—it may only need to be calibrated
  - Choose (phase) reference antenna wisely (ever-present, stable response)
- Increasing data volumes increasingly demand automated editing algorithms…

$\Rightarrow$ Bad data is worse than no data…

Editing Example



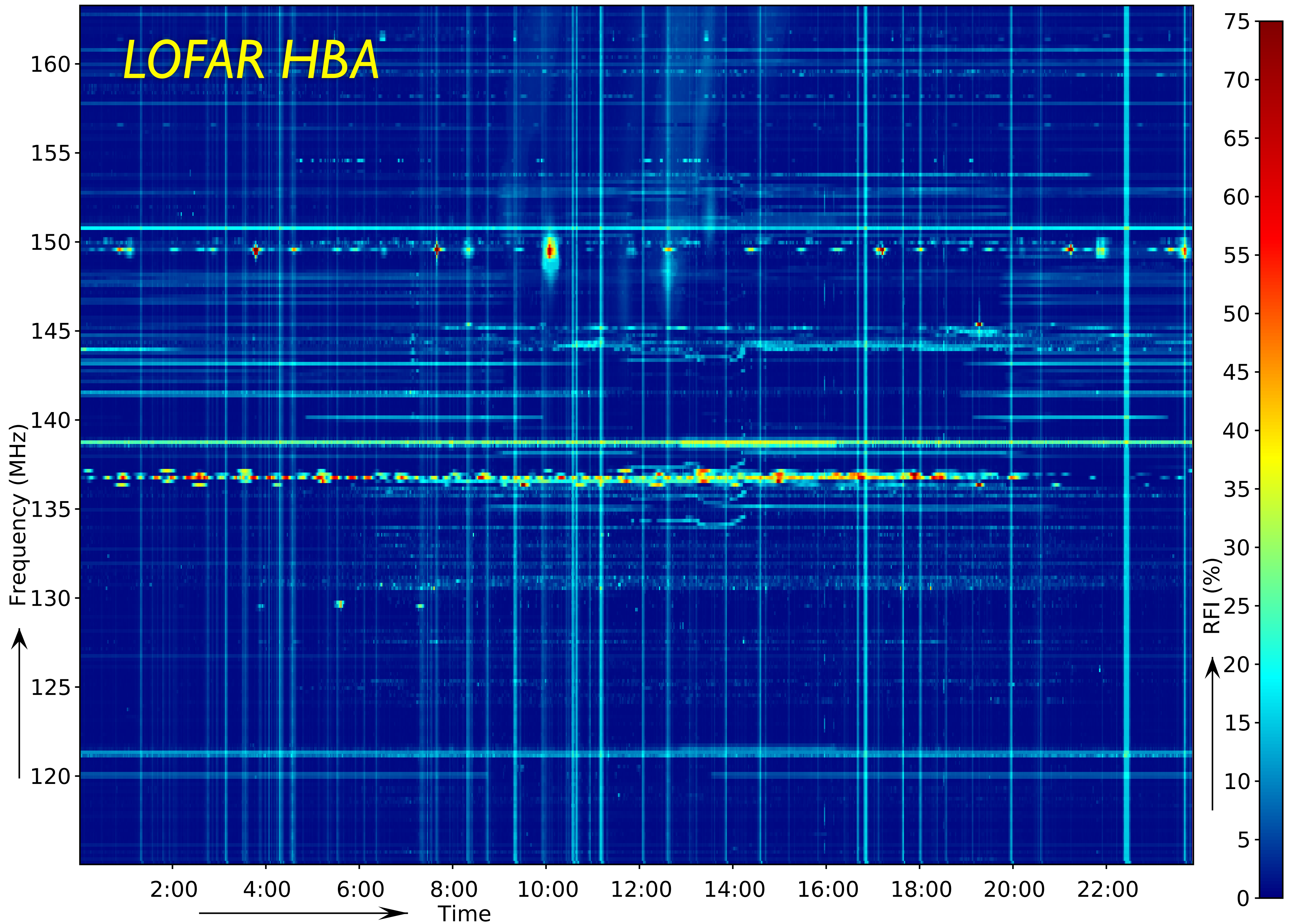




3C279 VLBA data at 43 GHz

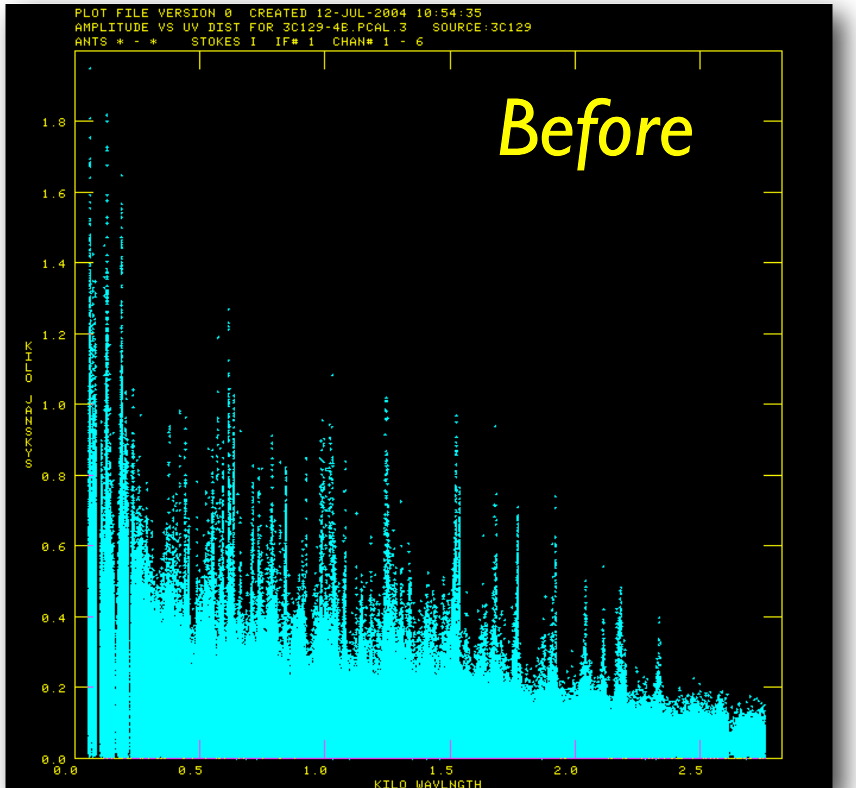
- Amplitude vs. uvdist shows outliers
- Amplitude vs. time shows outliers in last scan
- Amplitude vs. time without antenna 7 shows good data

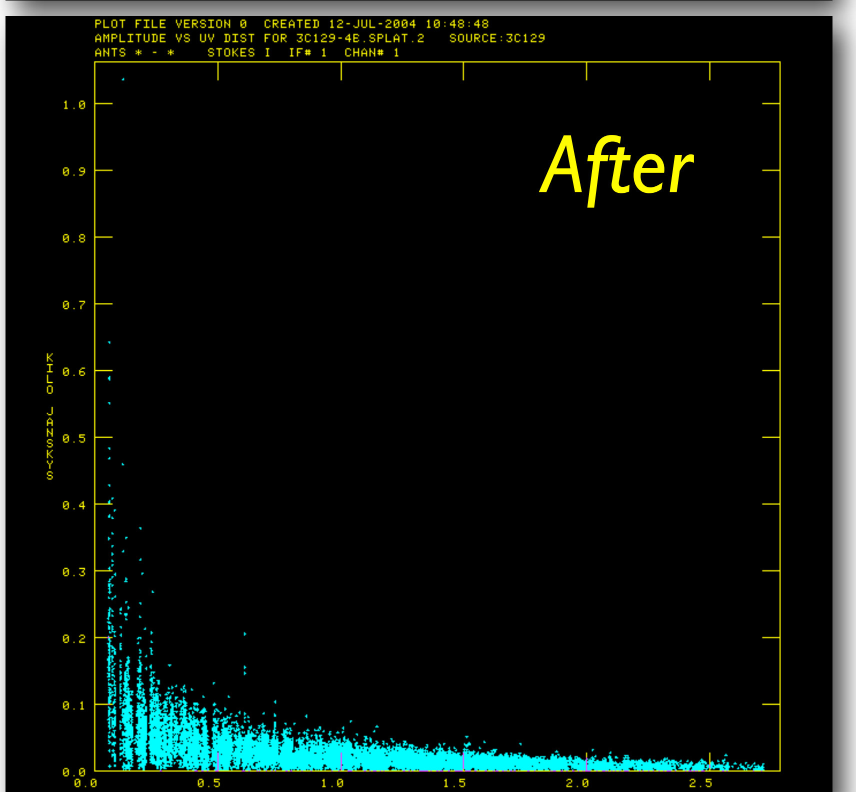
RFI Excision



Offringa et al. (2013)

RFI Excision in Practice



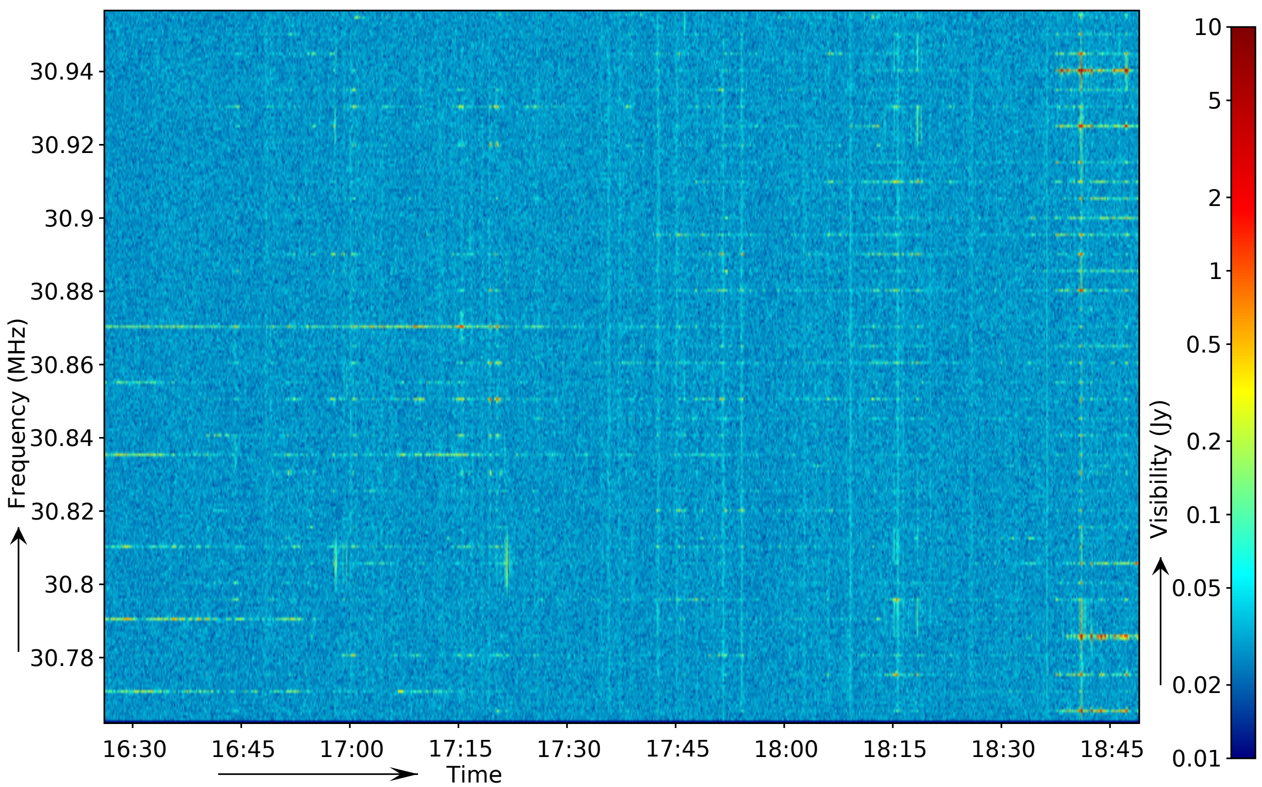


- RFI environment worse on short baselines
- Several 'types': narrow band, wandering, wideband, ...
- Wideband interference hard for automated routines
- Approach: averaging data in time and/or frequency makes it easier to isolate RFI, which averages coherently, from Gaussian noise, which does not
- Once identified, the affected times/baselines can be flagged in the un-averaged dataset
- Many tools available for manual editing: QUACK, SPFLG, TVFLG, UVPLT, UVFND, UVFLG, UVSUB, CLIP, FLGIT, FLAGR, ...

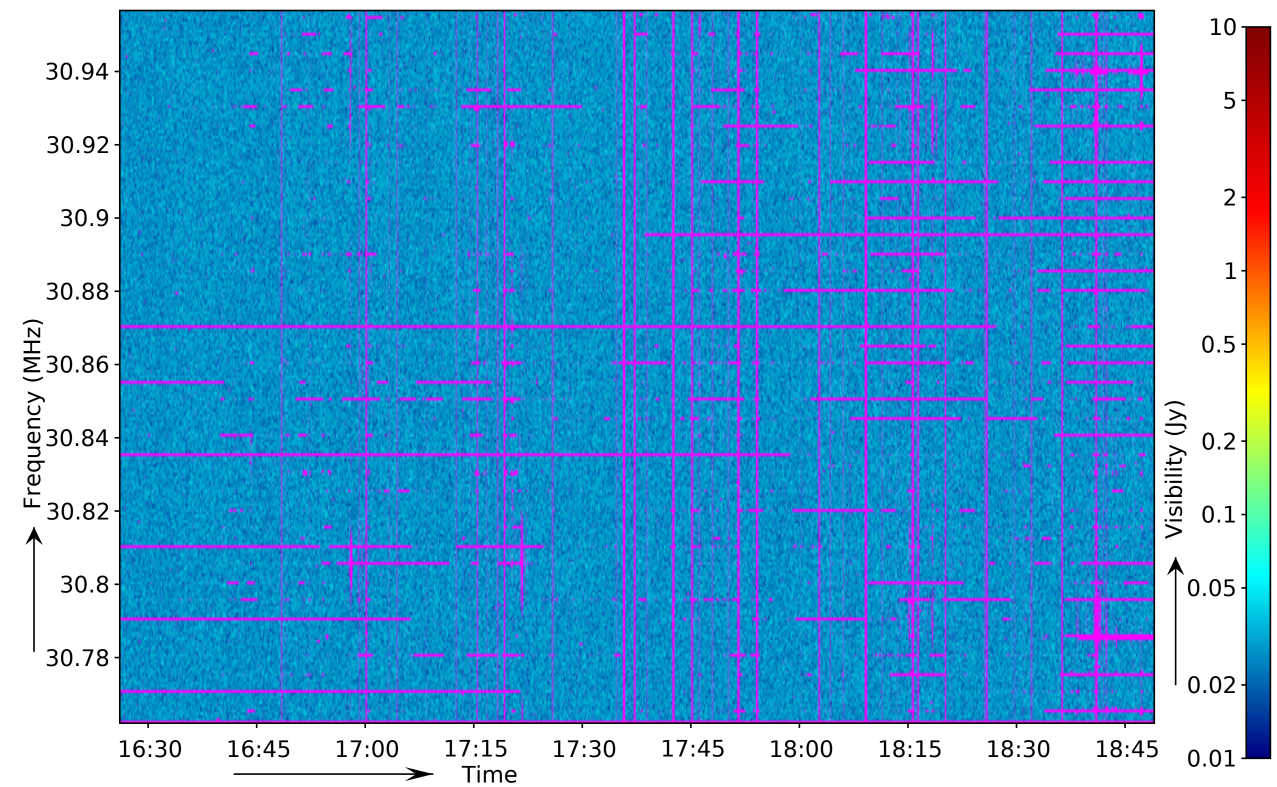
⇒ Data volumes increasingly require automated routines

Automated RFI Flagging

LOFAR LBA RFI



Automated Flagging



Offringa et al. (2013)

- LOFAR pipeline uses AOFLagger (*AO = André Offringa*)
- Runs in automated way, but can be run interactively from GU
- High spectral resolution (~1kHz by default) catches most RFI

Calibration Formalism

Calibration Equation

Fundamental calibration equation

$$V_{ij}(t) = g_i(t) g_j^*(t) V^{true}(t) + \varepsilon_{ij}(t)$$

$V_{ij}(t)$	Visibility measured between antennas i and j
$g_i(t)$	Complex gain of antenna i
$V^{true}(t)$	True visibility
$\varepsilon_{ij}(t)$	Additive noise

Calibration using a point source

- Calibration equation becomes

$$V_{ij}(t) = g_i(t)g_j^*(t)S + \varepsilon_{ij}(t)$$

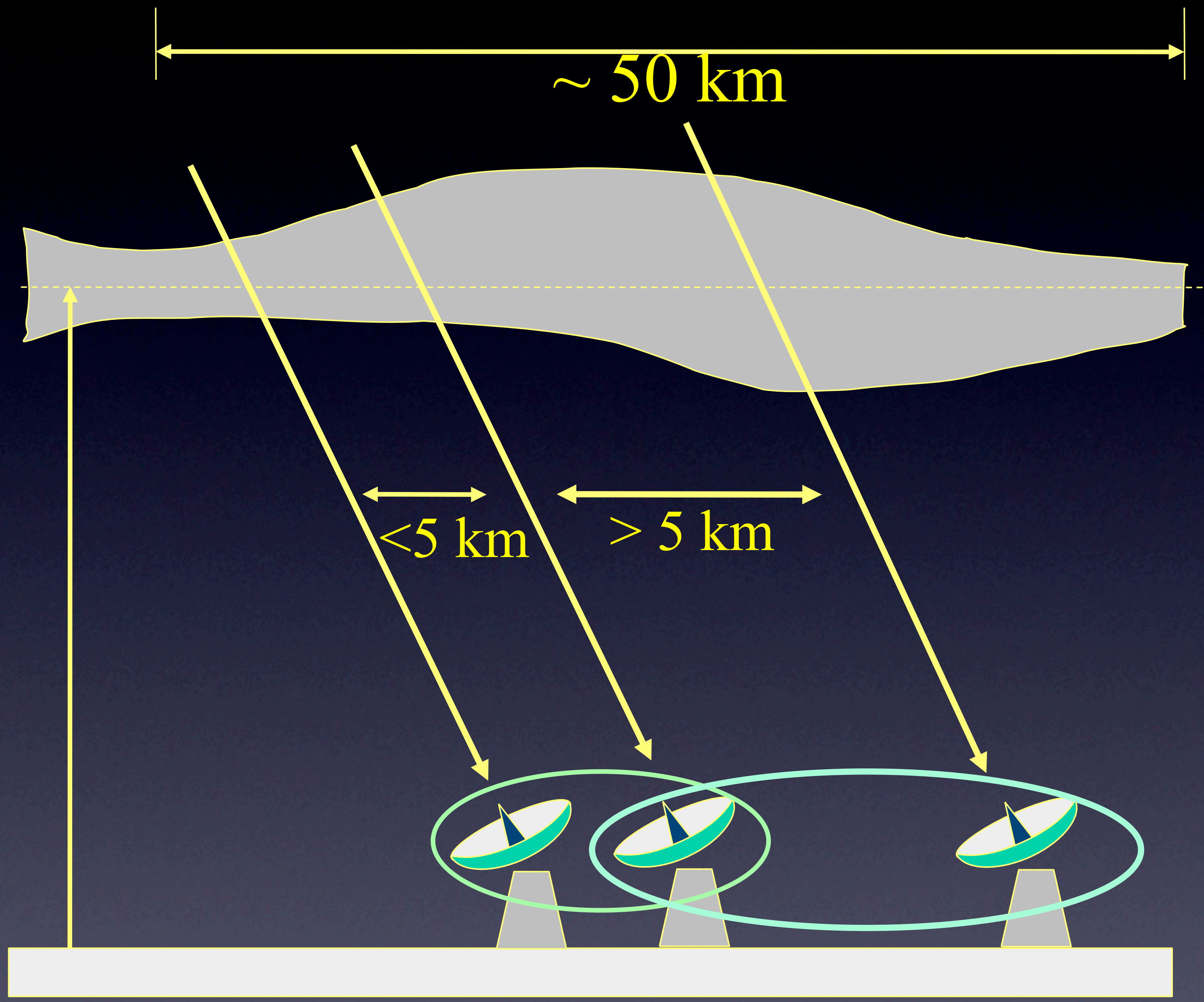
S Strength of point source

- Solve for antenna gains via least squares algorithm
- Works well - lots of redundancy
- (N-1) baselines contribute to gain estimate for any antenna
- Gains are antenna based and direction independent
- Also know as "cross calibration"

Why is *a priori* calibration insufficient?

- Initial calibration based on calibrator observed before and/or after target
- Gains were derived at a different time
  - Troposphere and ionosphere are variable
  - Electronics may be variable
- Gains were derived for a different direction
  - Troposphere and ionosphere are not uniform
  - $>$ 1 GHz $\Longrightarrow$ troposphere, $<$ 1 GHz $\Longrightarrow$ ionosphere
- Observation might have been scheduled poorly for the existing conditions

Ionospheric Structure



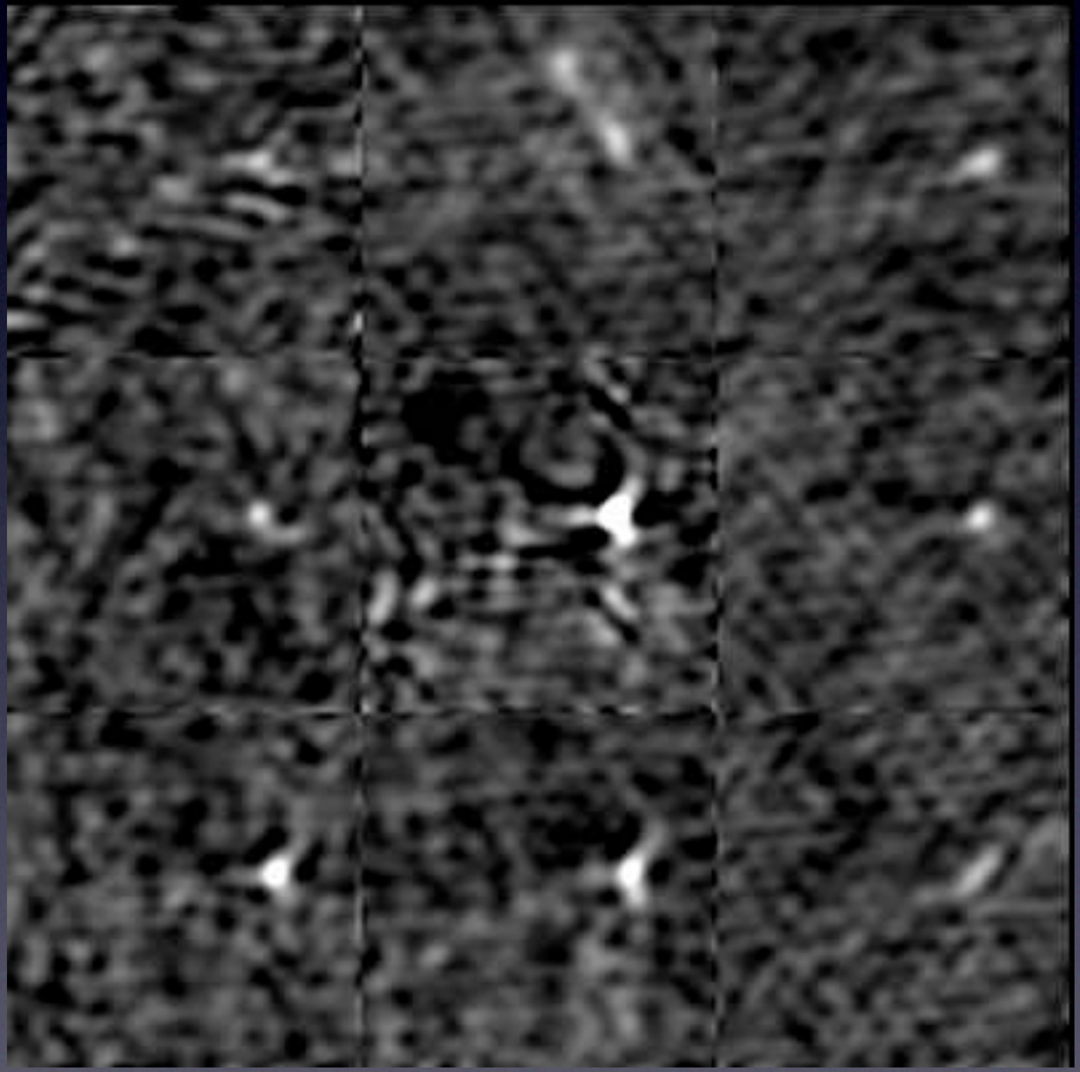
- Waves in the ionosphere introduce rapid phase variations (~1°/s on 35 km baselines)
- Phase coherence is preserved on BL < 5km
- BL > 5 km have limited coherence times
- Historically limited capabilities of low frequency instruments

⟹ Introduces timescale into calibration solutions

Effects of Ionosphere

Virgo A over the VLA (1.4 GHz)

Multiple sources in same field



(Courtesy B. Cotton)

- Ionospheric refraction and distortion
- Both global and differential refraction seen
- Time scales of 1 minute or less
- Equivalent length scales of 10 km or less

Calibration using a complex model

- Don’t need point source - can use model

$$V_{ij}(t) = g_i(t) g_j^*(t) V_{ij}^{\text{model}} + \varepsilon_{ij}(t)$$

V_{ij}^{model} Model visibility

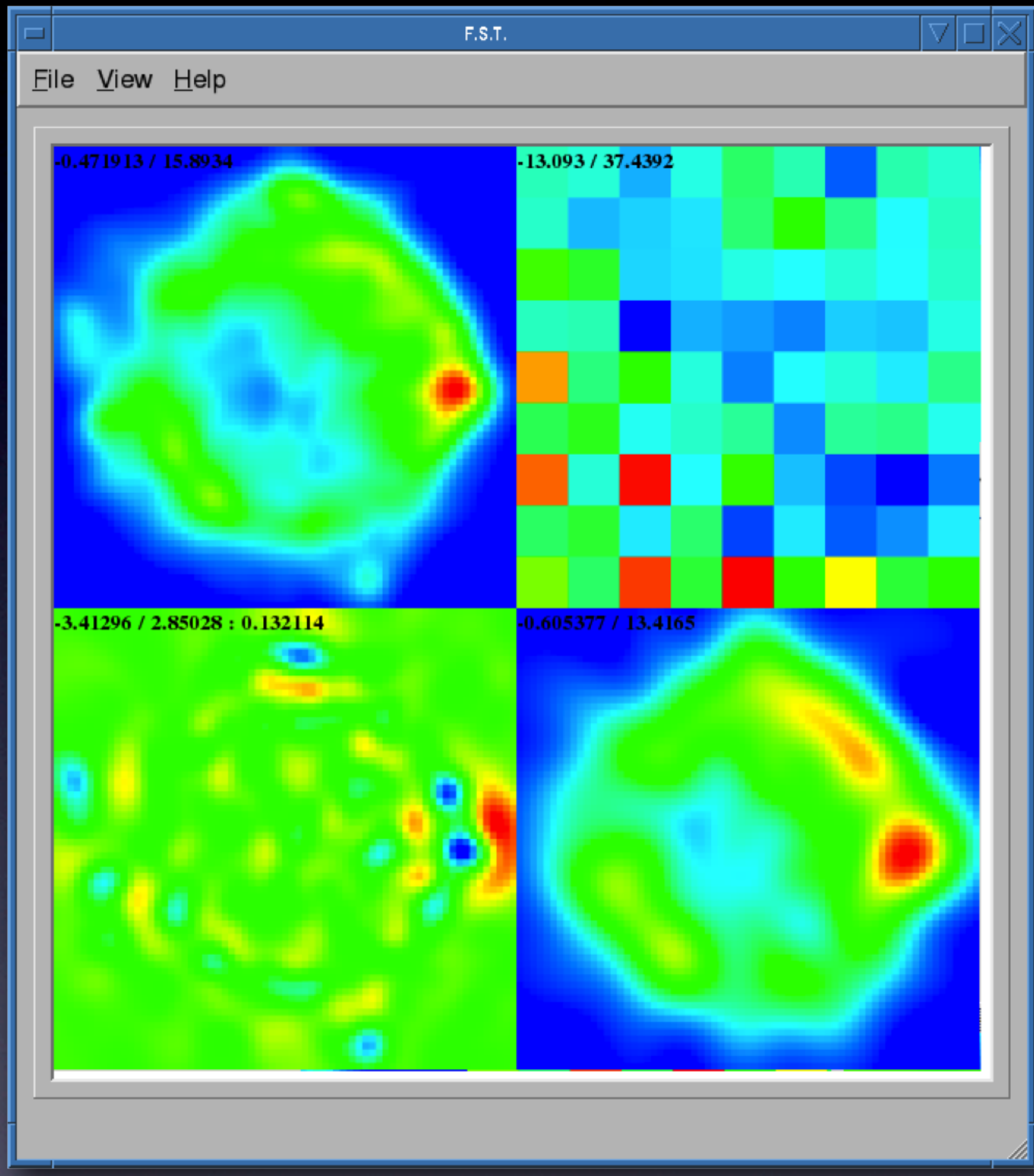
- Redundancy means that errors in the model average down
- Have N(N-1) equations with N unknowns
- Correct for estimated gains:

$$V_{ij}^{\text{cal}}(t) = \left(g_i(t) g_j^*(t)\right)^{-1} V_{ij}$$

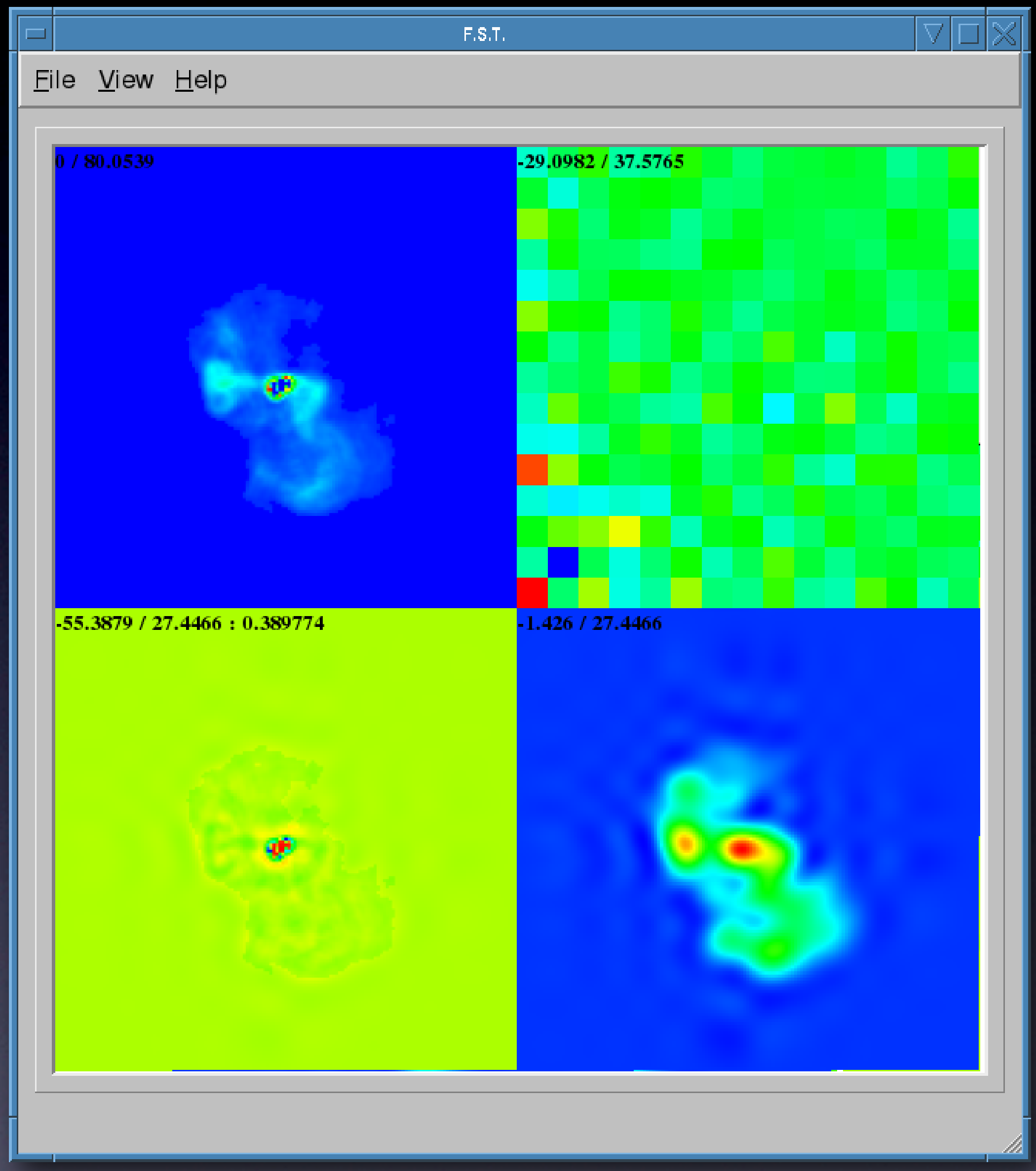
- Can smooth or interpolate gains if desired (be careful!)

Calibration using a complex model

Cas A (SNR)



Virgo A (M87)



⇒ *Can use shapelets for very complicated sources*

Self-Calibration

- Advantages
  - Gains derived for correct time --- no interpolation
  - Gains derived for correct position --- no atmospheric assumptions
  - Solution is fairly robust if there are many baselines
  - More time on-source
- Disadvantages
  - Requires a sufficiently bright source
  - Introduces more degrees of freedom into the imaging
  - Results might not be robust and stable
  - Absorbs position shifts (phase) and amplitude variations

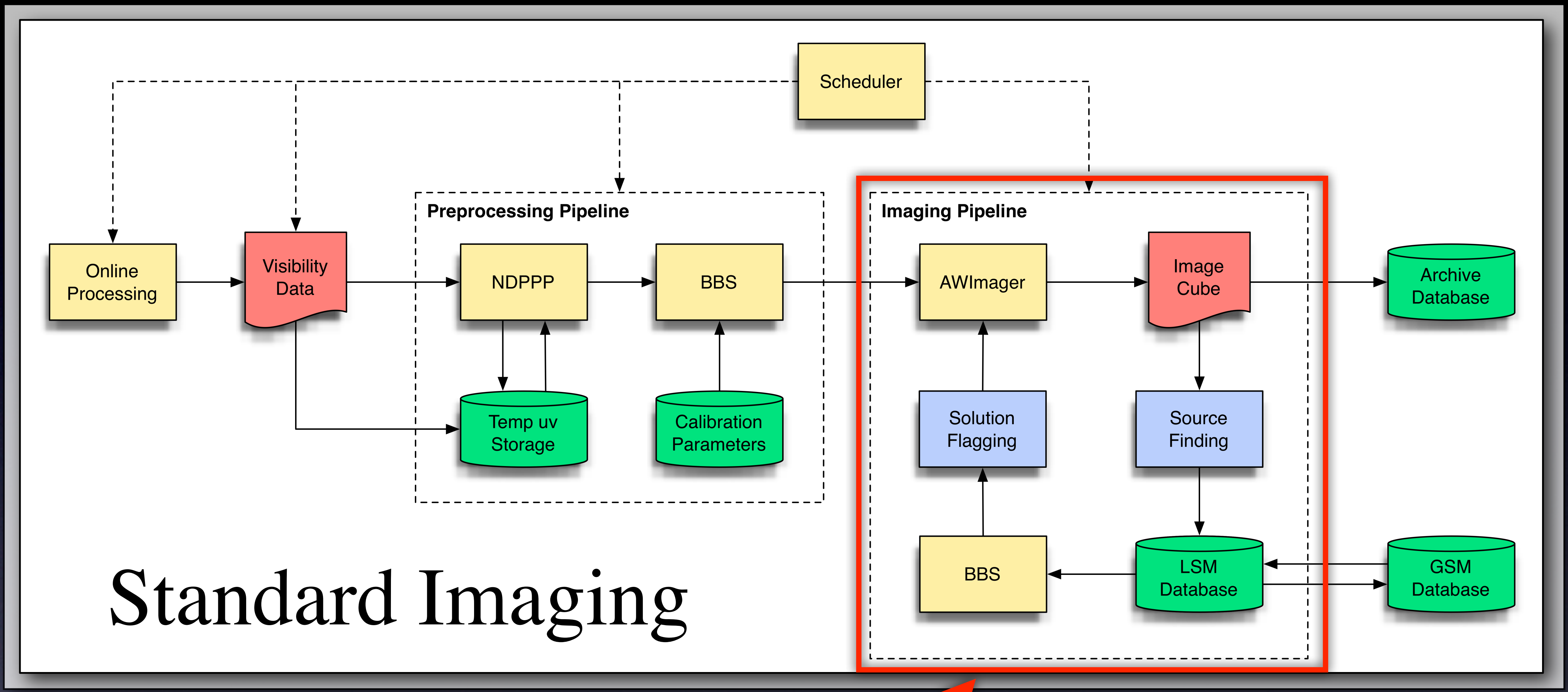
$\Rightarrow$ *Computationally expensive for crowded fields*

How to Self-Calibrate

1. Create an initial source model, typically from an initial image (or else a point source)
   - Use full resolution information from the clean components or MEM image NOT the restored image
2. Find antenna gains
   - Using "least squares" (L1 or L2) fit to visibility data
3. Apply gains to correct the observed data
4. Create a new model from the corrected data
   - Using for example Clean or Maximum Entropy
5. Go to (2),unless current model is satisfactory
   - shorter solution interval, different uv limits/weighting
   - phase ⇒ amplitude & phase

⇒ *You are solving for both gains* ***and*** *model*

How to Self-Calibrate

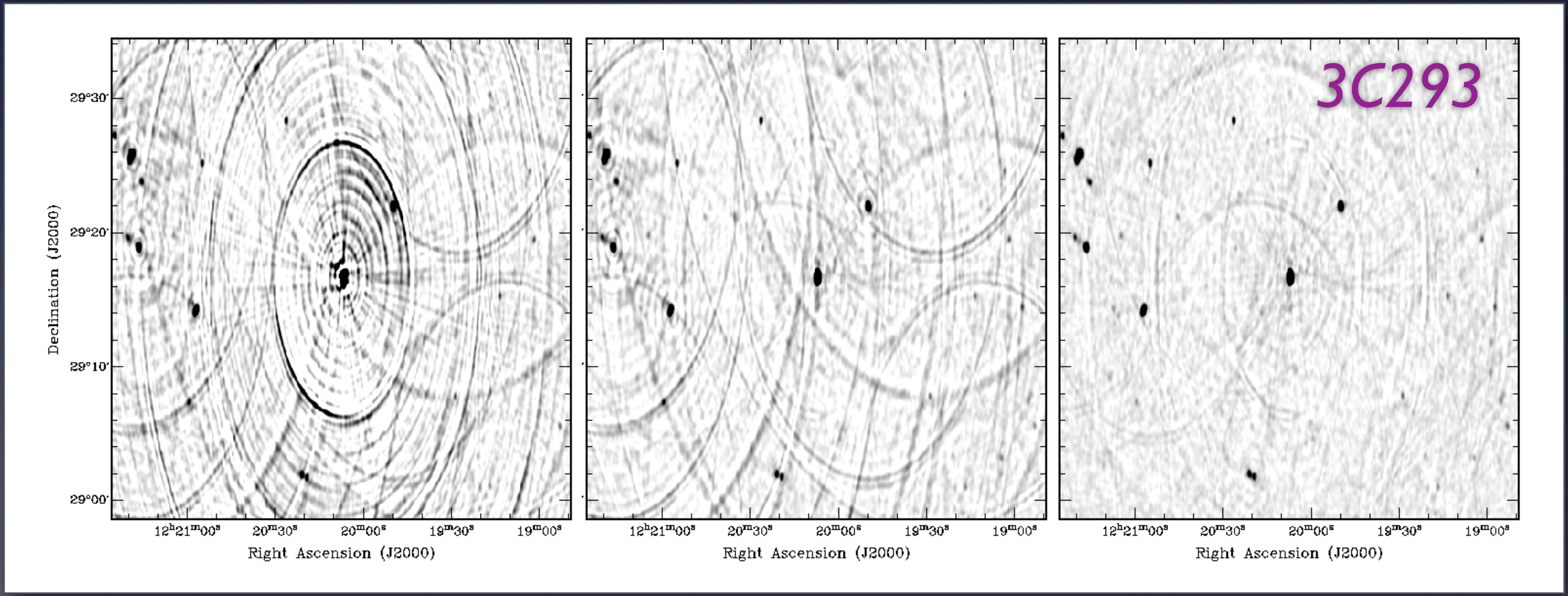


LOFAR Standard Imaging Pipeline

Self-Cal loop implemented as "Major Cycle"

To self-calibrate or not?

- Calibration errors may be present if one or both of the following are true:
  - The background noise is considerably higher than expected
  - There are convolutional artifacts around objects, especially point sources
- Don't bother self-calibrating if these signatures are not present
- Don't confuse calibration errors with poor Fourier plane sampling such as:
  - Low spatial frequency errors (fuzzy blobs) due to lack of short spacings
  - Multiplicative fringes (due to deconvolution errors)
  - Deconvolution errors around moderately resolved sources



Some Self-cal Guidelines

- Initial model
  - Point source often works well
  - Simple fit (e.g., Gaussian) for barely-resolved sources
  - Clean components from initial image (Don't go too deep!)
  - Simple model-fitting in (u,v) plane
- Self-calibrate phases or amplitudes?
  - Usually phases first (phase errors cause anti-symmetric image features)
  - For VLA and VLBA, amplitude errors tend to be relatively unimportant at dynamic ranges < 1000 or so
- Which baselines?
  - For a simple source, all baselines can be used
  - For a complex source, start with a compact components, and use longer baselines
- What solution interval should be used?
  - Use the shortest solution interval that gives "sufficient" signal/noise ratio (SNR)
  - Solutions will not track the atmosphere optimally

Sensitivity limit

- Can self-calibrate if SNR >1 on most baselines
- For a point source, the error in the gain solution is

Phase only $$\sigma_g = \frac{1}{\sqrt{N-2}} \frac{\sigma_V}{S}$$

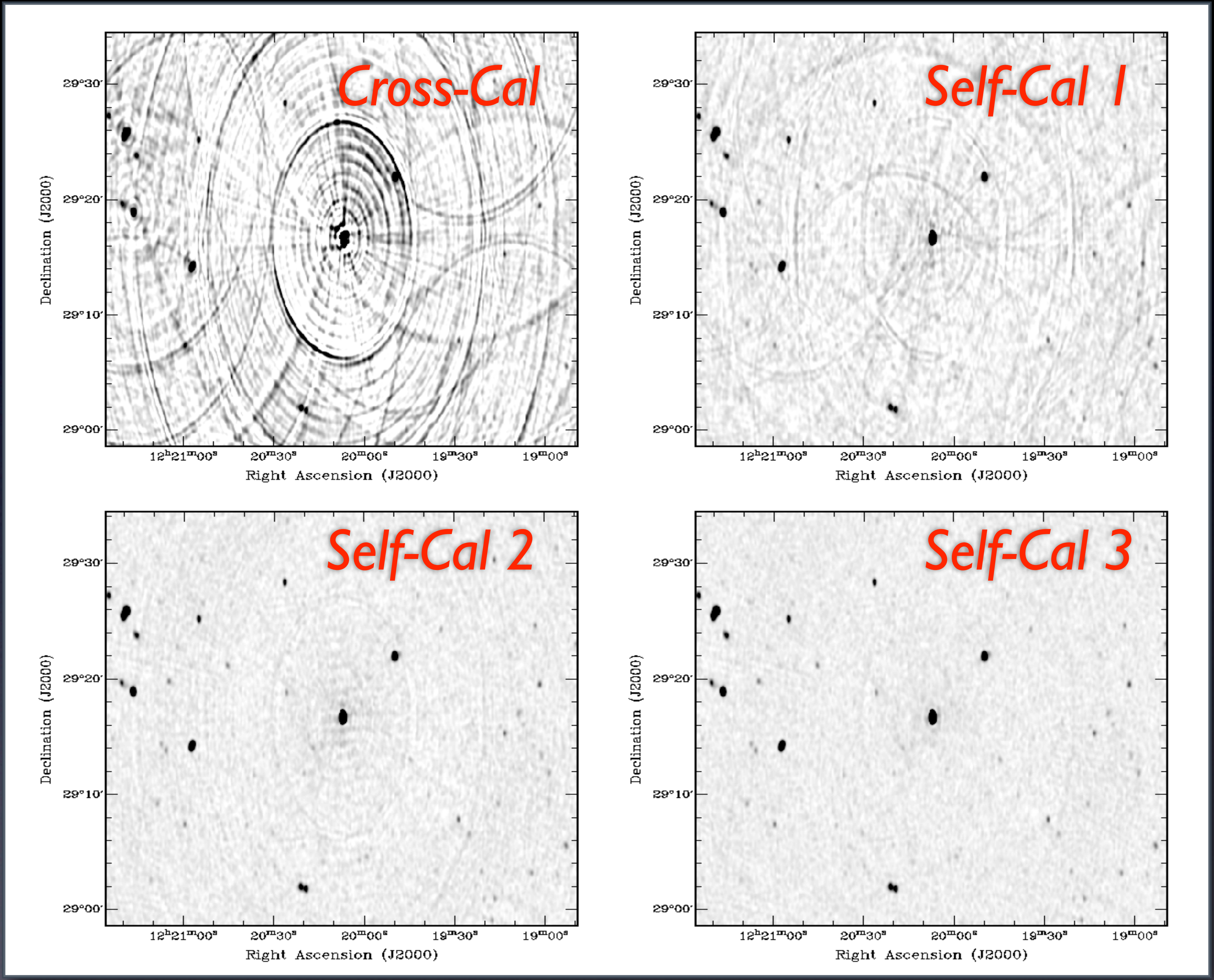
Amplitude and phase $$\sigma_g = \frac{1}{\sqrt{N-3}} \frac{\sigma_V}{S}$$

σ_V Noise per visibility sample

N Number of antennas

- If error in gain is much less than 1, then the noise in the final image will be close to theoretical

Self-cal Example: 3C293



Self-cal Example: Cygnus A

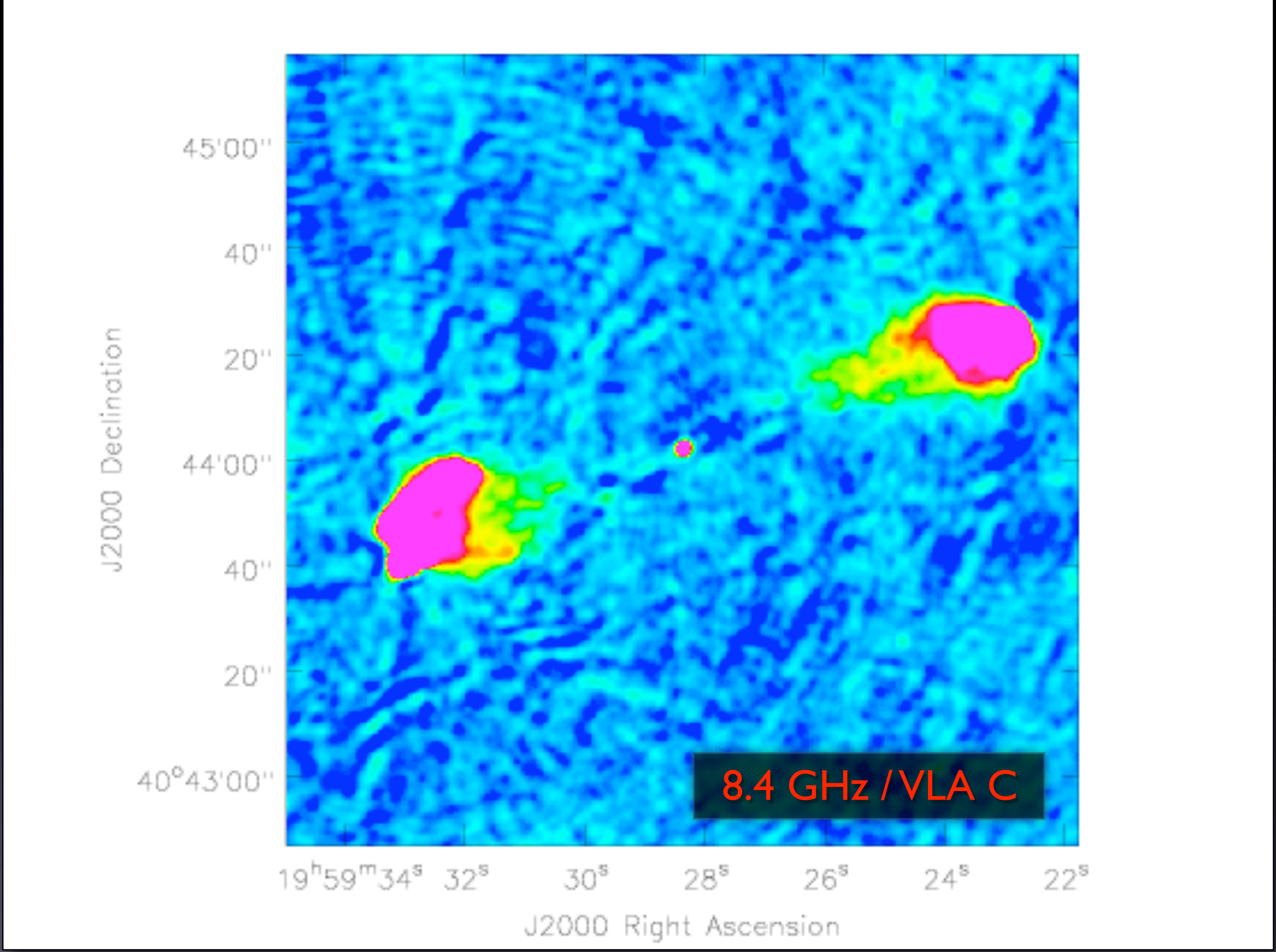
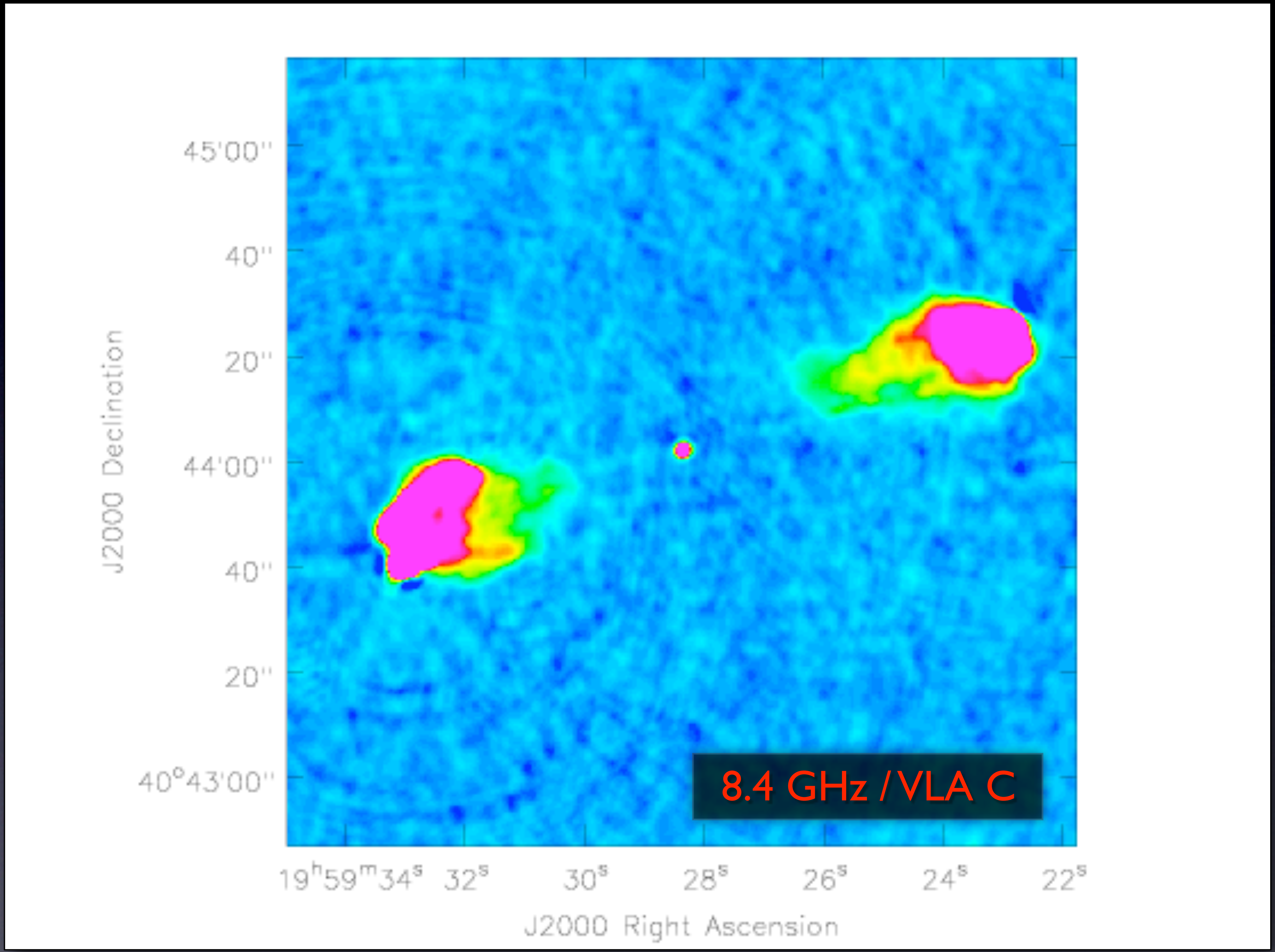


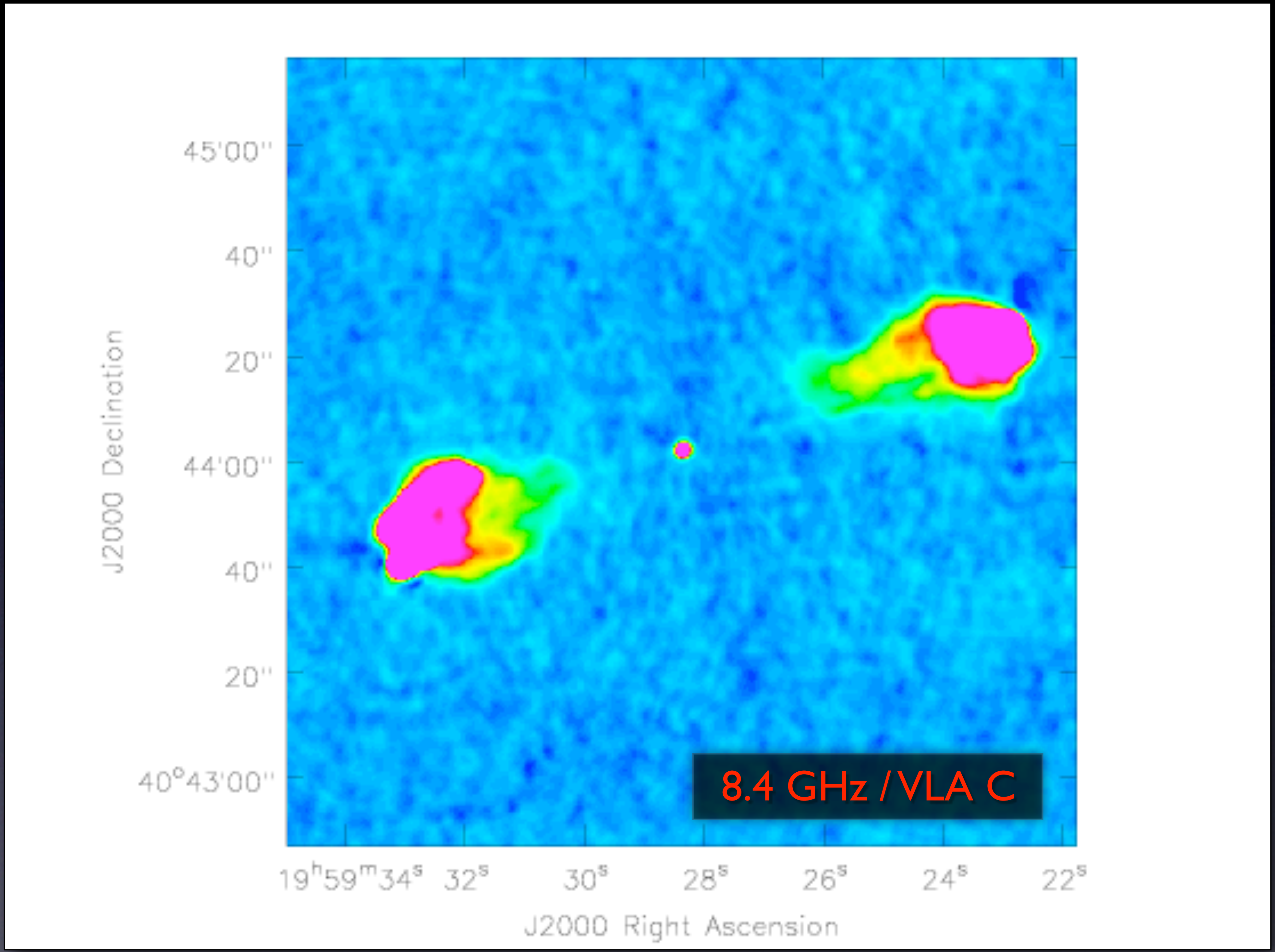
Image without self-calibration

Self-cal Example: Cygnus A



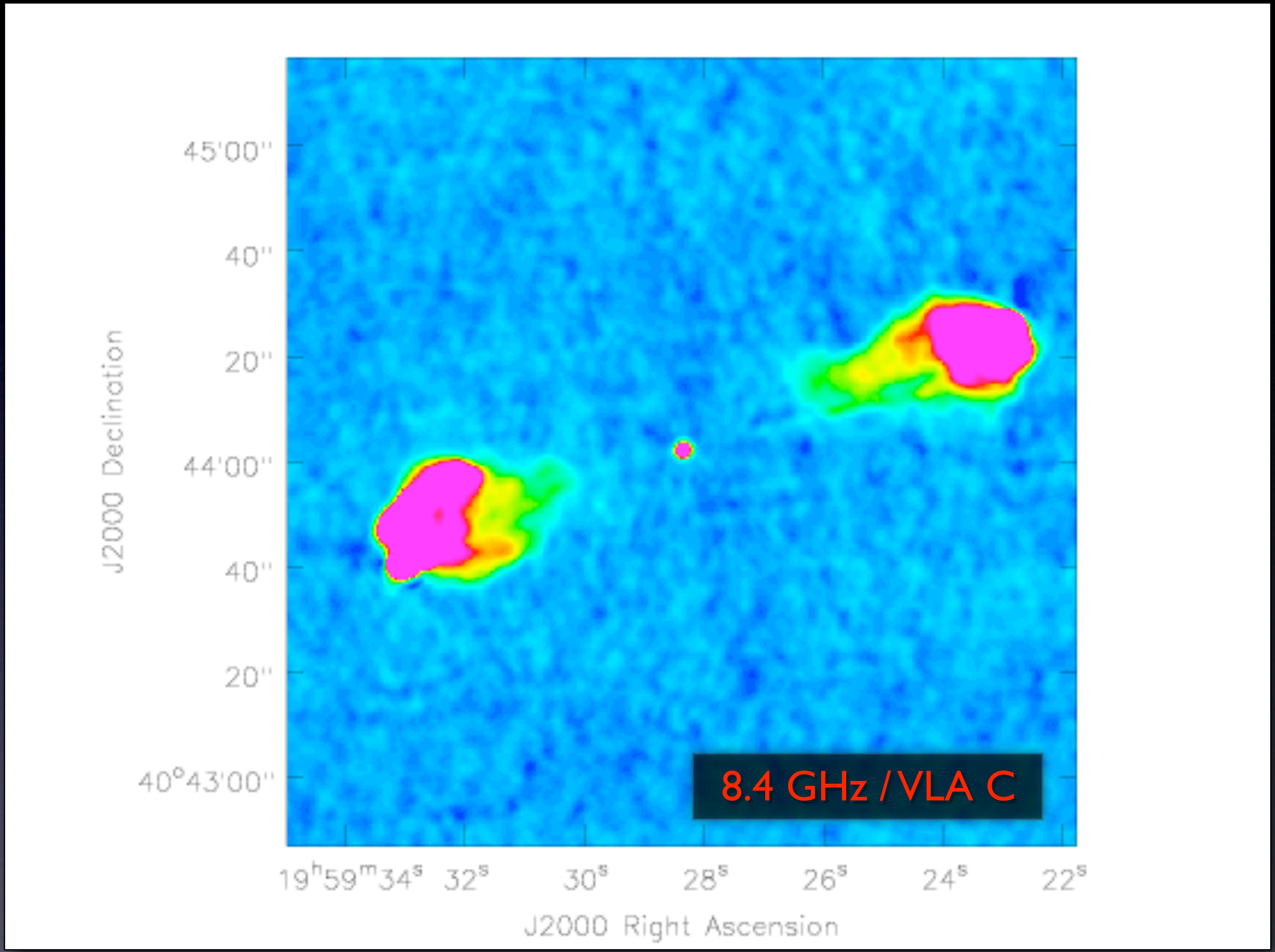
After 1 phase-only self-calibration

Self-cal Example: Cygnus A



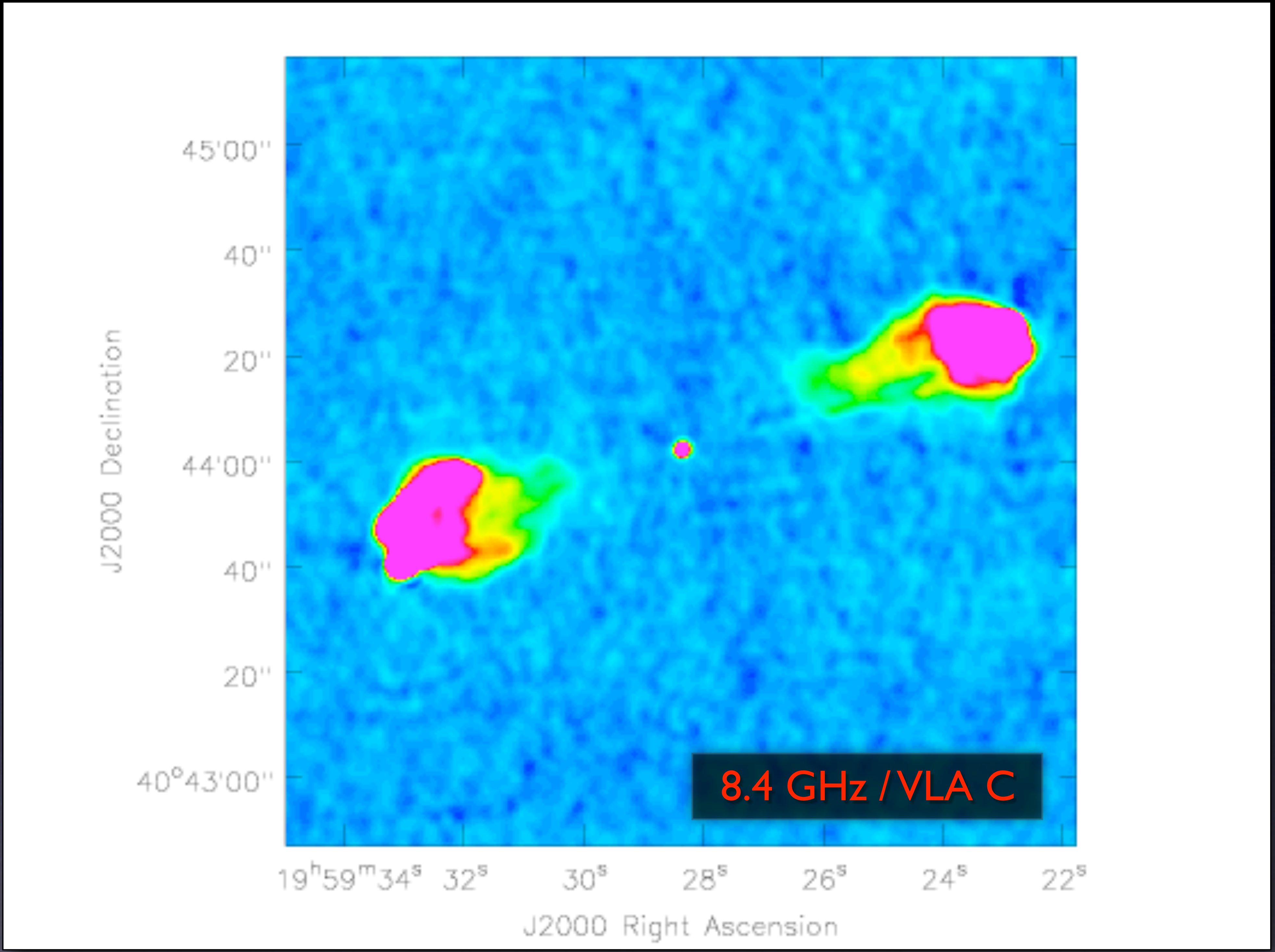
After 1 amplitude and phase calibrations

Self-cal Example: Cygnus A



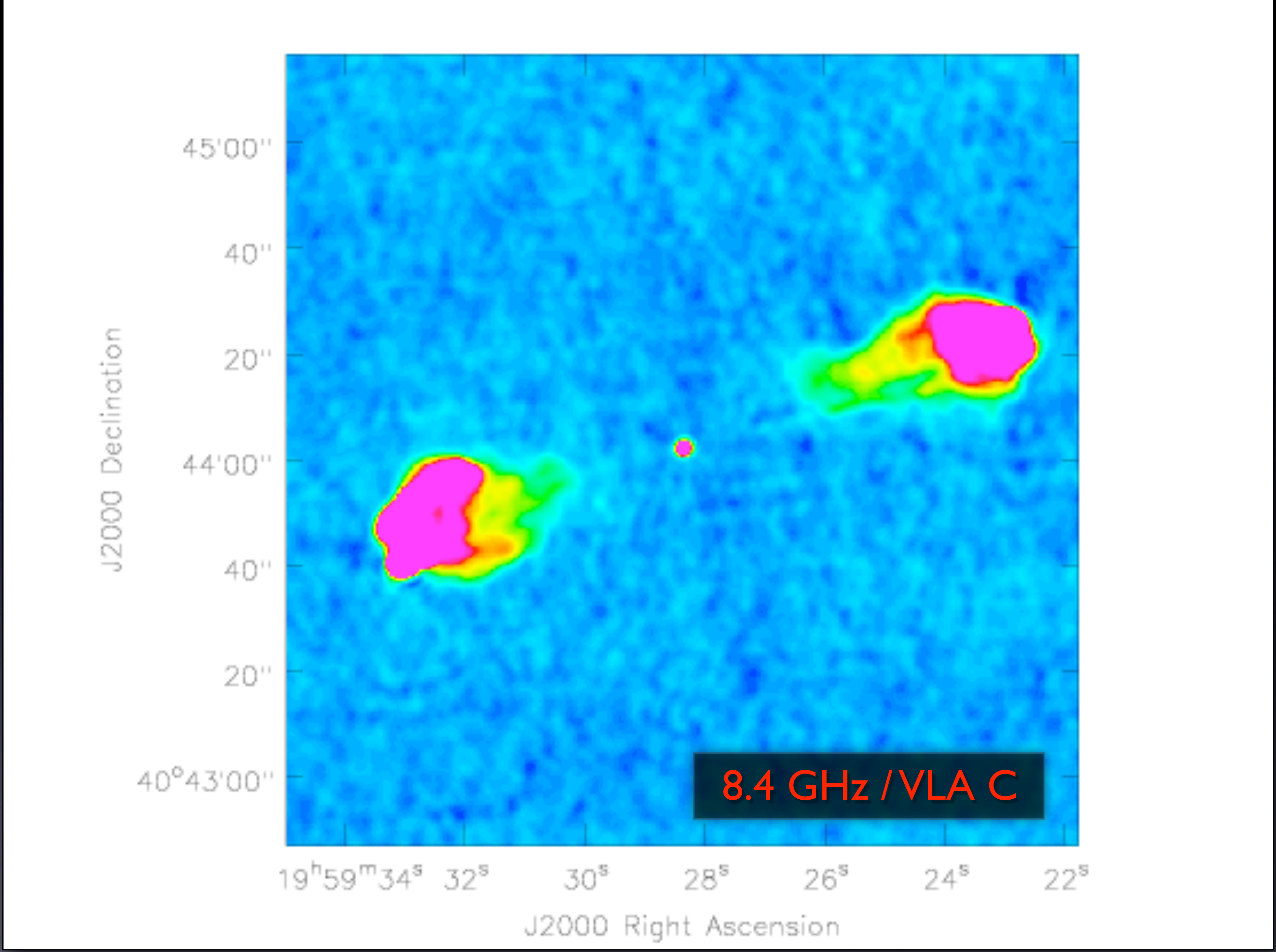
After 2 amplitude and phase calibrations

Self-cal Example: Cygnus A



After 3 amplitude and phase calibrations

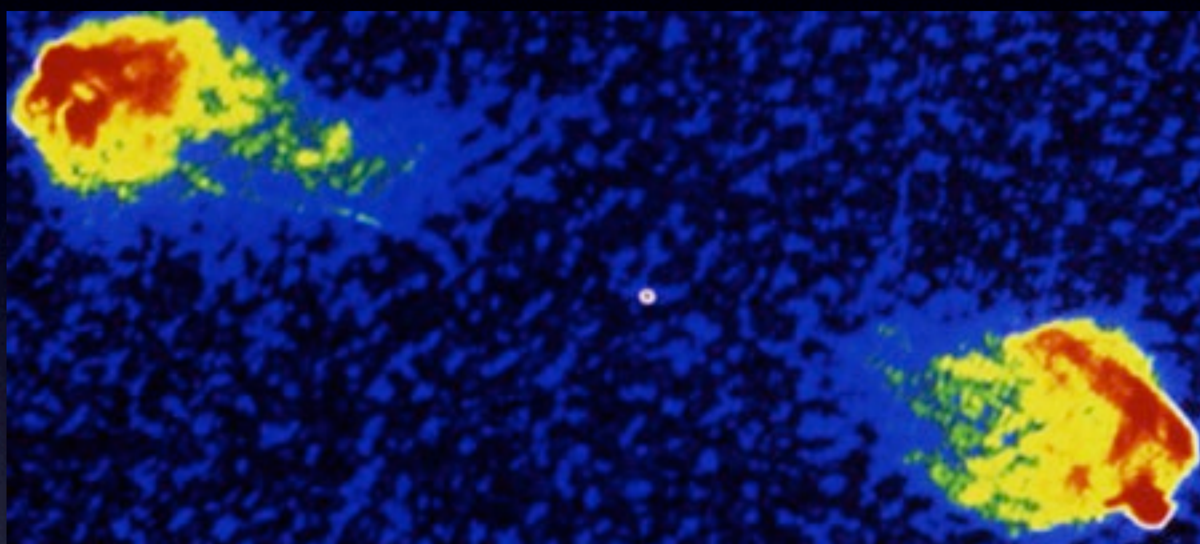
Self-cal Example: Cygnus A



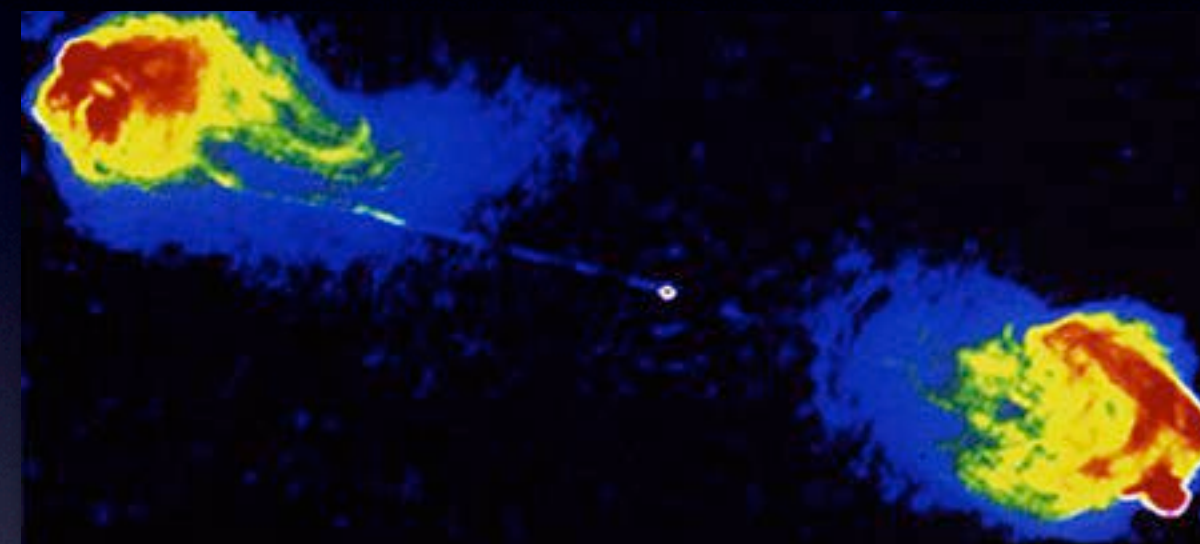
After 4 amplitude and phase calibrations

Self-cal Example: Cygnus A

Before Self-Cal



After Self-Cal



	Entire image			Off source		
	Max	**Minimum**	**RMS**	**Max**	**Minimum**	**RMS**
No selfcalibration	22.564	-0.179	0.409	0.072	-0.116	0.036
Phase only	22.586	-0.133	0.410	0.035	-0.035	0.013
1 Amp, Phase	22.976	-0.073	0.416	0.026	-0.033	0.012
2 Amp, Phase	22.912	-0.064	0.416	0.023	-0.033	0.012
3 Amp, Phase	22.887	-0.059	0.415	0.023	-0.033	0.012
4 Amp, Phase	22.870	-0.058	0.415	0.023	-0.032	0.012

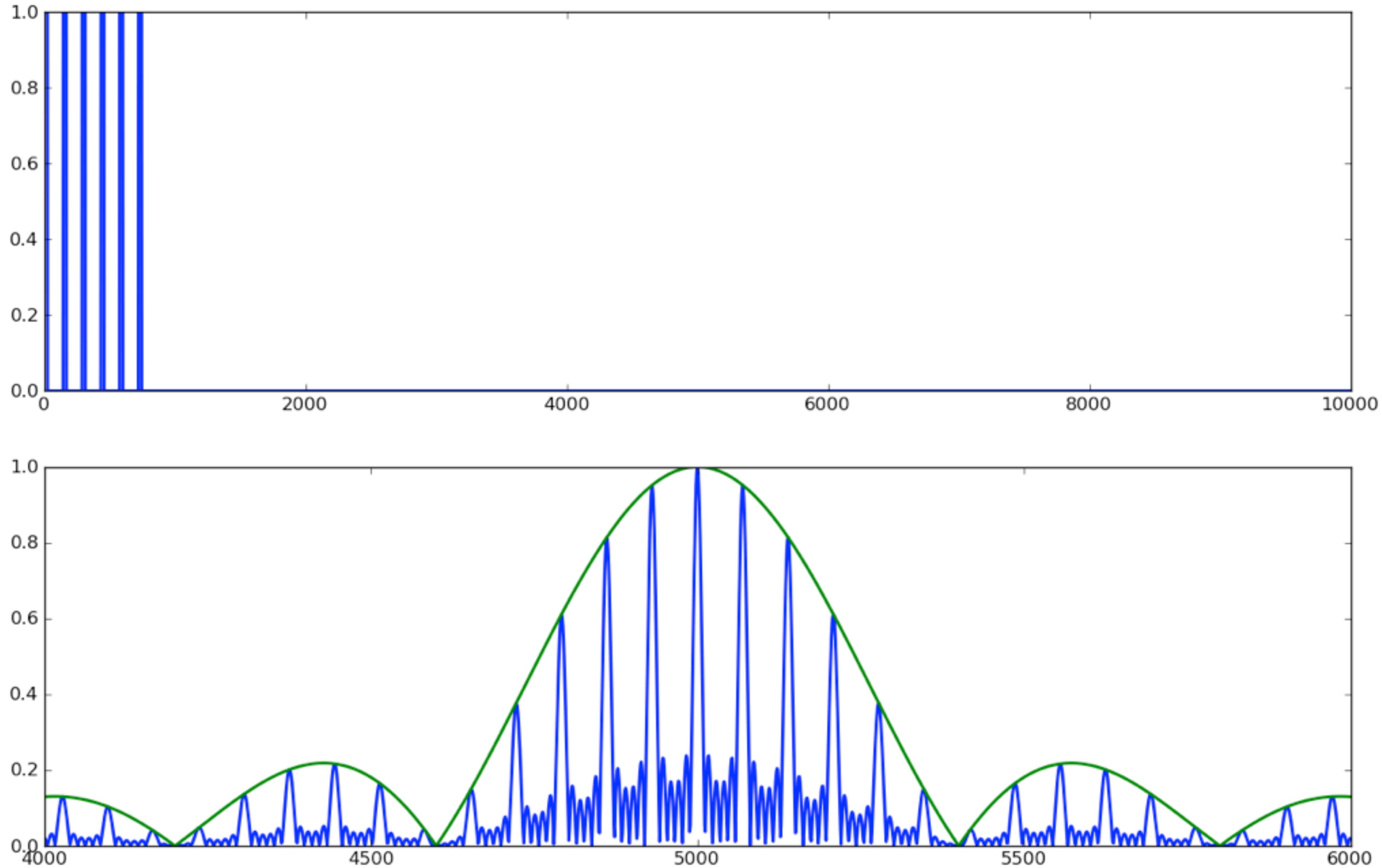
When Self-cal Fails

- Astrometry
- Variable sources
- Incorrect model
  - **barely-resolved sources**
  - **self-cal can embed mistakes in the data**
- Bad data
- Images dominated by deconvolution errors
  - **poor boxing**
  - **insufficient uv-coverage**
- Not enough flux density
  - **fast-changing atmosphere**
- Errors which are not antenna-based & uniform across the image
  - **baseline-based (closure) errors (e.g., bandpass mismatches)**
  - **imaging over areas larger than the isoplanatic patch**
  - **antenna pointing and primary beam errors**

Next Class

- The Measurement Equation
- Imaging and Deconvolution
- Image Quality, Noise, Dynamic Range
- Wide-band imaging, wide-field imaging
- Advanced Calibration Issues

Today's Practicum 2



- Use your interferometer simulation script
- Add errors to your perfect array (gain, position, etc.)
- Plot array response in presence of such errors

Questions?