UNIVERSITY OF AMSTERDAM

Master Astronomy and Astrophysics - 5214RAAS6Y




Radio Astronomy

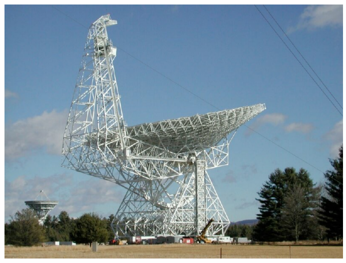
Lecture 7

The Techniques of Radio Interferometry II: Calibration

Lecturer: Michael Wise (wise@astron.nl)

May 8th, 2017



This lecture will be the first of two lectures on some of the analysis techniques used in radio interferometry.
Real data is very different than the pure mathematics of interferometry.
With experience, you can learn a lot of about your observation just from looking at the data itself.
We want to show you what real radio data looks like and begin to train your intuition.

Outline

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

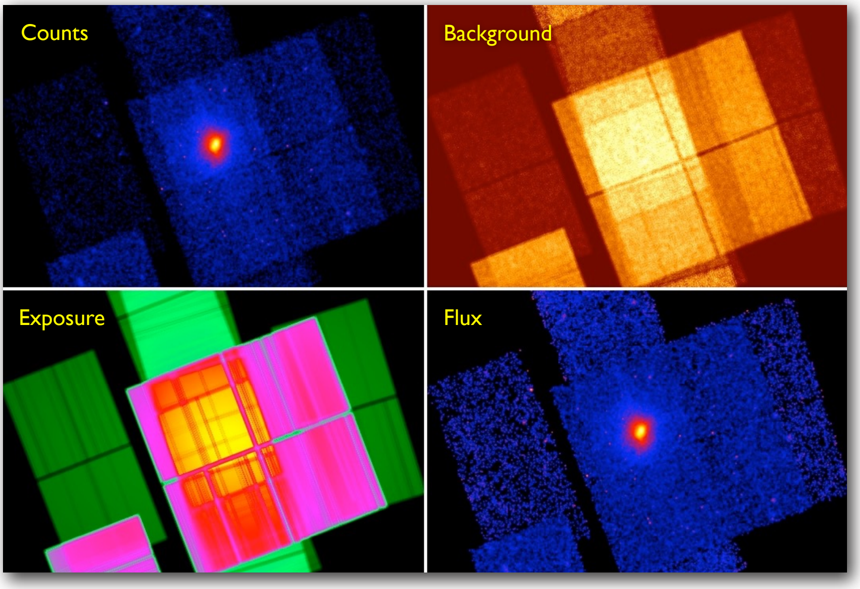
Definition of Calibration

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.*

Calibration means understanding both the instrument and the environment of your observation. Calibration is essential to separate what your data says about the astronomical source and everything else.
Calibration is the difference between being qualitative and quantitative.
Calibration is the difference between a wrong result and a correct one.

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

There are many sources of "error" or "noise" in an actual observation.
Some are related to the telescope itself and some may come from the surrounding environment.
Anything which is not related to the source of interest must be accounted for during the "calibration" process.
Calibration is a function of time, energy, direction, and instrument.

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

There are different techniques for calibration. Known or a priori calibrations are usually applied to the data before you begin your analysis.

We'll discuss two techniques, cross-calibration and self-calibration.

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

Almost any property of the source we want to study can be calibrated.
Flux density, position, and polarization are some of the more basic properties we try to calibrate.
We often rely on other well-known and well-studied sources to set the scale of these measure properties.

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:

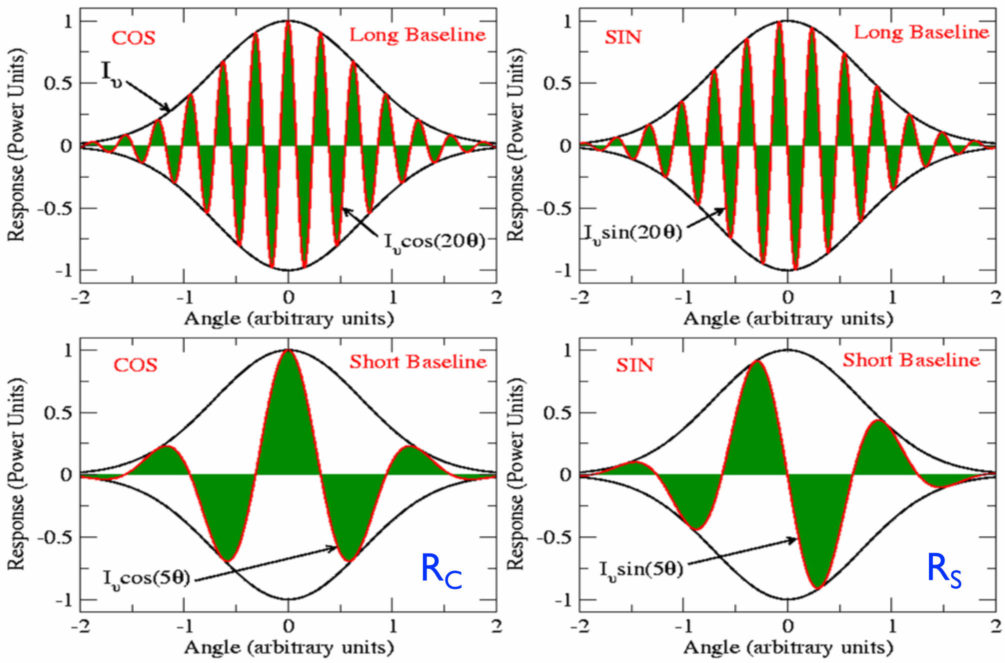
$$V_\nu(\vec{b}) = R_C - iR_S = \iint I_\nu(\vec{s}) e^{-2\pi i\nu\vec{b}\cdot\vec{s}/c} d\Omega$$

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

Jason discussed the interferometer equation in detail in lecture 6.
The visibility function (what we measure) is related to the sky surface brightness (what we want to know).
The measured visibility function is the Fourier transform of the sky surface brightness.
This equation assumes a "perfect" instrumental response.
Real telescopes are not perfect, so using this nice equation is tricky in practice.

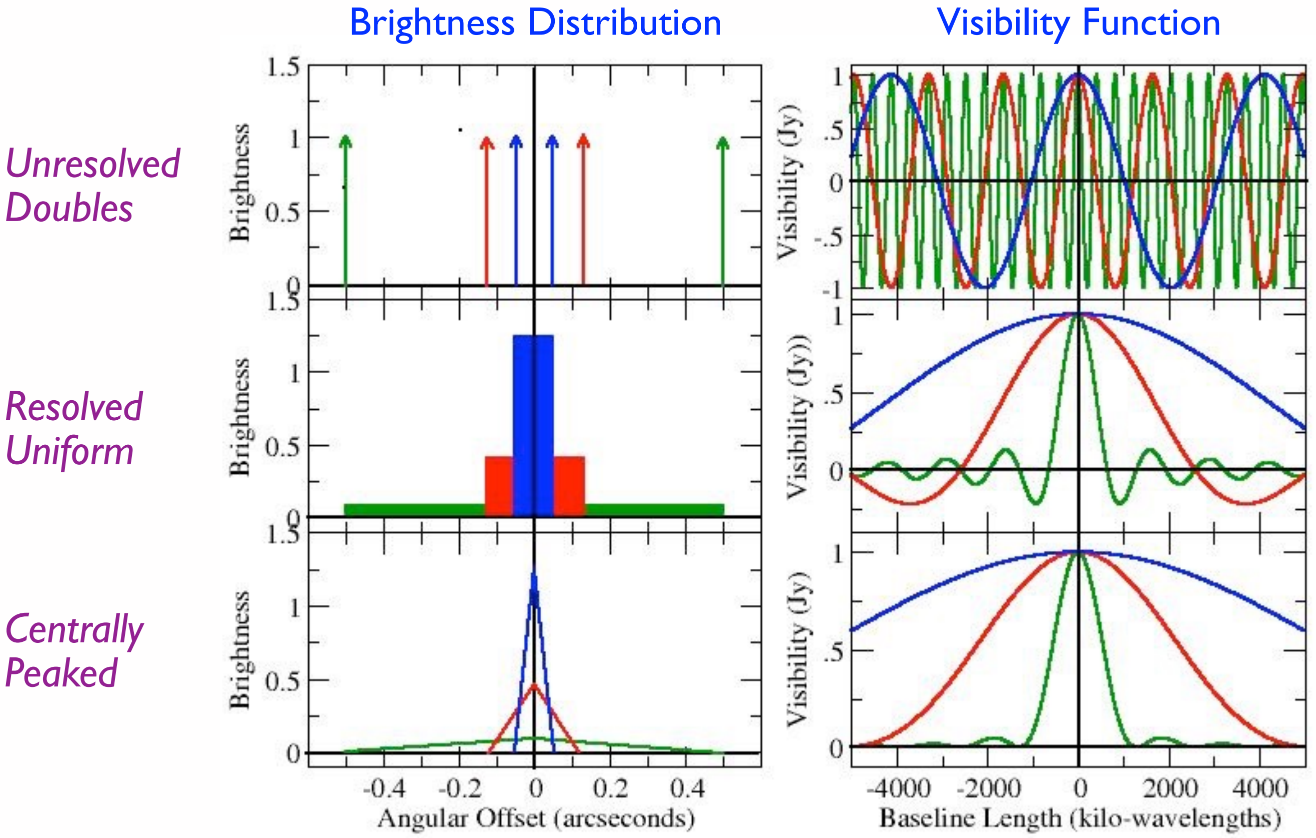
Visualizing Visibilities

- 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)



Visibilities are the "raw", unprocessed (mostly) output from an interferometer.
They are complex quantities with real and imaginary parts.
Can be expressed in terms of amplitude and phase.
For a simple point source, amplitude is related to the source brightness and phase is related to position.

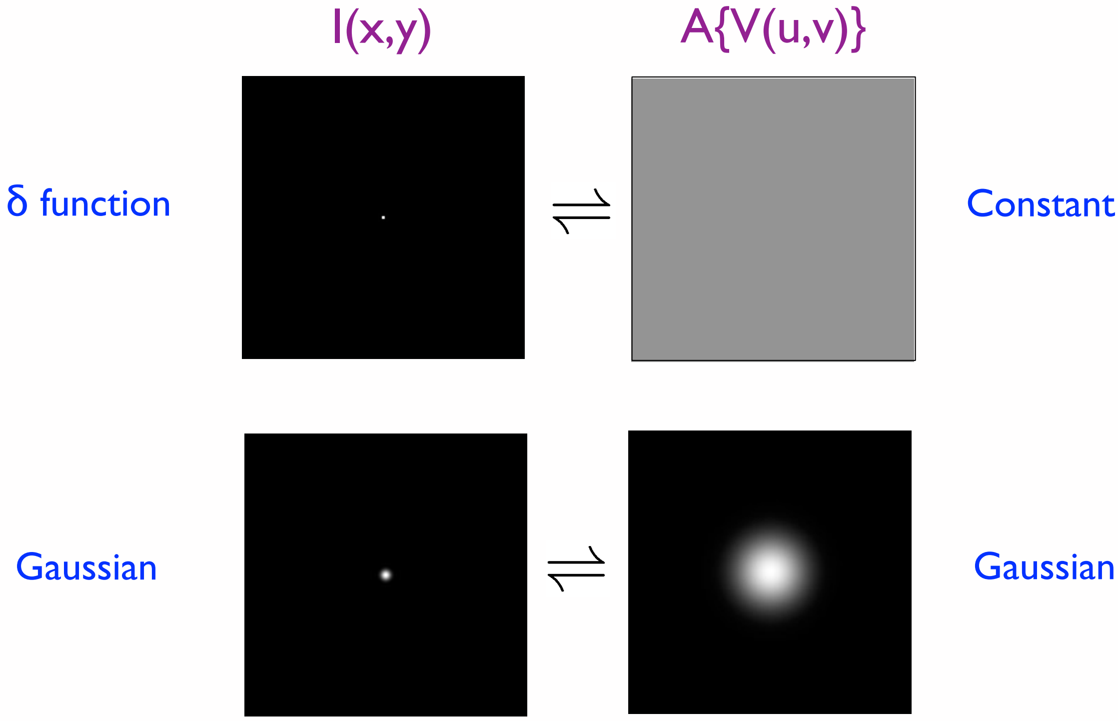
Simple Visibility Functions



The measured visibility pattern contains information about both the telescope and the source.
Characteristic source brightness patterns create recognizable visibility patterns.
With experience you can learn things about your source just by looking at the “raw” data.

Fourier Transform of *I(x,y)*

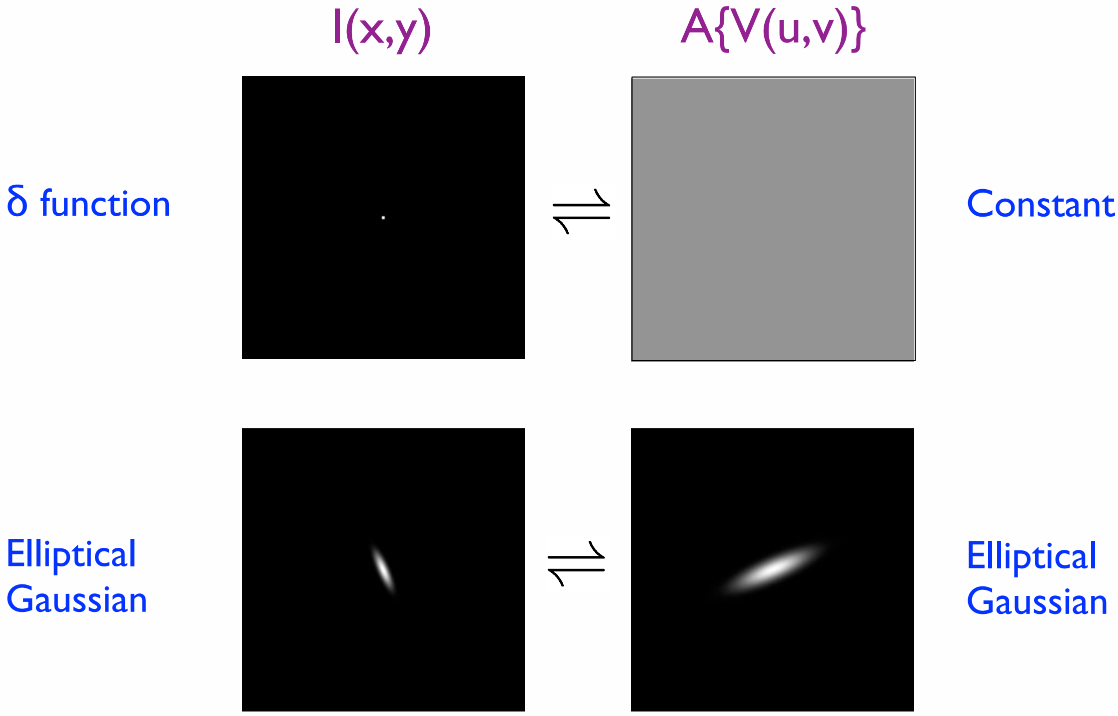
- V(u,v) is a complex quantity expressed as (real, imaginary) or (amplitude, phase)
- Narrow features transform into wide features (and vice-versa)



Some examples of simple source structures and their corresponding visibility functions.

Fourier Transform of *I(x,y)*

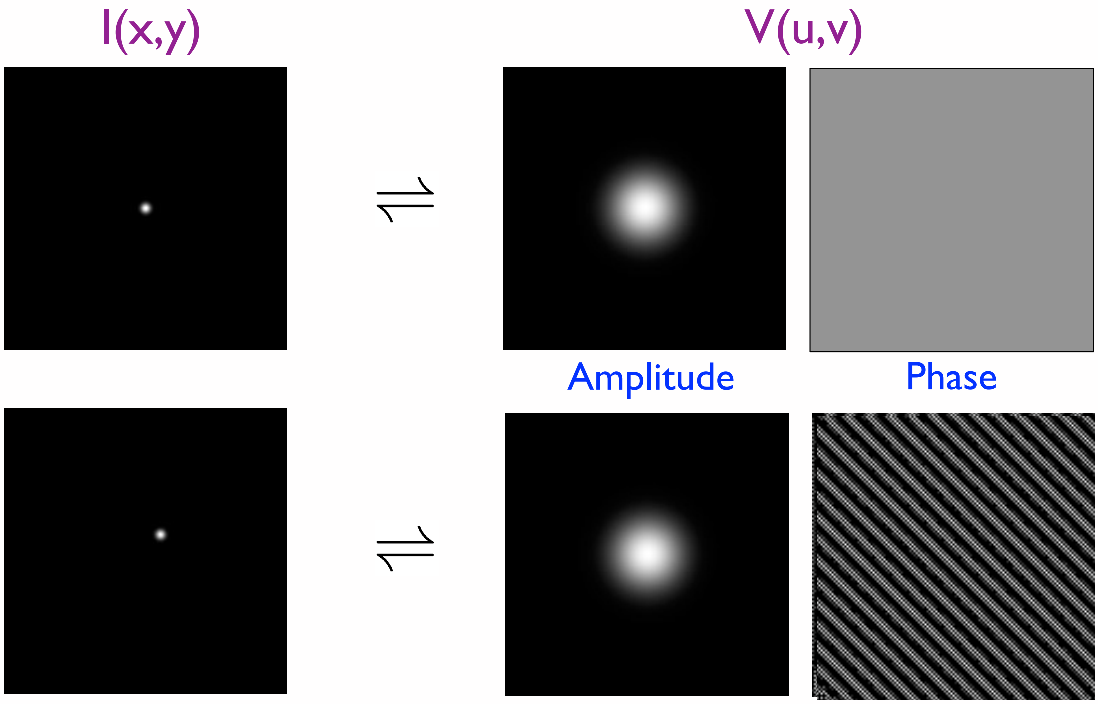
- V(u,v) is a complex quantity expressed as (real, imaginary) or (amplitude, phase)
- Narrow features transform into wide features (and vice-versa)



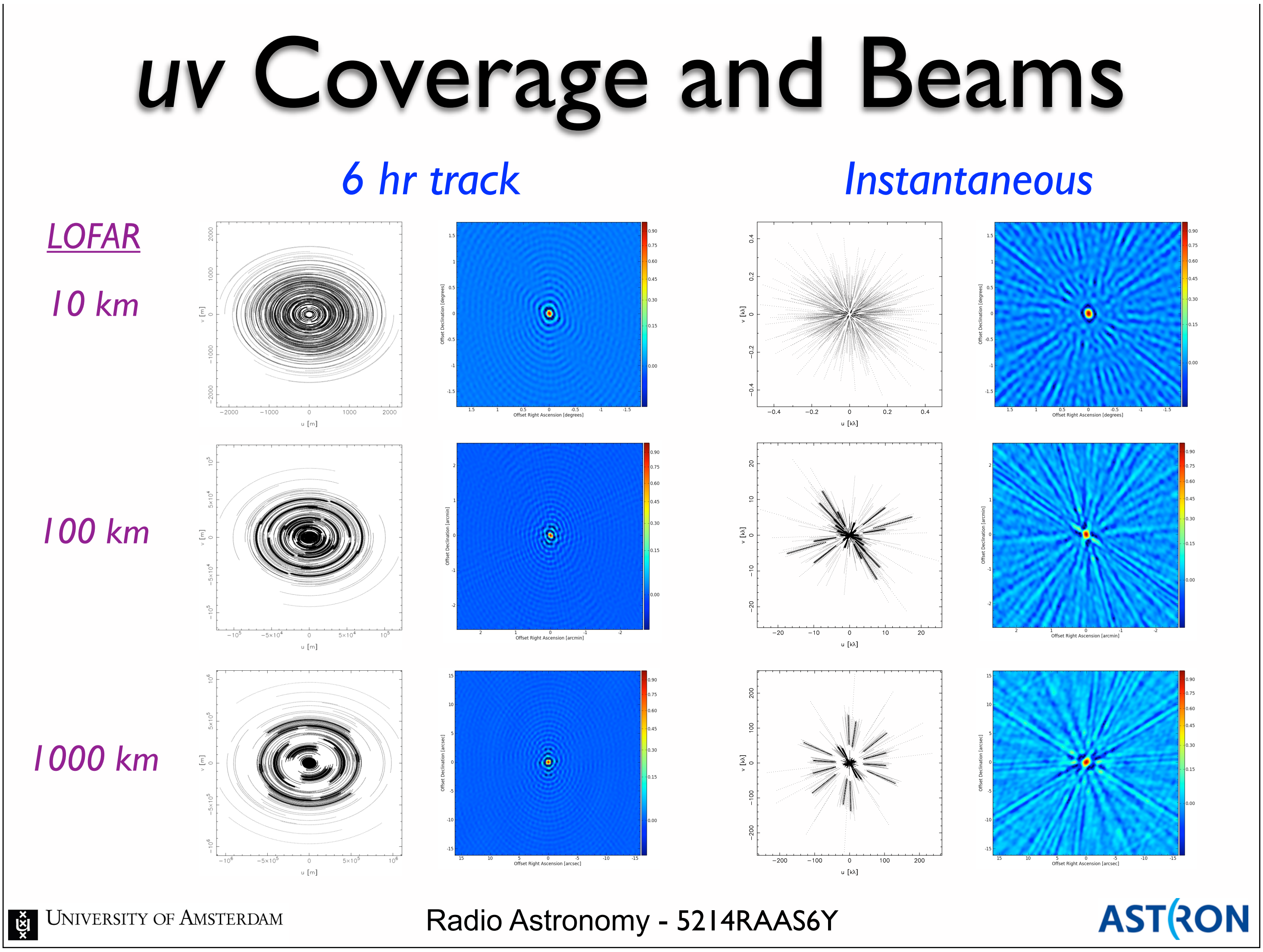
Some examples of simple source structures and their corresponding visibility functions.

Amplitude and Phase

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



Notice that the shift in position on the sky (relative to the pointing direction) produce a phase shift.



Some example visibility plots from LOFAR and the corresponding beam patterns they generate.
General rule of thumb, more complete coverage of (u,v) plane gives better beams.
Better generally means more symmetric and withouy strong sidelobes.
The (u,v) coverage of a given telescope is like its unique fingerprint.
You can learn to recognize what telescope was used just by seeing the (u,v) coverage plot.

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*

Visibilities are stored in the form of a big table. An entry for each time step, baseline, frequency channel, etc. The amount of data scales as the number of baselines squared. The size of the data from modern telescopes is becoming a real bottleneck for analysis these days. More on that in lecture 11.

Data Contents

- Usually presented to astronomer as Vij(ν,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

The data content for visibility data is very similar for different telescopes.
Mostly tables of the visibilities and various kinds of "metadata" (data about data) to describe them.
The format of the files, however, can be quite different for different telescopes.
Radio astronomy is moving toward standardizing these formats so common software can be developed.
Still not there yet, many different software packages around (CASA, AIPS, Miriad, LOFAR, etc.)

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

Newer telescopes like LOFAR (and eventually the SKA) are generating HUGE amounts of data these days.
Way more data than can be stored or reduced on a single laptop.
The single user, single desktop paradigm is increasingly rare.
The next generation of telescopes (LOFAR, ASKAP, SKA, etc.) will require supercomputers and HPC.
More in lecture 11.

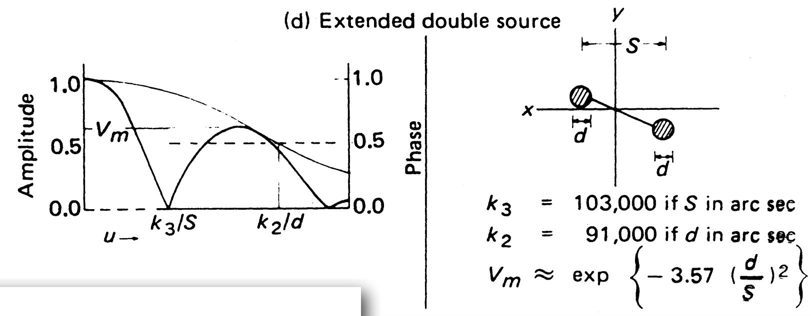
Inspecting Visibility Data

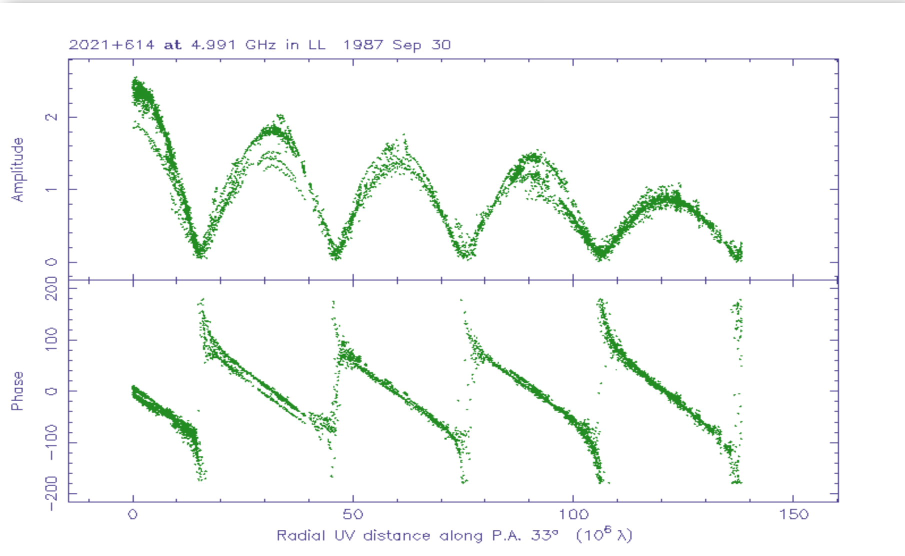
- **Useful visualizations**
  - Sampling of the (u,v) plane
  - Amplitude and phase vs. radius in the (u,v) plane
  - Amplitude and phase vs. time on each baseline
  - Amplitude variation across the (u,v) plane
  - Projection onto a particular orientation in the (u,v) plane
- **Advantages to inspecting *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

Inspecting your visibility data is a good habit to develop and can tell you how your analysis is going. Plot everything and plot often.
With experience, you will be able to see how your calibration is working just by examining the raw data.
You can also save yourself time by spotting problems before you spend a lot of time calibrating.

Estimating Initial Model

Can derive an initial model by inspection!

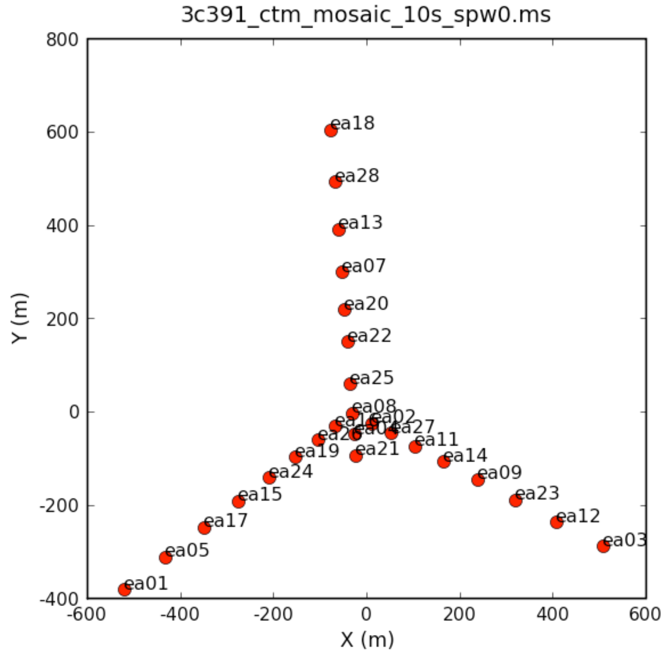




For simple source structures, you can learn to predict that source structure just from the visibilities.

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

We're going to look at some visualizations of a typical dataset from the VLA.
The dataset includes several pointings of our target as well as three calibration sources.

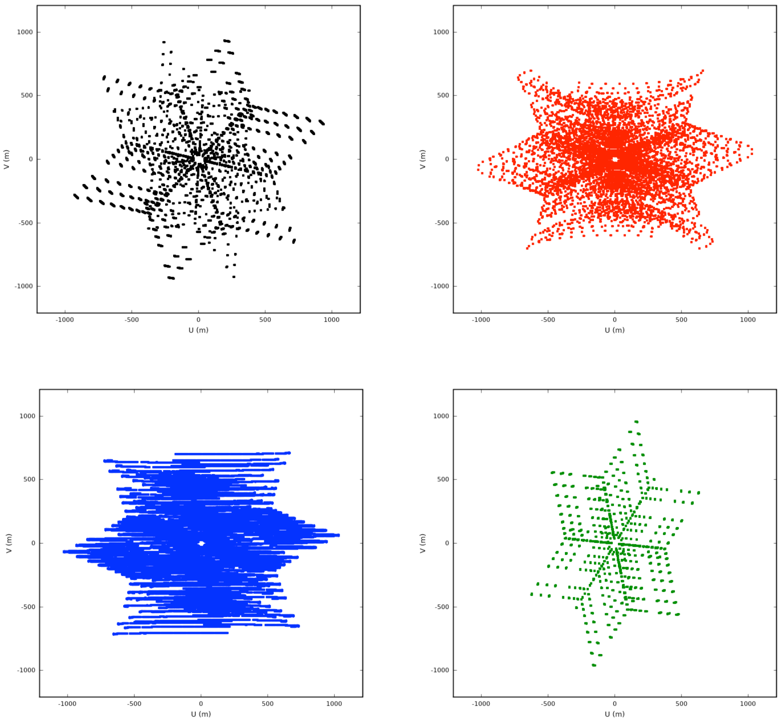
Observed *uv* Coverages

3C286
(Flux Density)

J1822-0938
(Gain Calibrator)

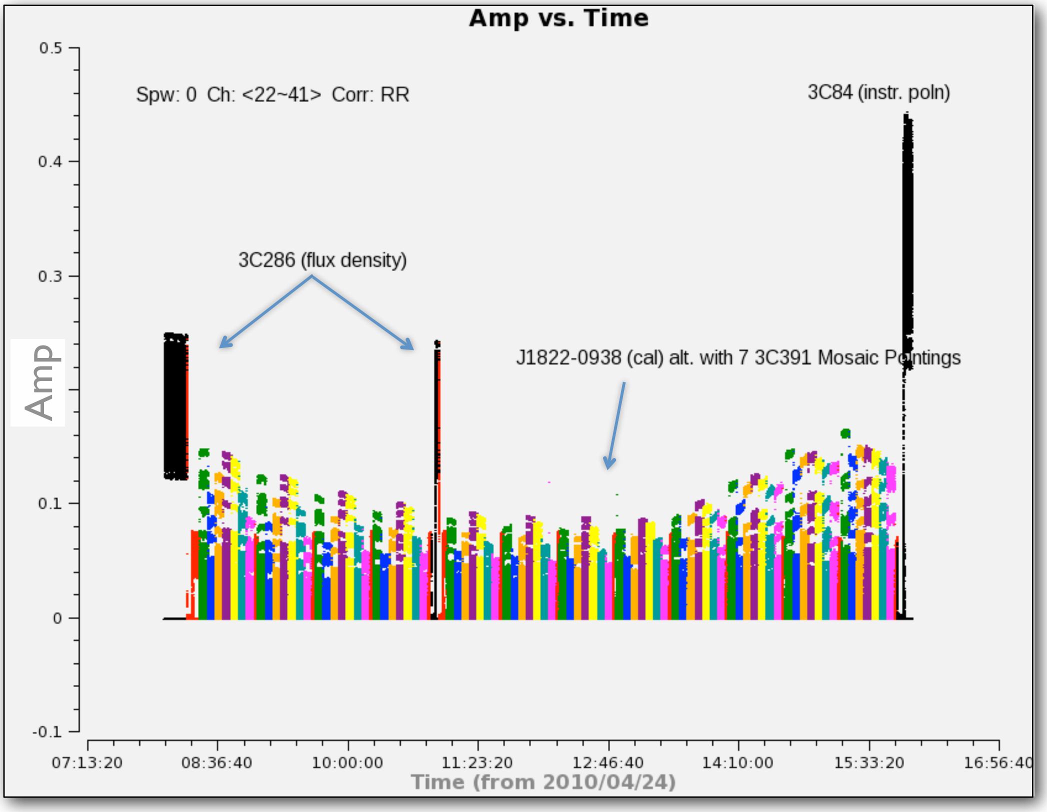
3C391
(Science Target)

3C84
(Pol. Calibrator)



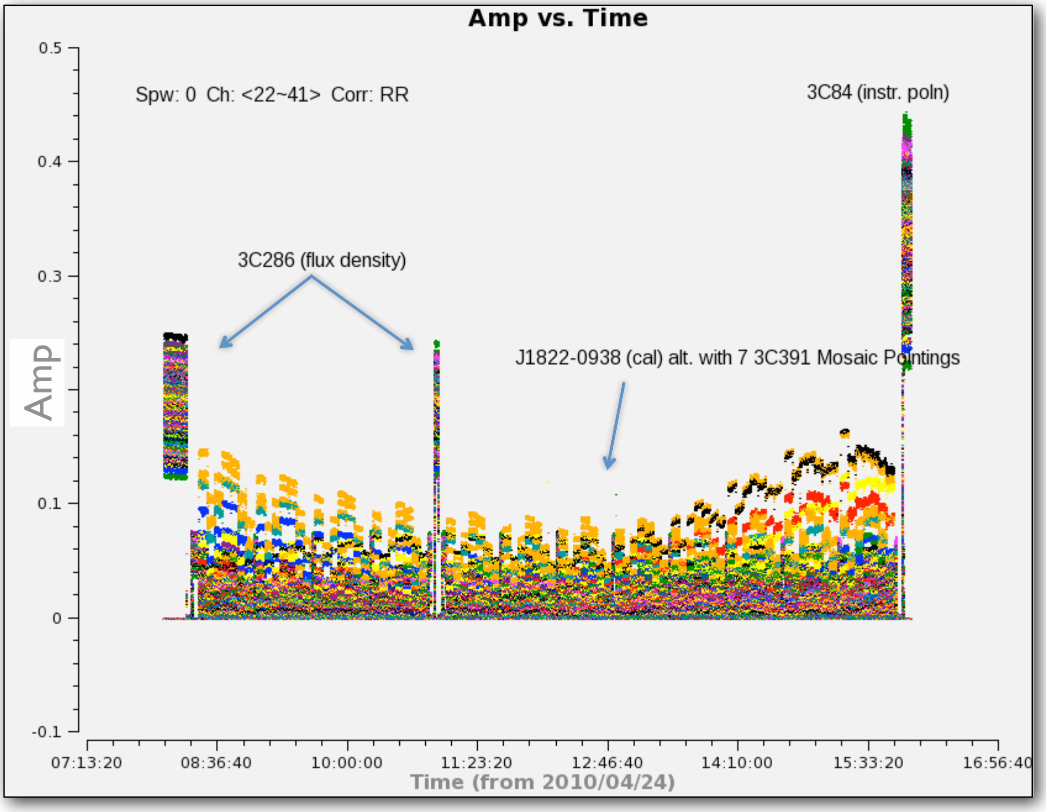
Plot of the (u,v) coverages for each source. Notice the distinctive “Y”–shaped (u,v) coverage.
Also notice how the calibrations are snapshot observations and the target has a more filled (u,v) coverage.
Finally, why do some of the (u,v) plots look squashed?

Visibilities (source colors)



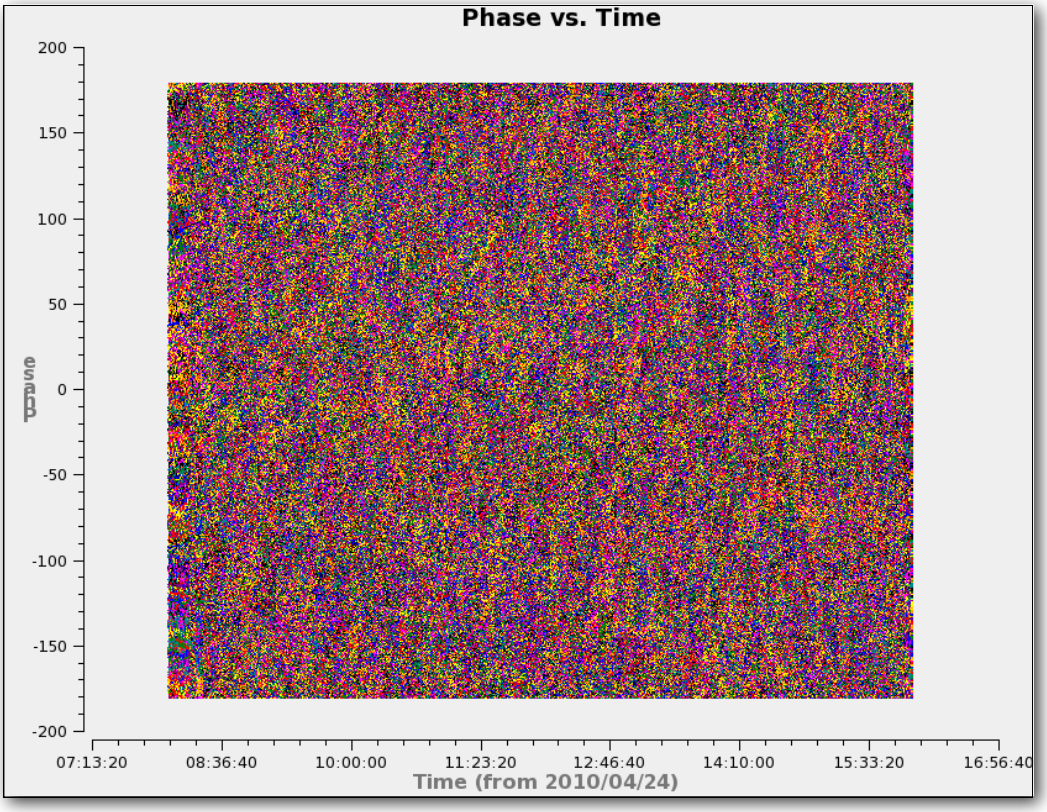
Plot of the visibility amplitudes for the entire dataset, color-coded for each source.
The different, multiple pointings for the target source each have a different color.
Notice the snapshot pointings of the calibrator sources.

Visibilities (baseline colors)



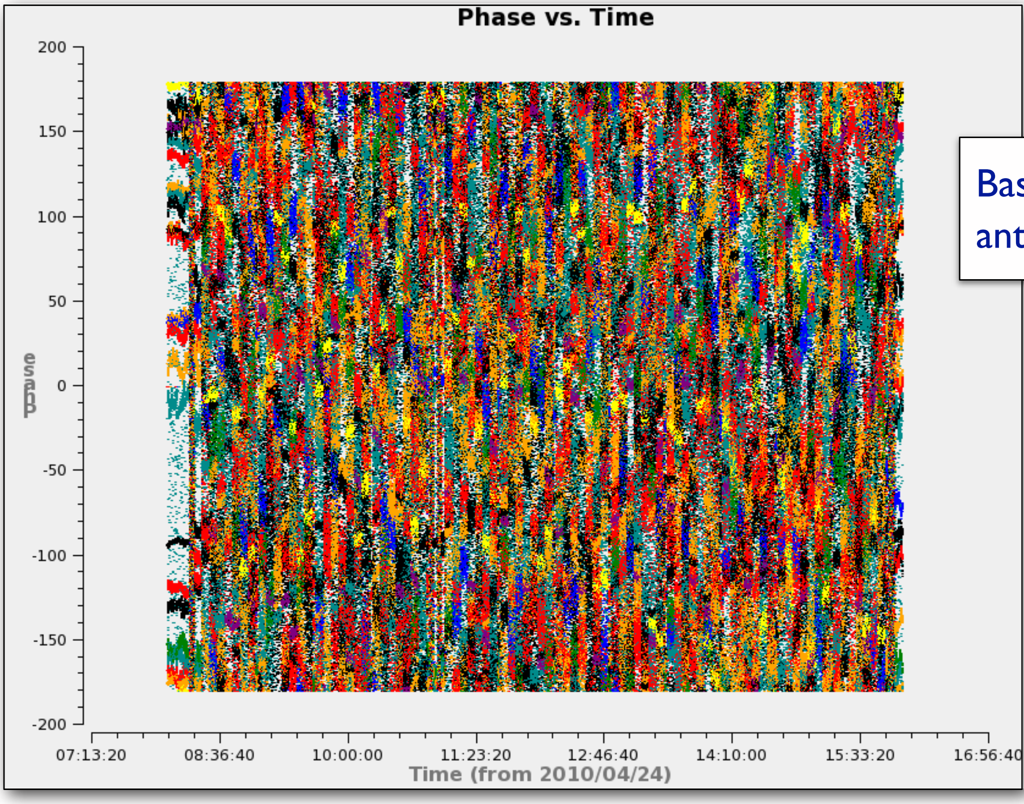
Same plot, but now the colors map to different baselines.
Notice some of the gaps in the bars. These correspond to baselines where there was little or no flux detected.
Could be related to the source structure, background, or calibration problems.

Visibilities (baseline colors)



Plot of all the visibility phases color-coded by baseline.
A bit hard to interpret!

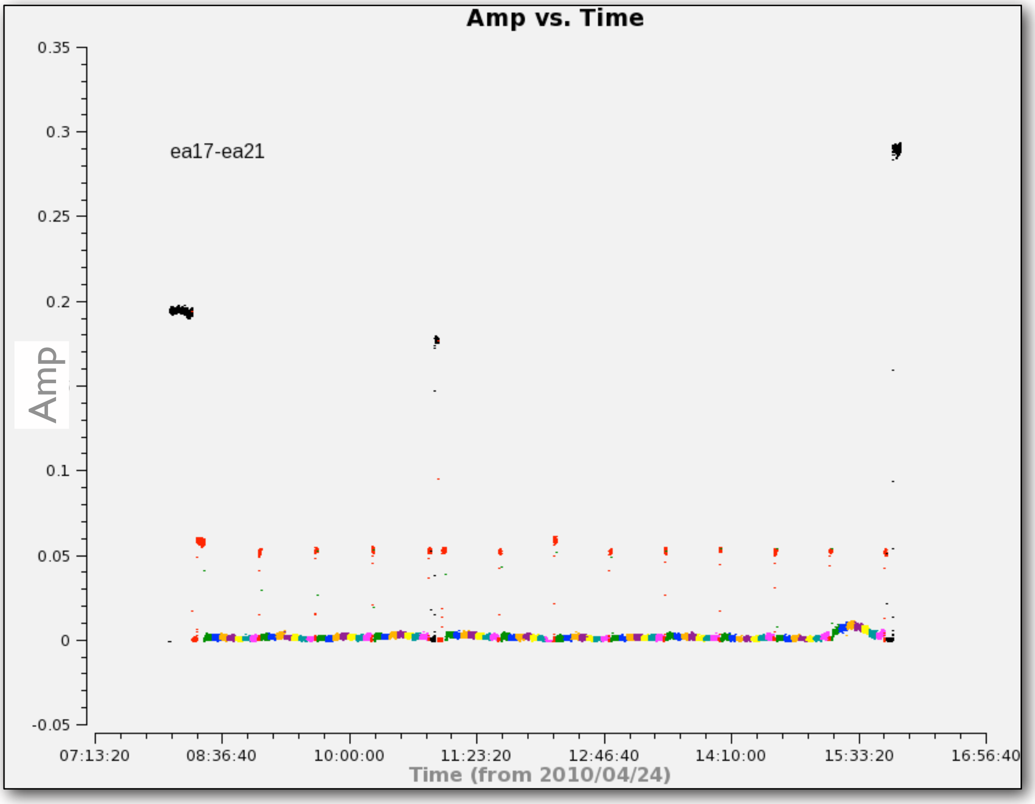
Visibilities (baseline colors)



Baselines to antenna ea21

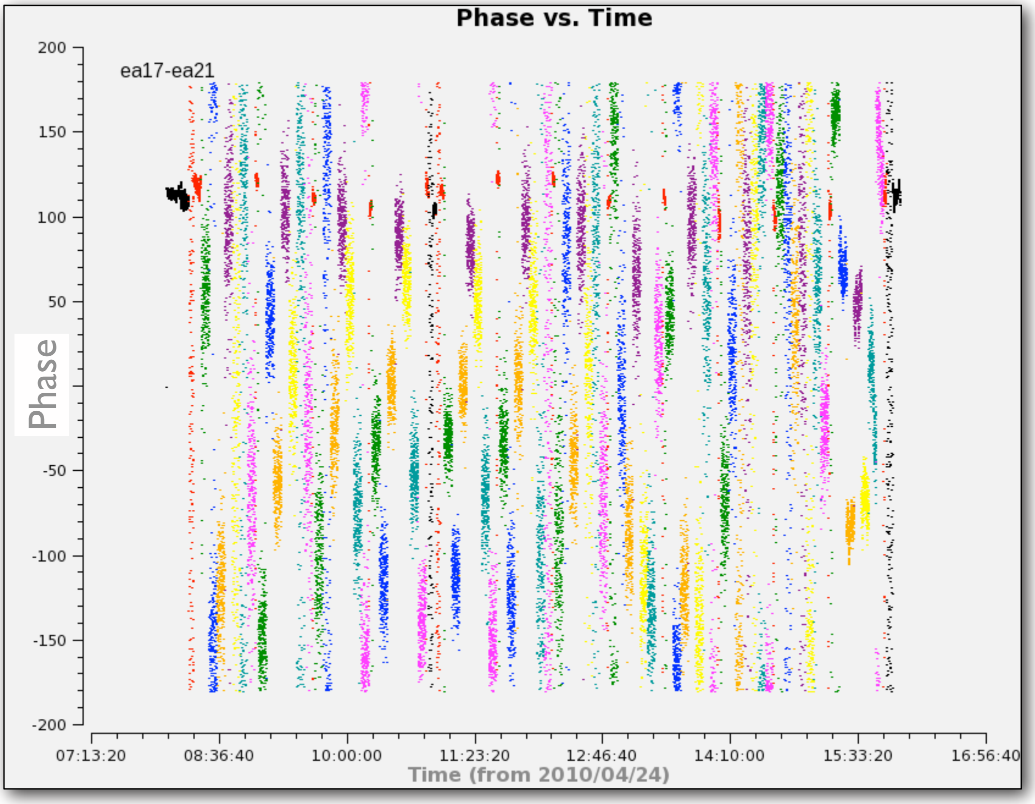
Same plot, but now only for those baselines that include antenna EA21.

Single Baseline (Amplitude)



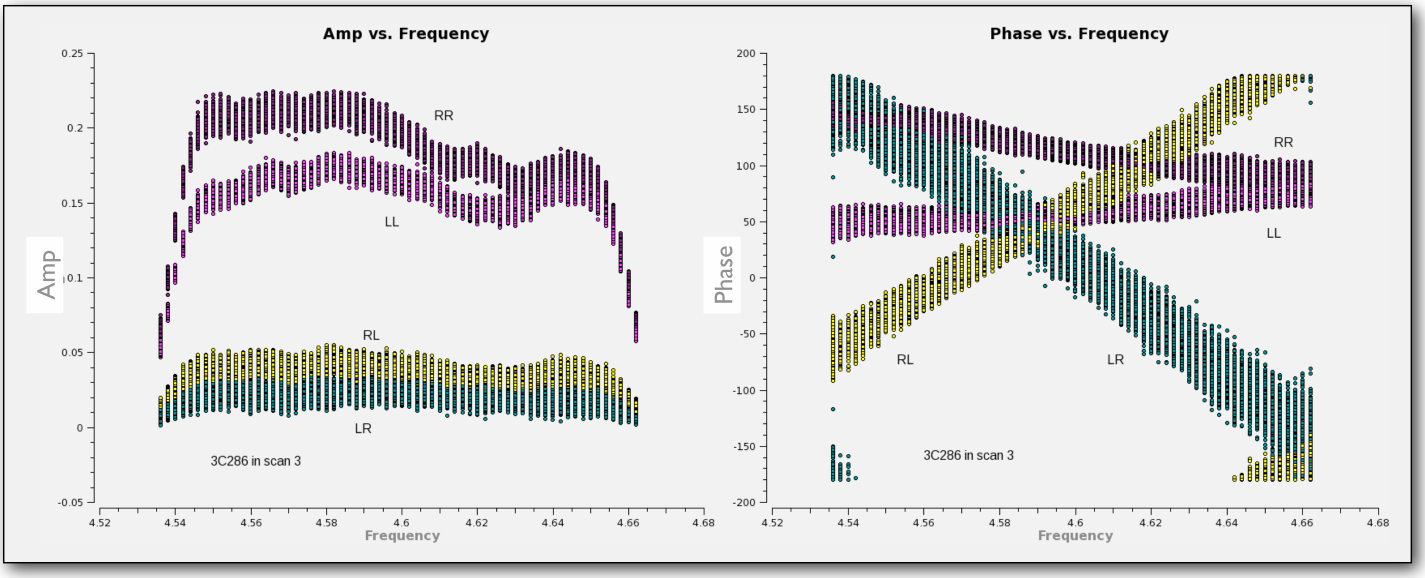
Amplitude versus time plot for a simple baseline.
Very easy to see the three calibrator sources now.
Notice that the amplitude of the calibrator J1822-0938 is very constant with time.
That's a good indication that the system was fairly stable over the course of observation.

Single Baseline (Phase)



Phases versus time plot for a simple baseline.

Single Baseline Spectra



Baseline ea17-ea21 (all 4 polarizations)

Spectra response of the system for a single baseline. One point for each target pointing.
This characteristic spectral shape is (partially) a function of the system.
We have to take out the inherent spectral response of the system to derive the true spectral shape of the target.

Intermission

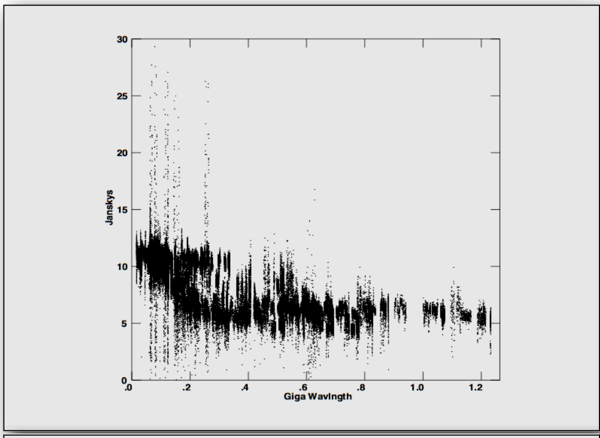
Data Editing

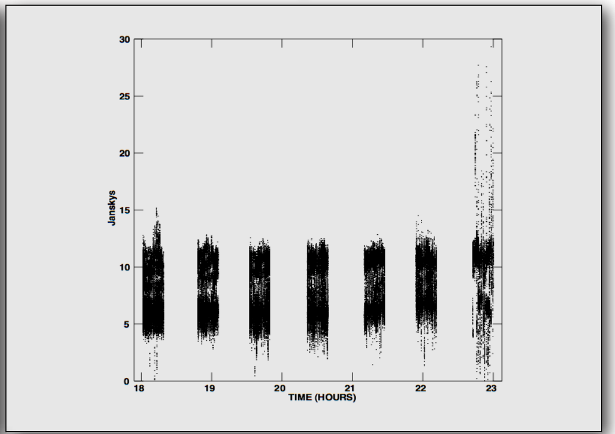
- Initial data examination and editing very important
- What to edit (much of this is automated):
  - Some real-time flagging occurred during observation
  - Any such bad data left over?
  - Any persistently "dead" antennas?
  - Periods of especially poor weather?
  - Amplitude and phase should be continuously 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…

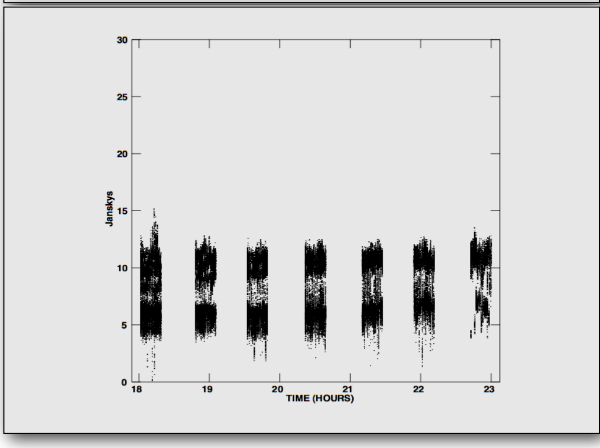
$\Rightarrow$ *Bad data is worse than no data…*

The first step in the calibration process is getting rid of bad data.
We call it editing, but the "bad" data is still there. We just flag it so that the software ignores it.
A little bit of bad data can ruin your whole calibration process (nearby sources of noise are often *much* brighter than the target source on the sky).
Can edit data by hand or increasingly using automated algorithms.

Editing Example





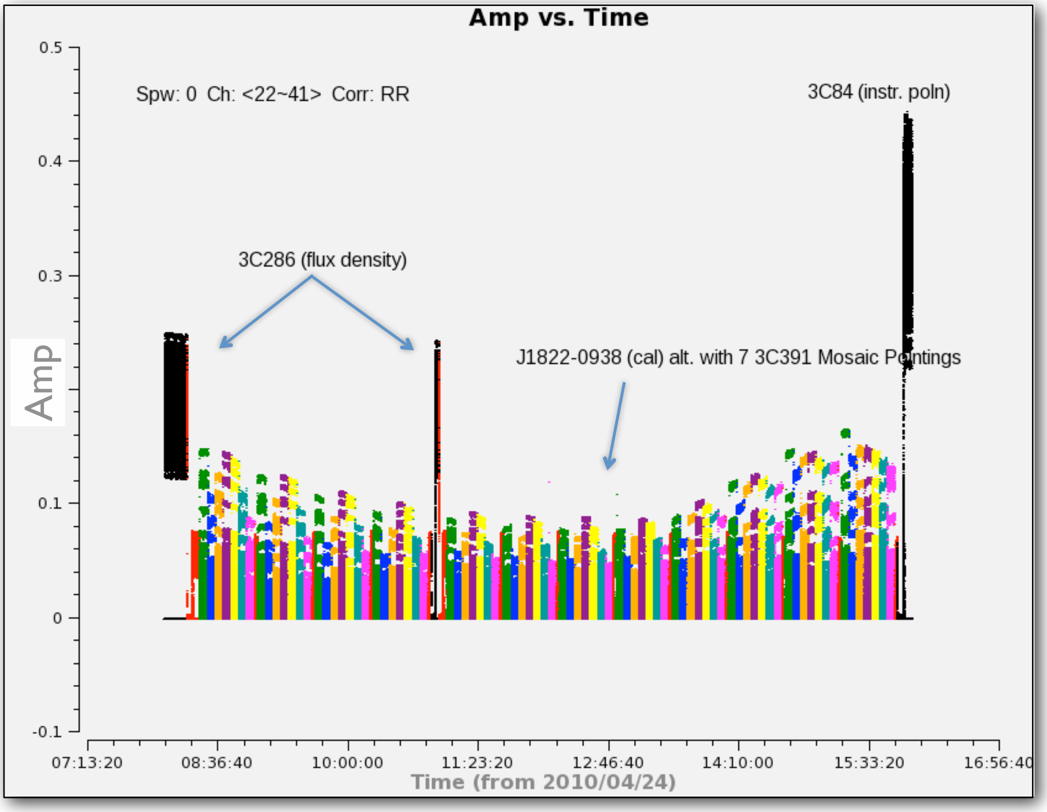


3C279 VLBA data at 43 GHz

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

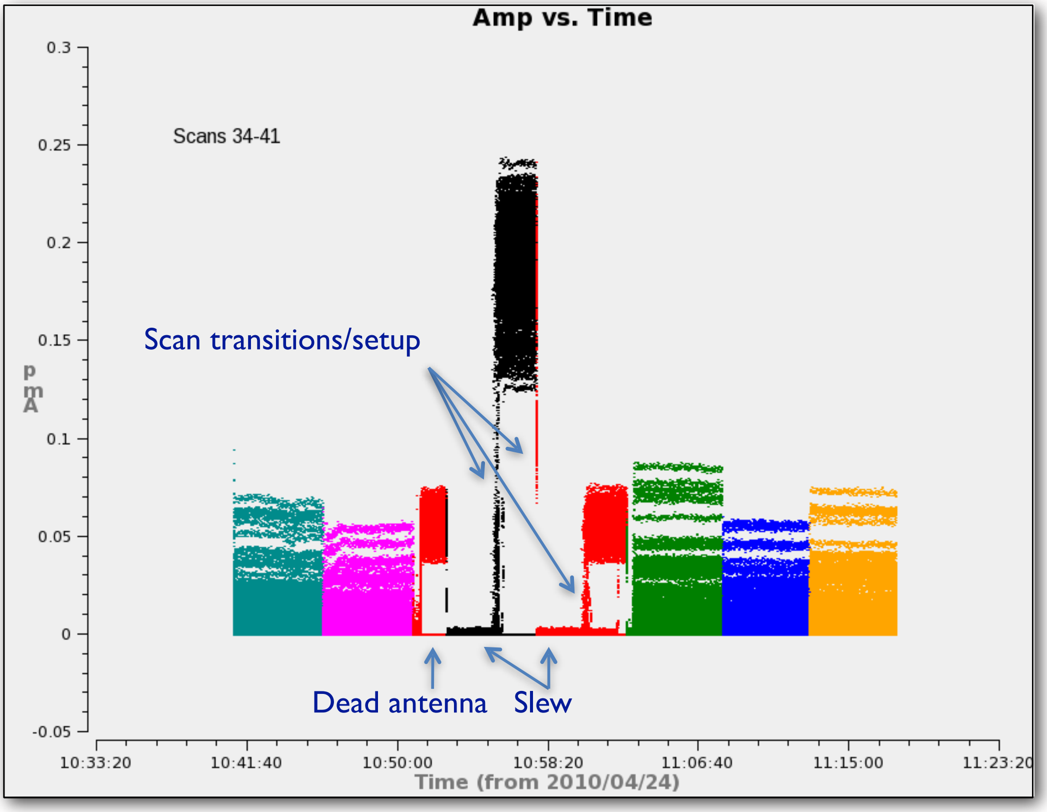
Plot showing visibility amplitudes as a function of uv distance and time.
Notice the outlier points with very high amplitudes. These are suspicious.
In the plot versus time, its clear that they all occur due one time slice (and one antenna).
After editing, the outliers are gone.

Editing Example



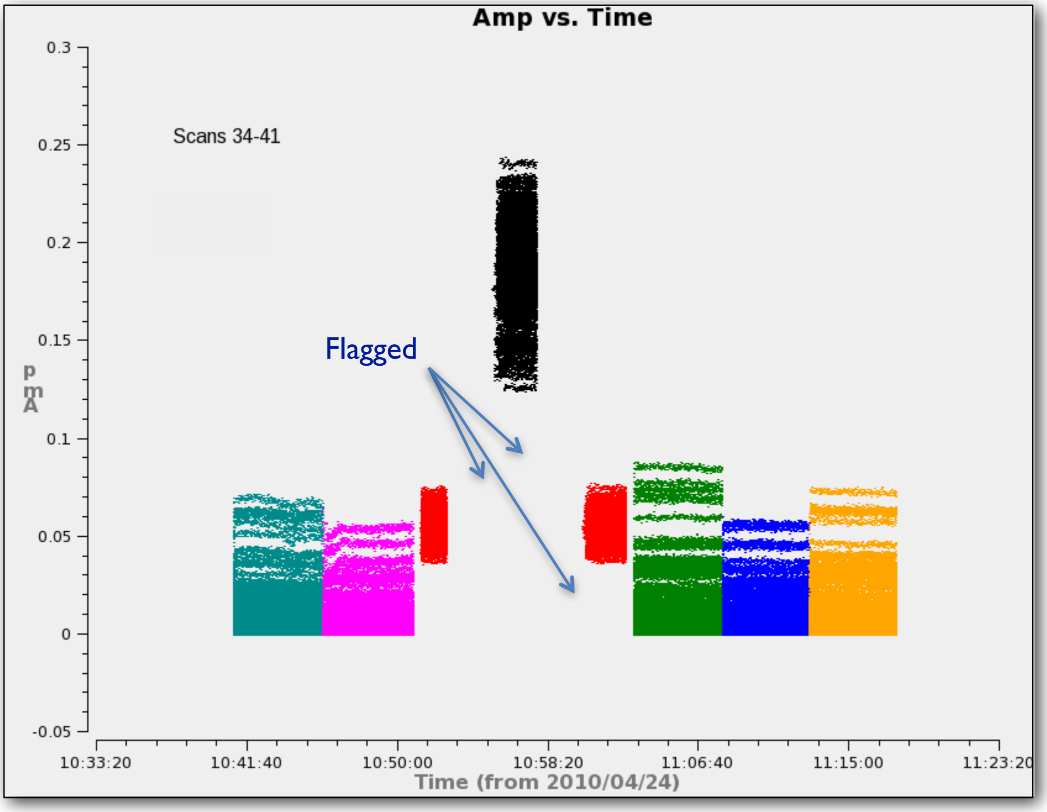
Same plot as before of VLA dataset for 3C391. Notice the gaps in some of the bars.

Editing Example



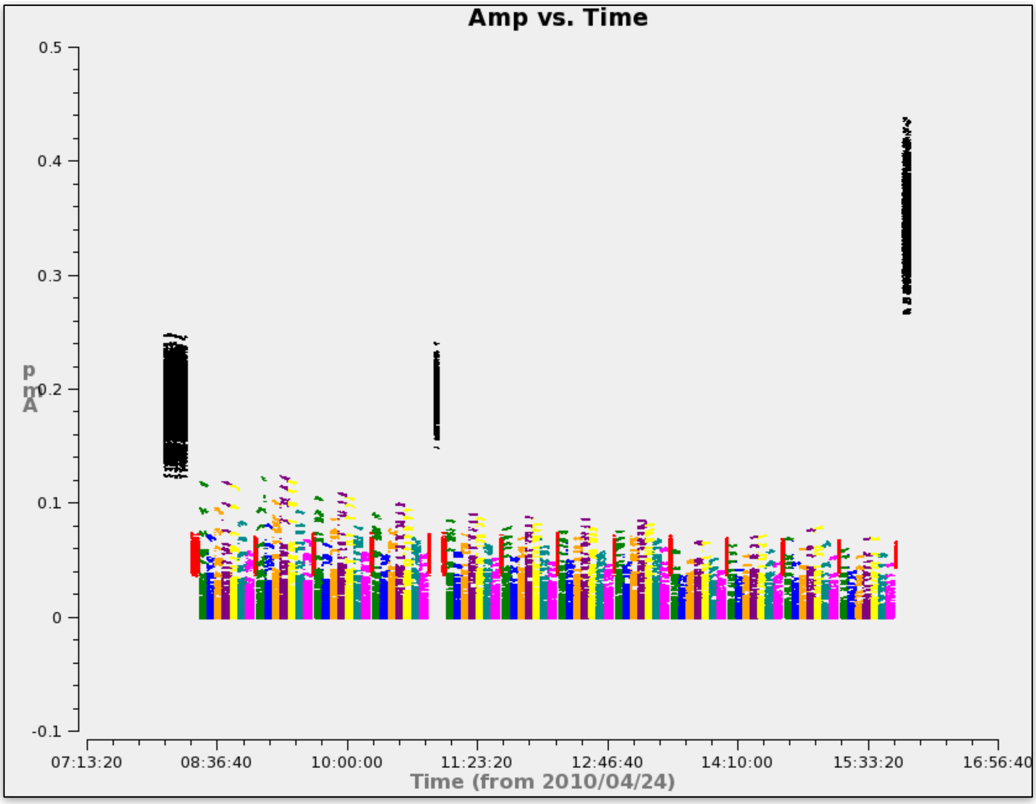
Zooming in on a few time steps.
The gaps are indicative of bad data for various reasons.

Editing Example



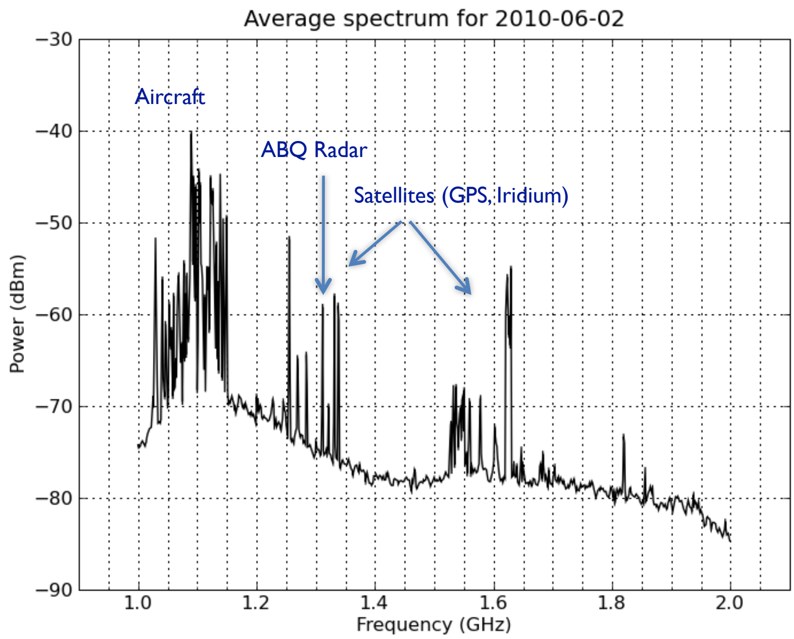
After flagging, these data points are not included in subsequent calibration or analysis.
Smooth ranges of data points with few outliers is indicative of good data.

Editing Example



Zooming back out, we see the full dataset after editing. The weird, partially populated gaps are gone.

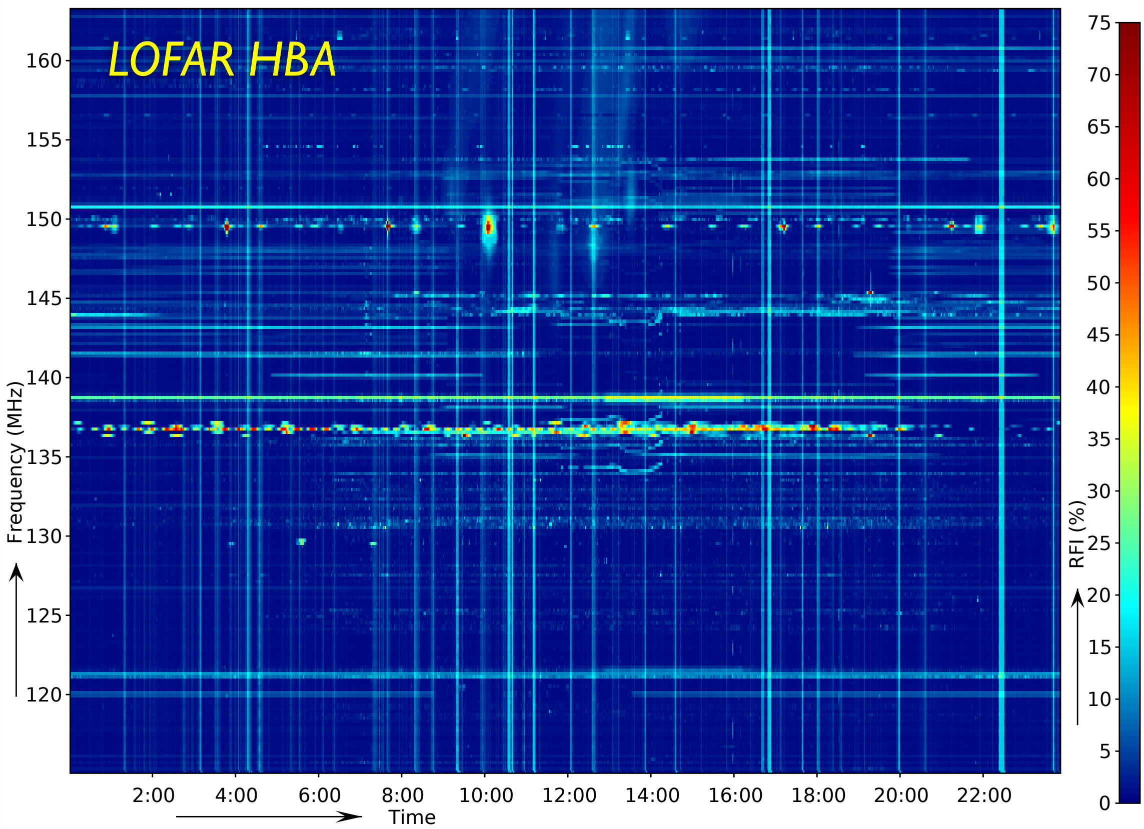
Radio Frequency Interference



- RFI originates from man-made signals
- Generated in the antenna electronics or by external sources (e.g., satellites, air traffic, cell-phones, radio and TV stations, automobile ignitions, microwave ovens, computers and other electronic devices, etc.)
- Adds to total noise power in all observations, thus decreasing the fraction of desired natural signal passed to the correlator —> reduced sensitivity.
- Can correlate between antennas if of common origin and baseline short enough, thereby obscuring natural emission in spectral line observations
- Least predictable, least controllable threat to a radio astronomy observation

RFI can completely ruin your observation depending on what frequency you want to observe.
Its a strong function of time and frequency.
Is usually orders of magnitudes brighter than celestial sources, so a little RFI will be brighter than your target.
Avoiding RFI is why we put radio telescopes far away from people. Its why the SKA will be built in the desert.
The equivalent of light pollution for optical telescopes.

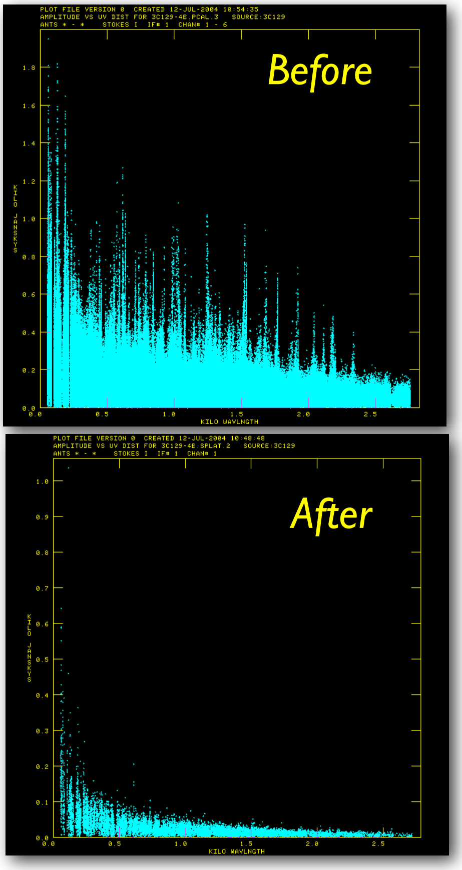
RFI Excision



Offringa et al. (2013)

An example of the RFI environment near the core of LOFAR.
Dynamic spectrum of frequency versus time.
Notice the thin lines. These represent narrow frequency RFI or short burst of RFI in time (or both).
This sort of RFI is fairly easy to filter out.

RFI Excision in Practice



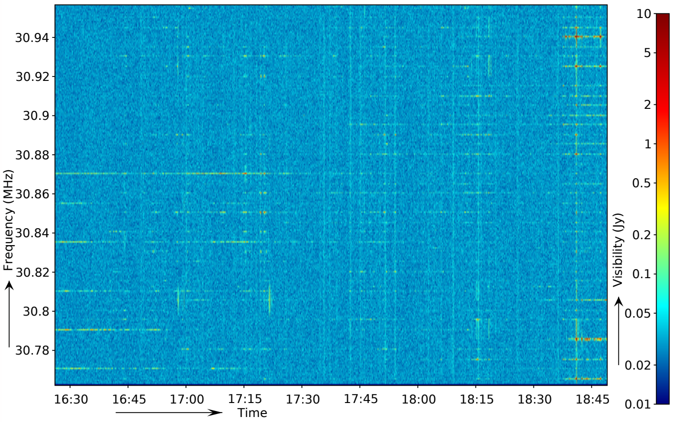
- RFI environment worse on short baselines
- Several 'types': narrow band, wandering, wideband, ...
- Wideband interference hard for automated routines
- Approach: averaging data in time and/or frequency makes it easier to isolate RFI, which averages coherently, from Gaussian noise, which does not
- 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

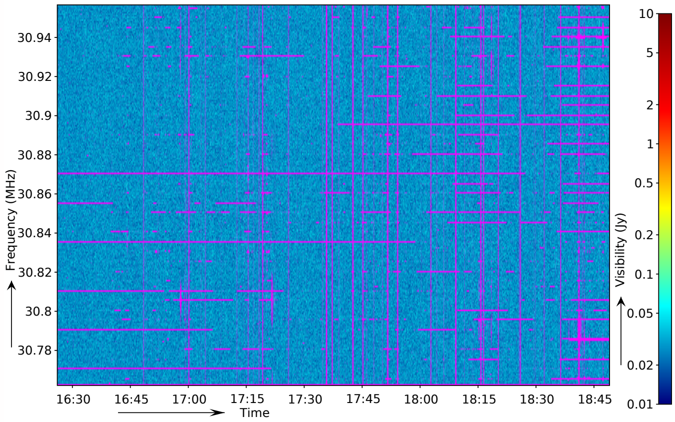
RFI detection and filtering algorithms can easily catch localized RFI (narrow in frequency or time).
Use “sigma-clipping” algorithms to remove all data that is statistically much above the mean.
These algorithms don’t work as well for broad-band RFI.

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

Results showing how well the automated RFI flagging works for LOFAR data.
The green, thin lines on the left indicate RFI.
The purple colored regions on the right are what the system automatically flagged as RFI.
LOFAR loses less than 5% of its total bandwidth to narrow-band 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

This equation gives the basic calibration equation to be solved.
It relates measured visibilities on the lefthand side to the true visibility function on the righthand side.
The gains for each antenna need to be determined in order to invert this equation and determine the true V's.
Solving this set of coupled, linear equations is the calibration process.

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

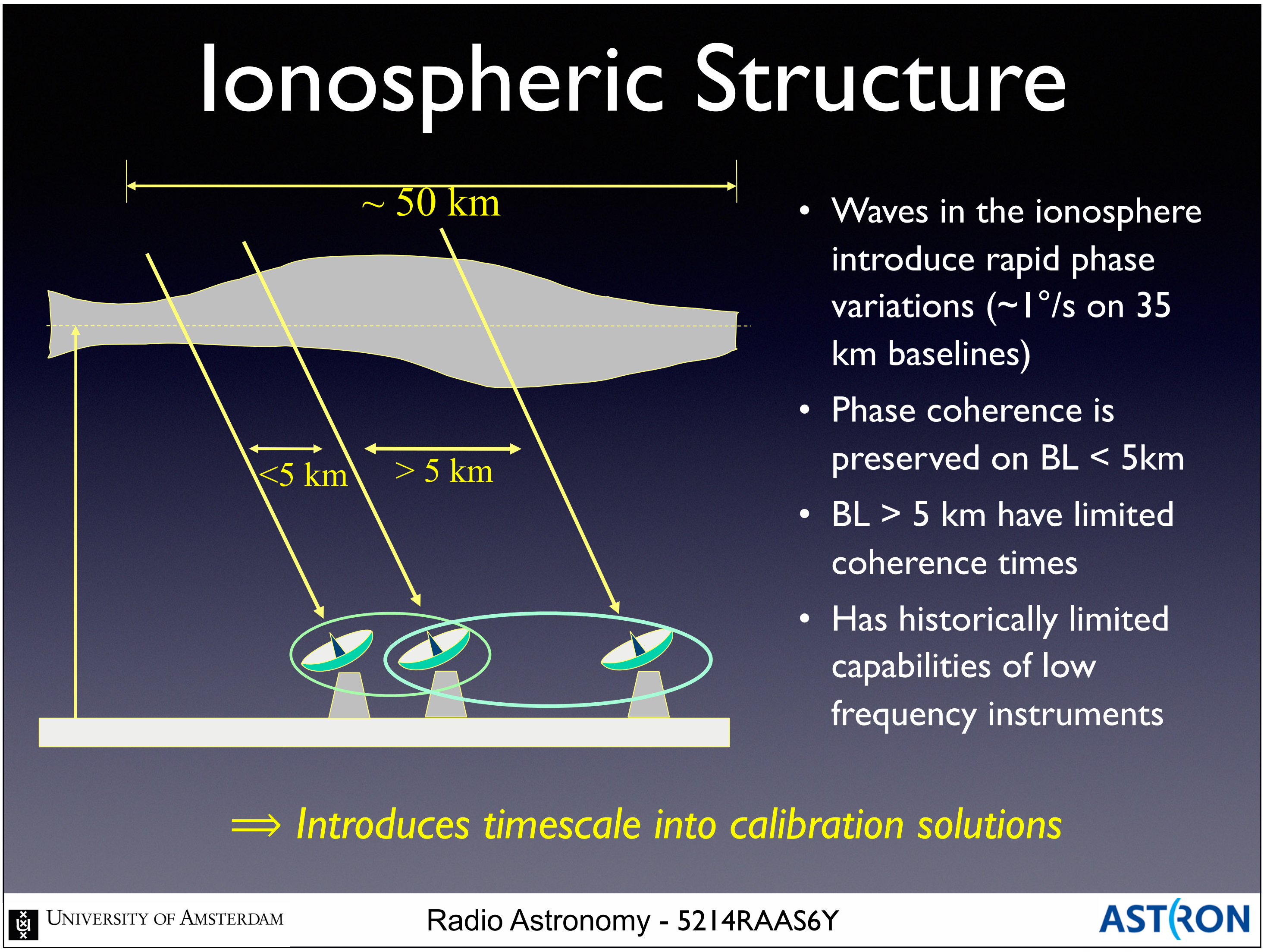
- Solve for antenna gains via least squares algorithm
- Works well - lots of redundancy
- (N-1) baselines contribute to gain estimate for any antenna
- Gains are antenna based and direction independent
- Also know as "cross calibration"

The true visibility function depends on the nature of the source.
Extended sources have different intensities on different scales.
For a point source, its easier since the source is the same on all baselines.

Why is *a priori* calibration insufficient?

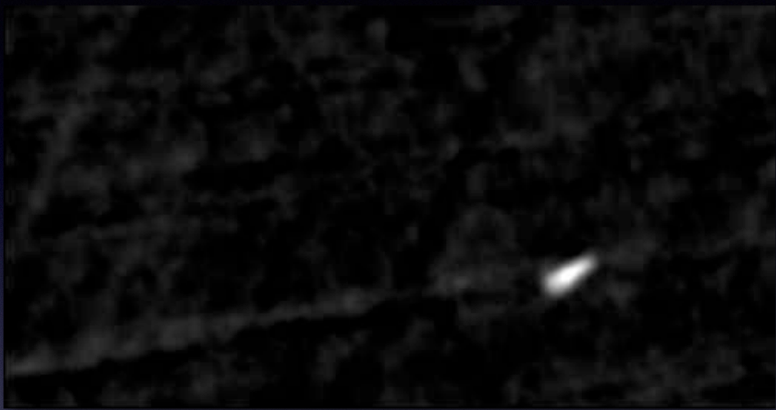
- Initial calibration based on calibrator observed before and/or after target
- Gains were derived at a different time
  - Troposphere and ionosphere are variable
  - Electronics may be variable
- Gains were derived for a different direction
  - Troposphere and ionosphere are not uniform
  - > 1 GHz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

Why can't we just measure the gains for each antenna once and be done with it?
The physical properties of the system itself change with time (due to temperature, RFI, etc.)
Environmental effects like the ionosphere and troposphere vary with time and position.
Generally you need the gains derived near your source observation (in time and on the sky).



The changes in the ionosphere introduce a timescale into our calibration.
Timescales are related to the size of turbulence regions in the ionosphere.
Typical sizes for these regions can be 10’s of kilometers.
These disturbances translate into timescales for things to change of minutes to seconds.
The effect of the ionosphere is a strong function of wavelength, makes low frequencies very difficult.

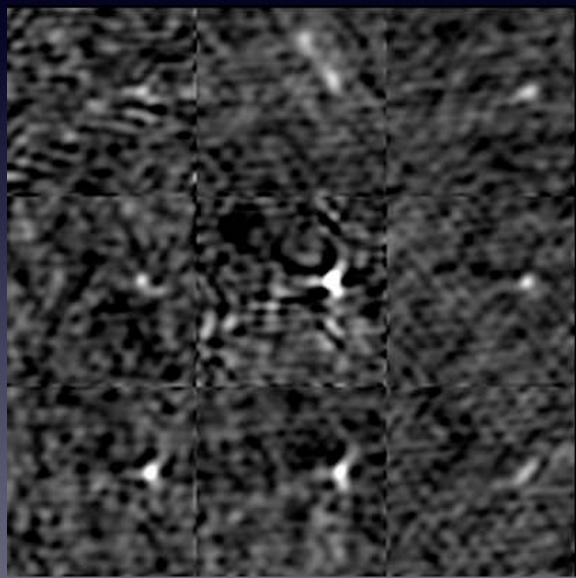
Effects of Ionosphere



Virgo A over the VLA (1.4 GHz)

- 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

Multiple sources in same field



(Courtesy B. Cotton)

A movie of the effects of the ionosphere on the image quality of a point source observed with the VLA.
The ionosphere produces changes in position, intensity, and source shape.
The good news is that if we record our calibration information on the right timescales, we can take these effects out during the calibration step.

Calibration using a complex model

- Don't need point source - can use model

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

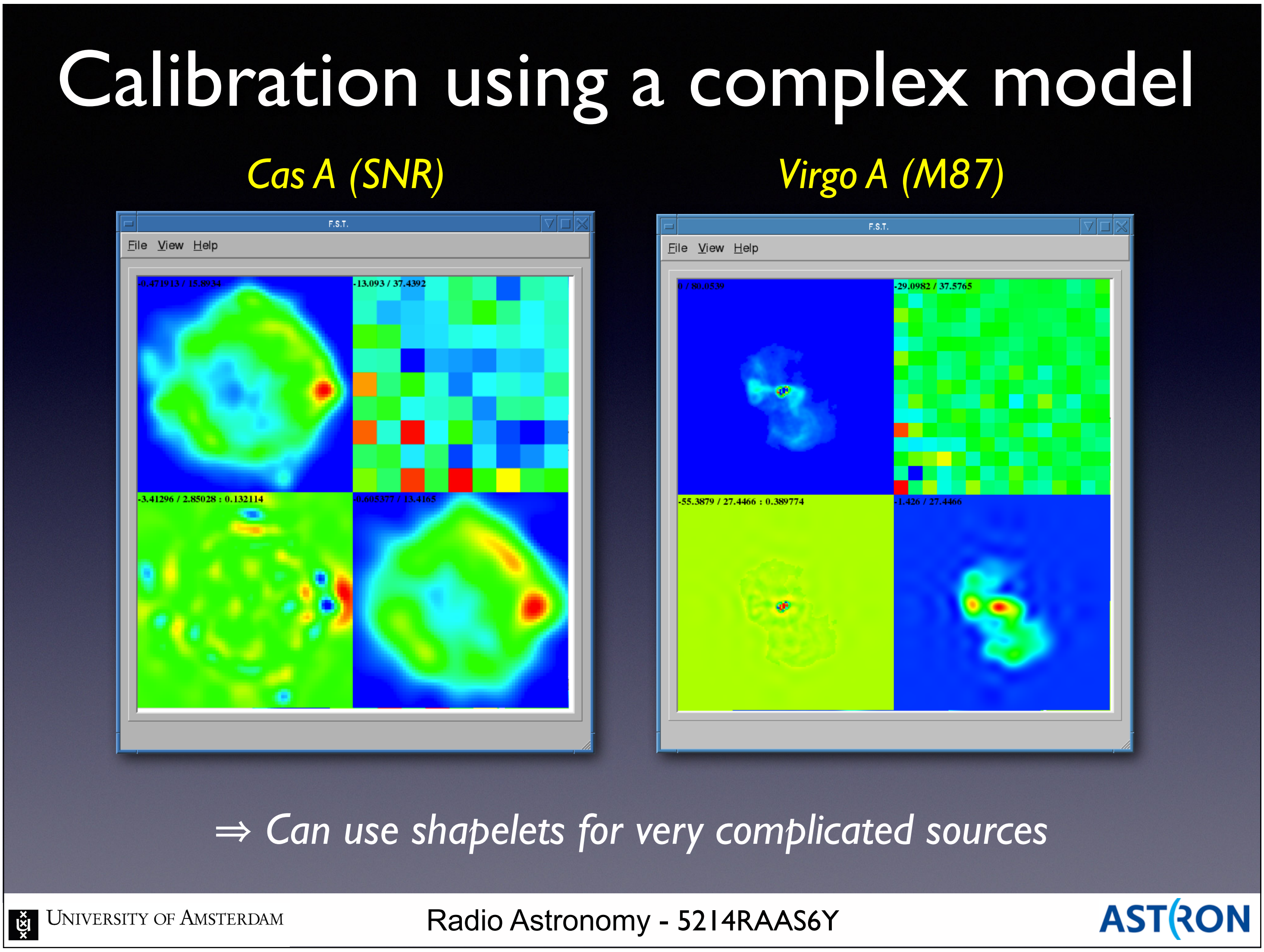
V_{ij}^{model} Model visibility

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

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

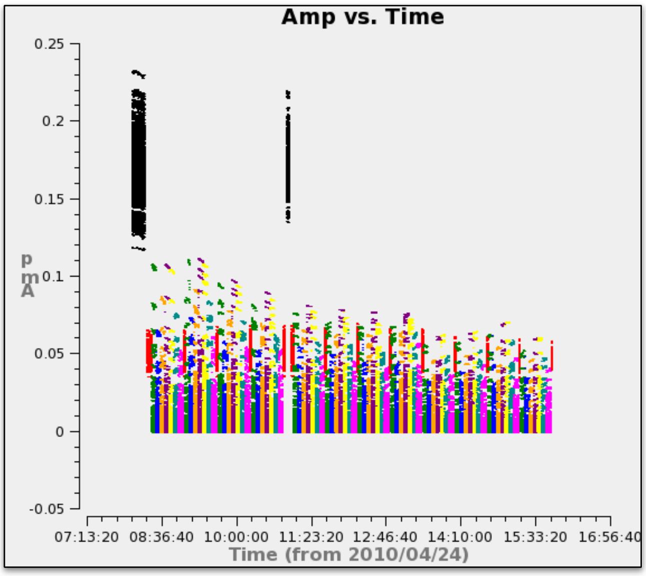
- Can smooth or interpolate gains if desired (be careful!)

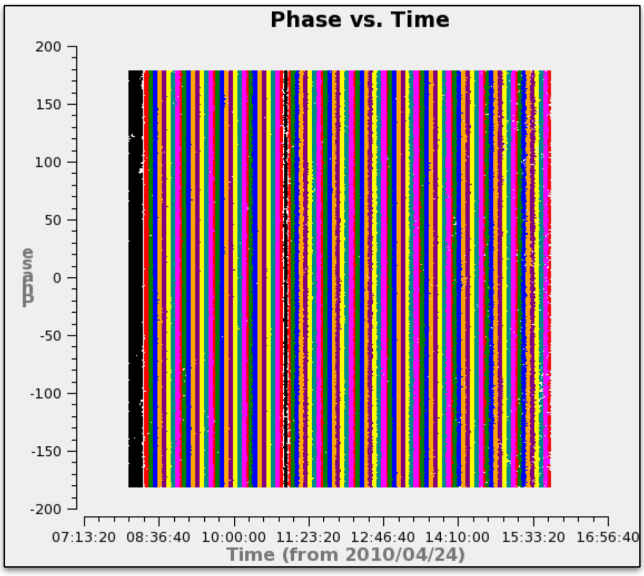
Many sources are not point sources, or at least not point sources on all baselines.
Can substitute a more complicated model for the true visibility function.
A more complicated model for the source means additional unknowns to derive.
We can get away with more complicated source models because we often have many constraints (baselines).



An example of extended sources where we need to use more complicated source models.
Can think of creating a model of any arbitrarily complicated source as the sum of many point sources.

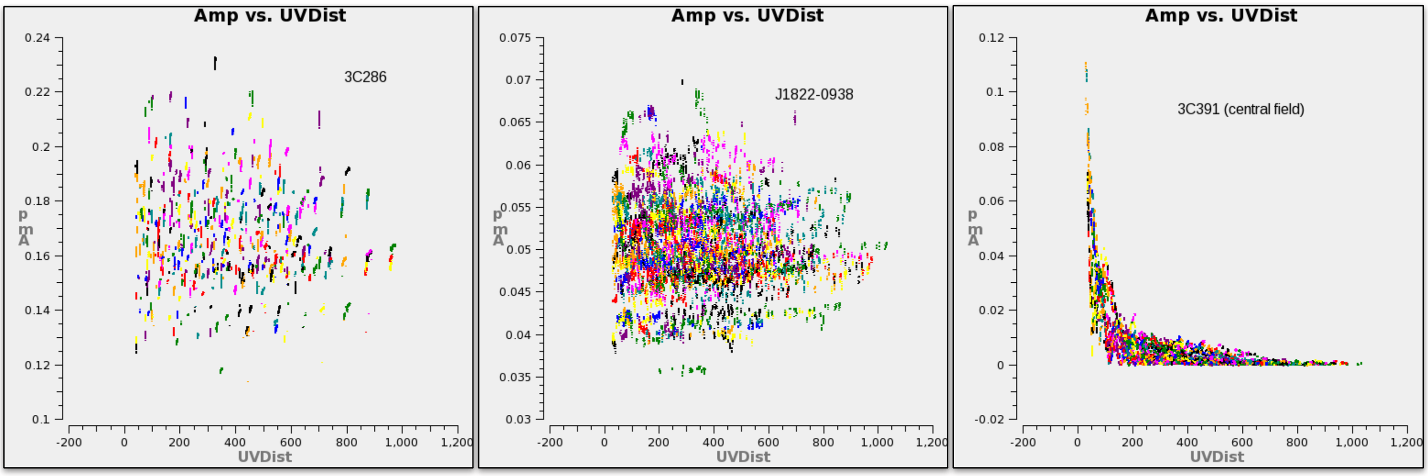
Uncalibrated Data





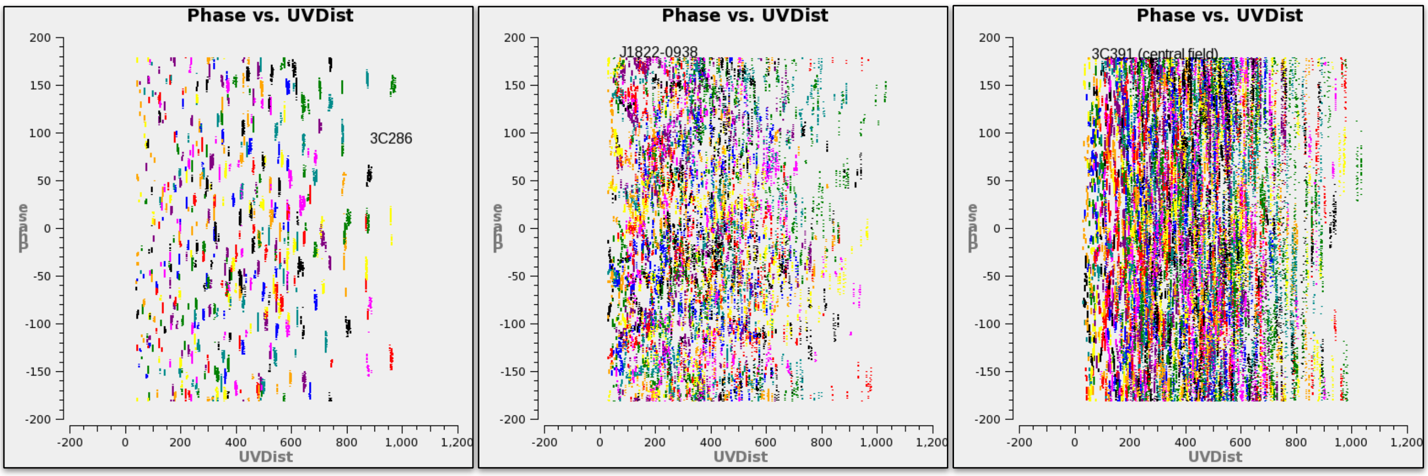
Lets look at the actual data before and after the calibration process.
Same plots as before of the amplitude and phase for the VLA dataset for 3C391.
Data is shown after editing and before calibration.

Uncalibrated Data



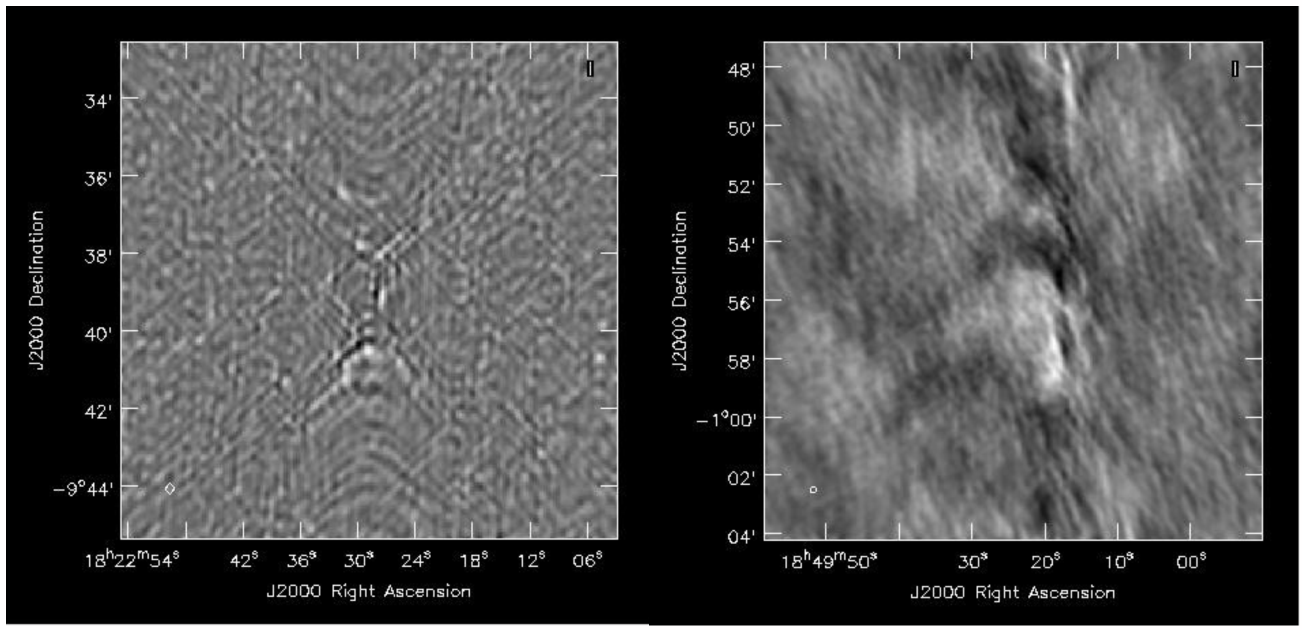
Plots of the amplitudes versus uvdist for the target and two of the calibration sources before calibration.

Uncalibrated Data



Plots of the phases versus uvdist for the target and two of the calibration sources before calibration.

Uncalibrated Images

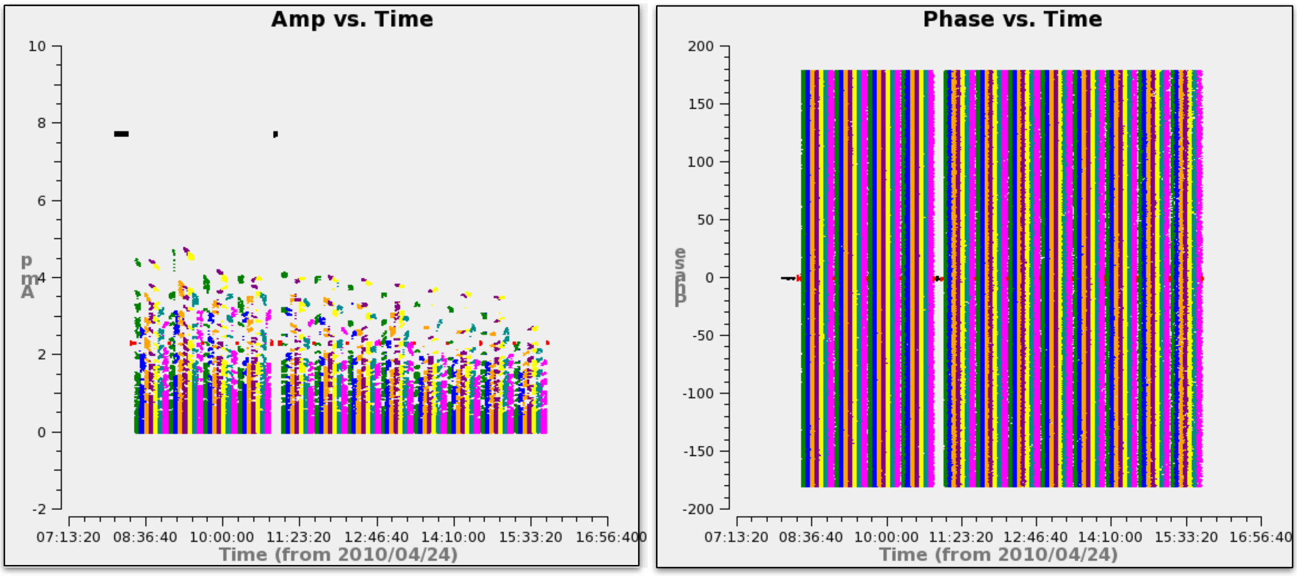


J1822-0938
(calibrator)

3C391
(science target)

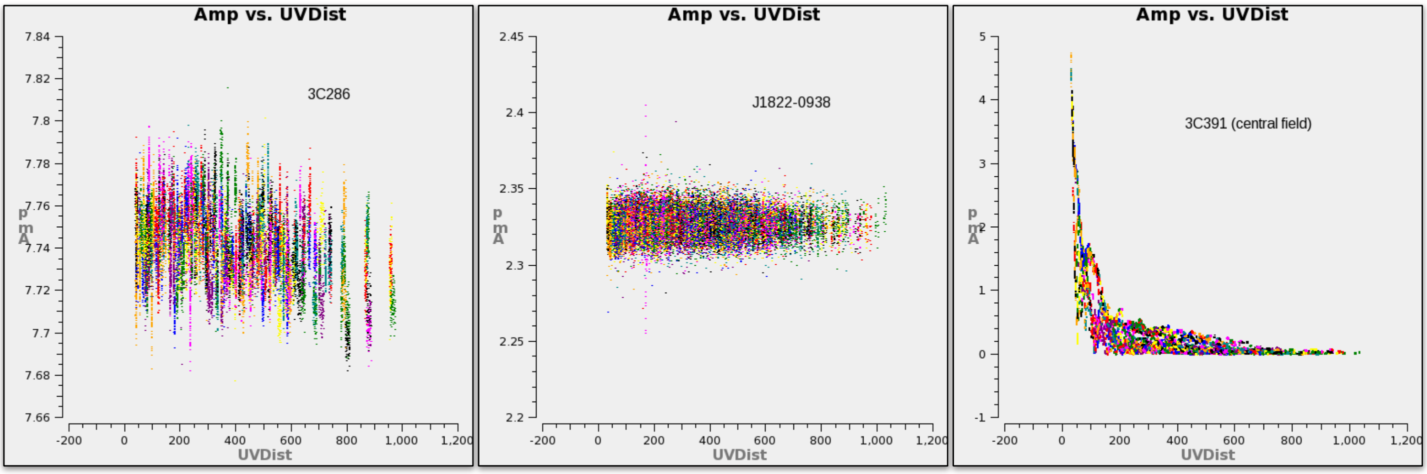
Images made from the uncalibrated visibilities for the target and one of the calibration sources. Neither source is really visible or recognizable. What should the image of the calibrator source look like?

Calibrated Data



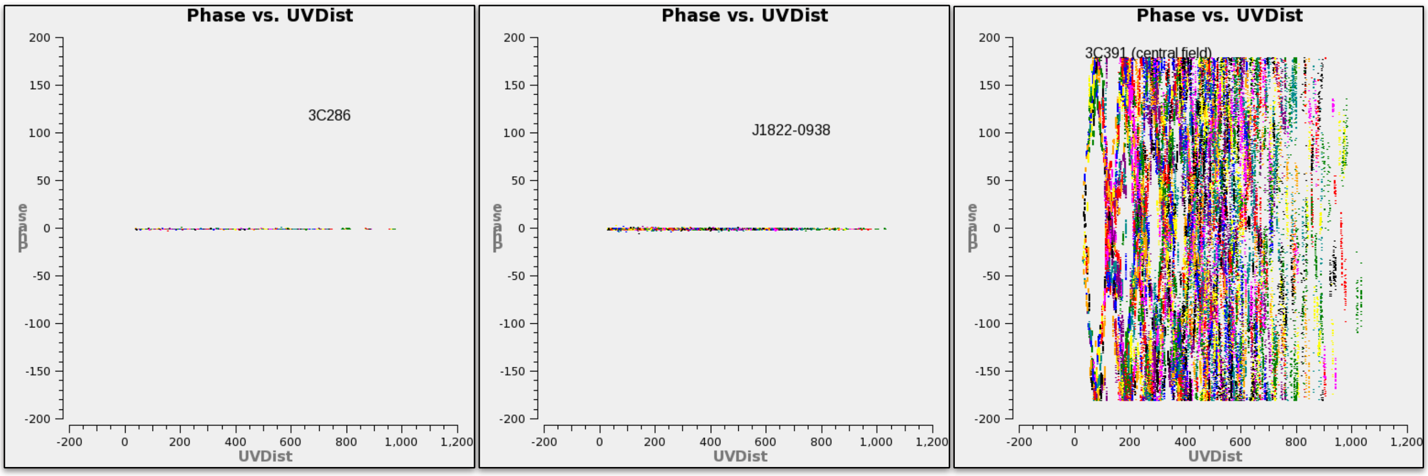
Amplitude and phase for the VLA dataset for 3C391 after calibration.
Looks similar in many ways, what's different?
First notice the scale of the amplitudes has changed dramatically.
Also notice the calibrator source flux densities are all now the same, as they should be.

Calibrated Data

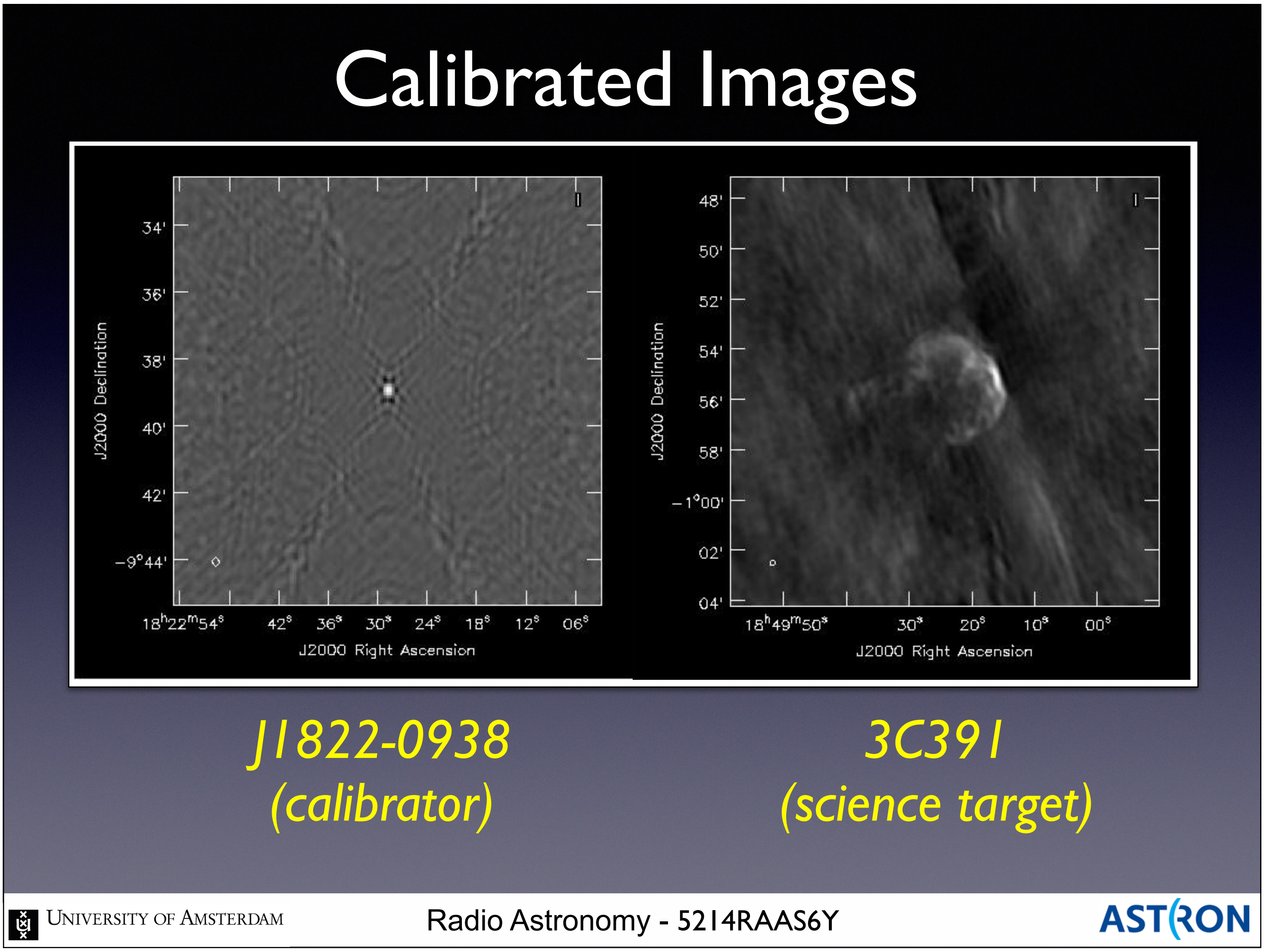


Plots of the amplitudes versus uvdist for the target and two of the calibration sources after calibration. Notice that the calibrator amplitudes are more or less constant (to a few %).

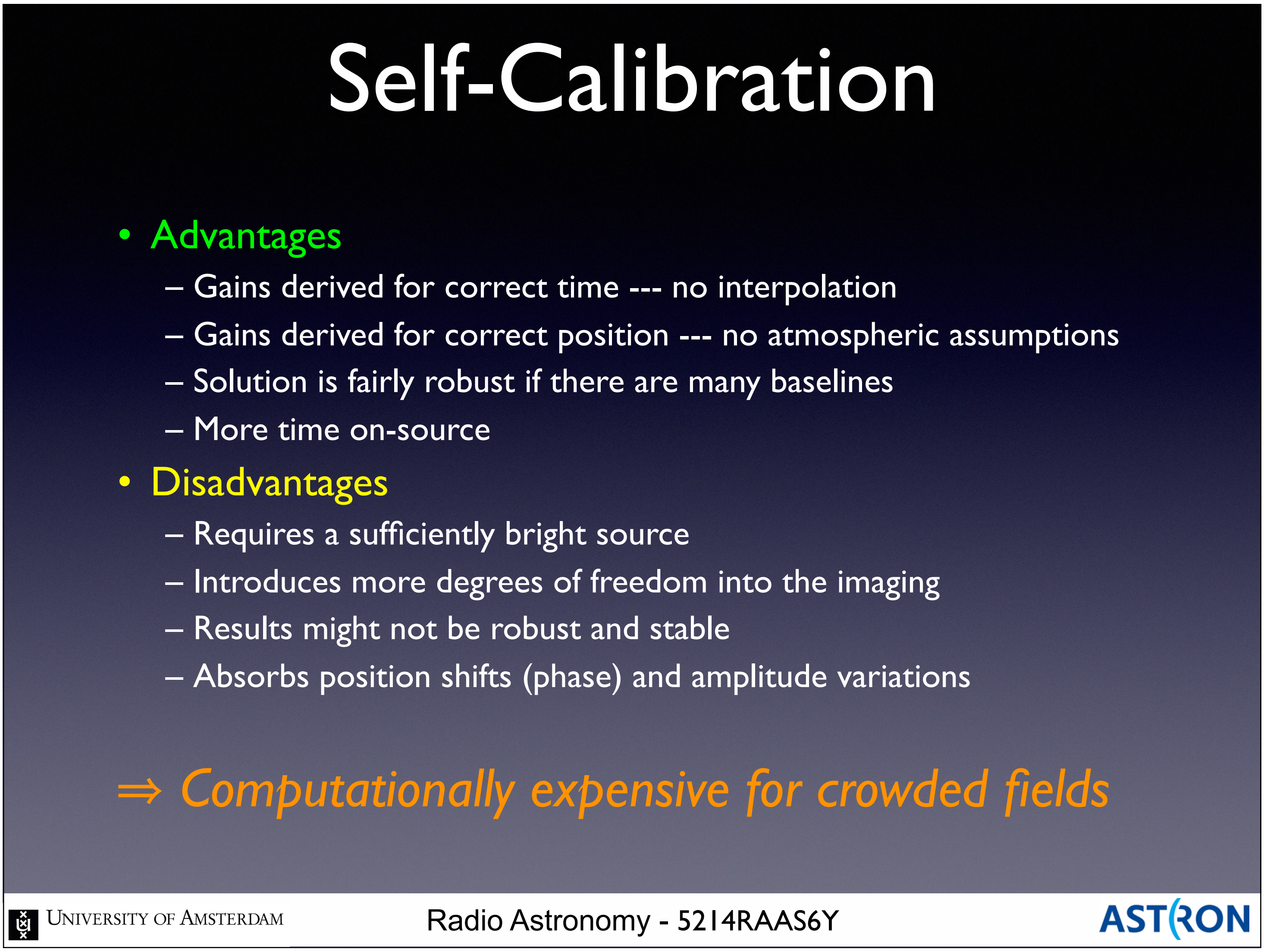
Calibrated Data



Plots of the phases versus uvdist for the target and two of the calibration sources after calibration. Notice that the calibrator amplitudes have constant, zero phases (as they should if they are point sources).



Images made from the calibrated visibilities for the target and one of the calibration sources.
Notice the calibrator now looks like a point source… as it should.
The target source looks like a supernova remnant, as it should.



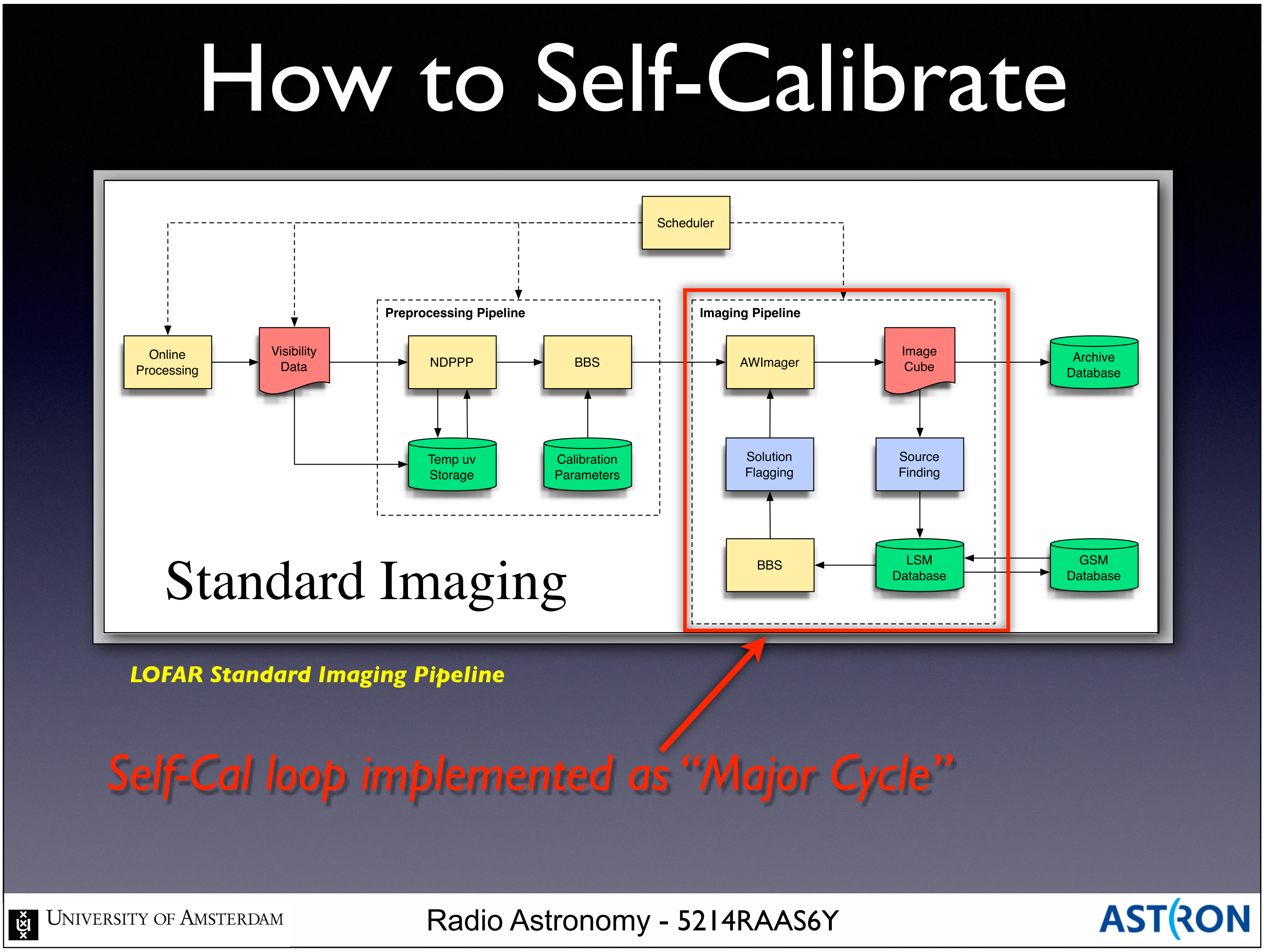
If your data is of sufficient quality, you can often “self-calibrate” the data.
Works if you have good data and more constraints than free parameters.
Doesn’t require external calibrators.
Does require (potentially) several iterations to get a good solution.
Can be computationally expensive.

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*

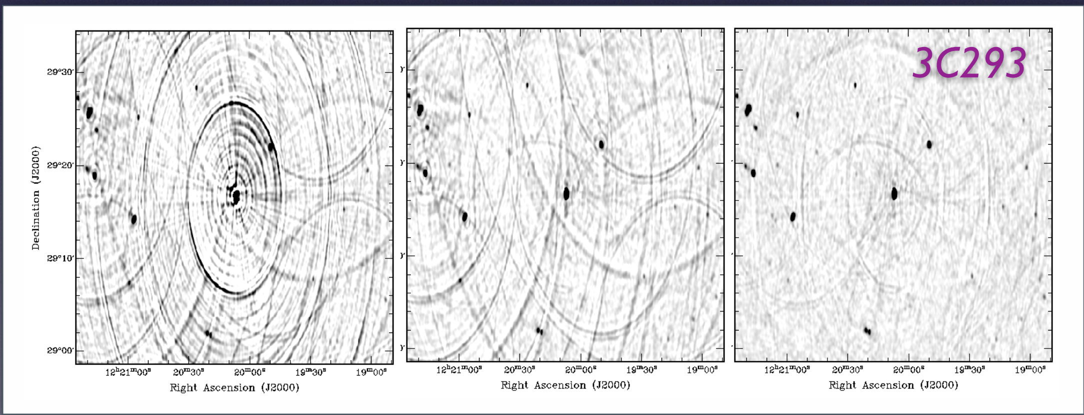
The source model becomes just another set of free parameters for which to solve.



The LOFAR imaging pipeline incorporates and automated, iterative self-cal loop as part of the standard processing.

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



A few guidelines about when and when not to try self-calibration.

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

A few guidelines about when and when not to try self-calibration.

Sensitivity limit

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

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

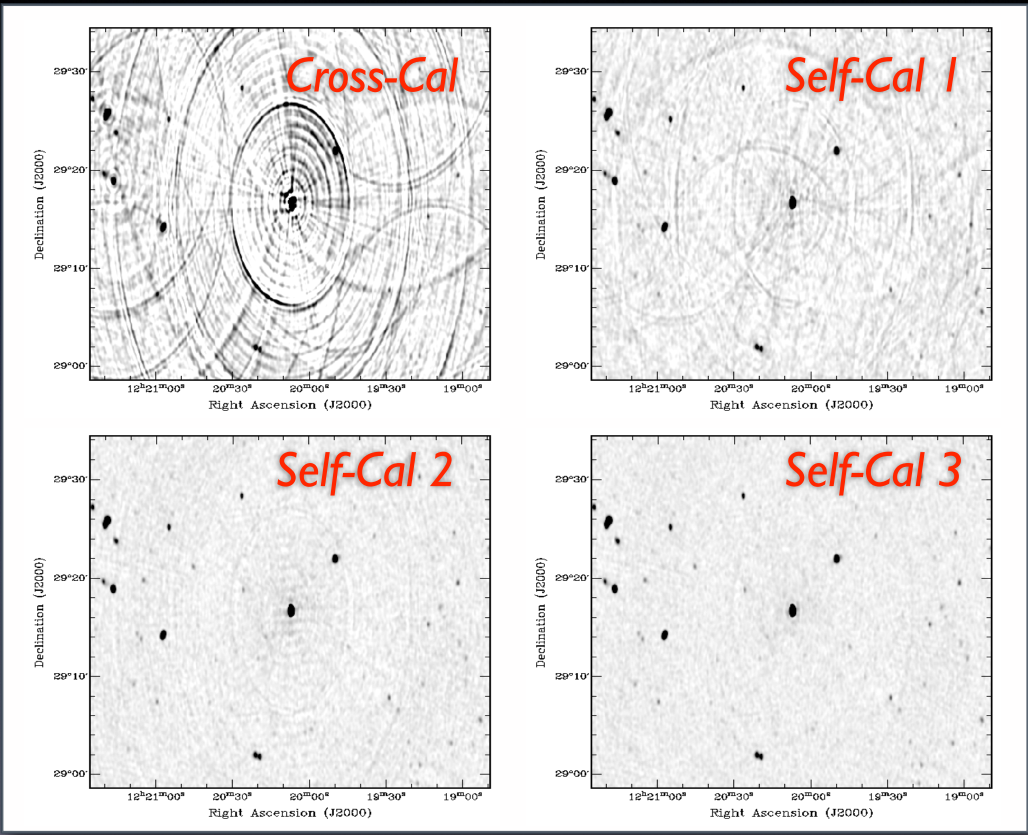
Amplitude and phase $$\sigma_g = \frac{1}{\sqrt{N-3}} \frac{\sigma_V}{S}$$

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

You can use the following formula to estimate whether your data have sufficient signal-to-noise to allow a self-calibration to work.

Self-cal Example: 3C293



An example of the improvement self-calibration can produce with enough iterations.
In general, self-cal procedures iterate until the process converges.
Convergence is usually measured as when the noise in the image does not improve.

Self-cal Example: Cygnus A

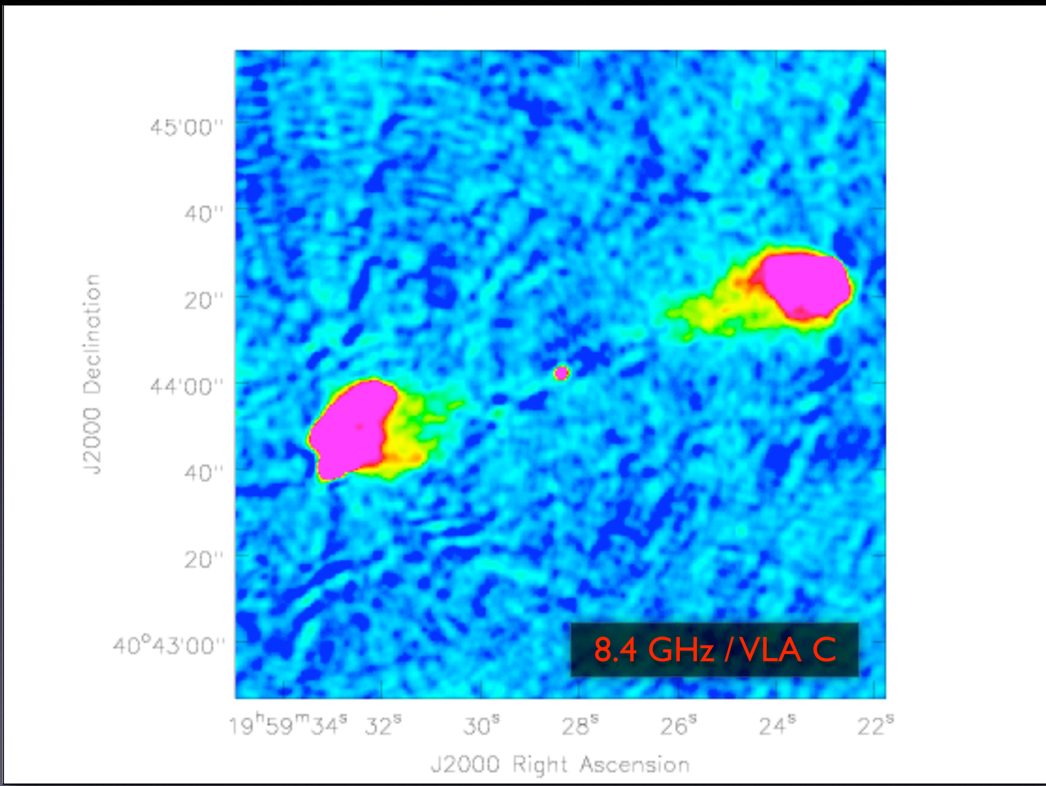
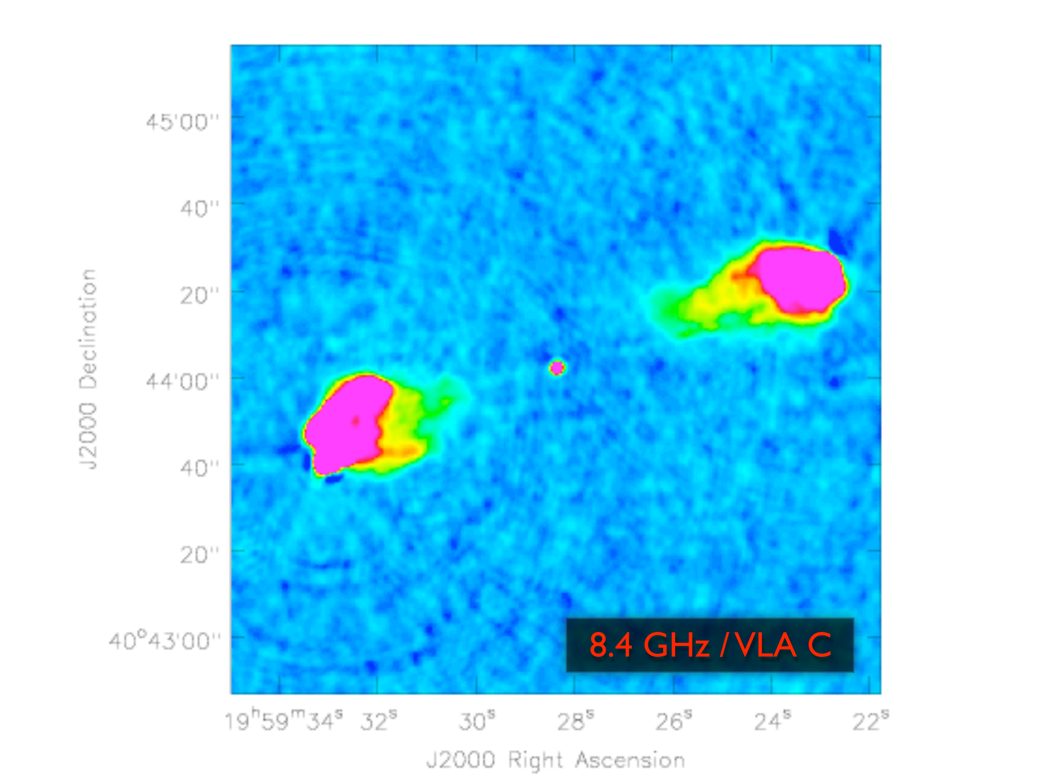


Image without self-calibration

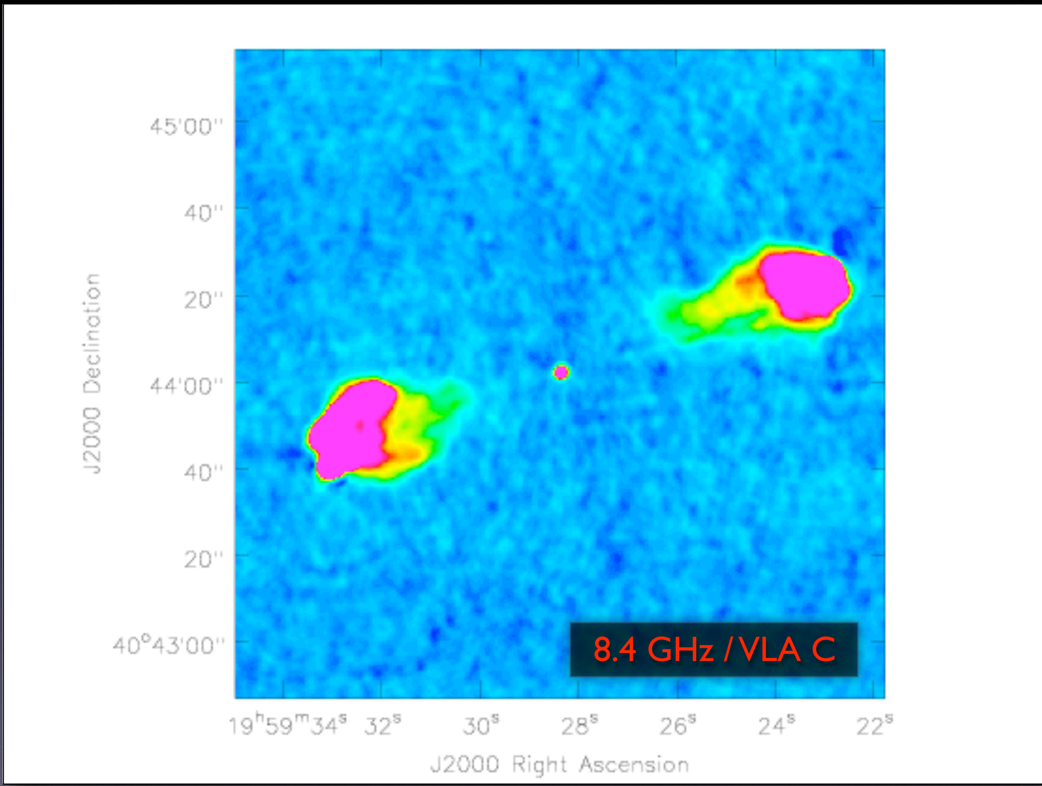
An example of the improvement that self-cal can produce.
This image shows the resulting image using normal cross-calibration.

Self-cal Example: Cygnus A



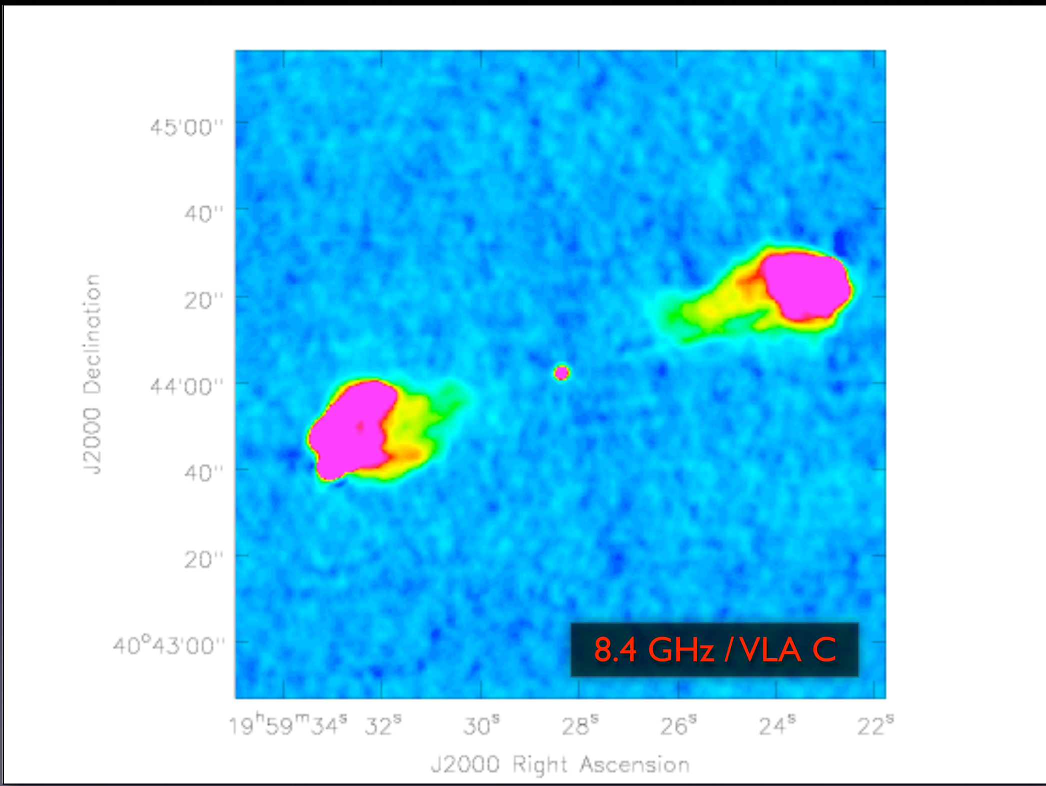
After 1 phase-only self-calibration

Self-cal Example: Cygnus A



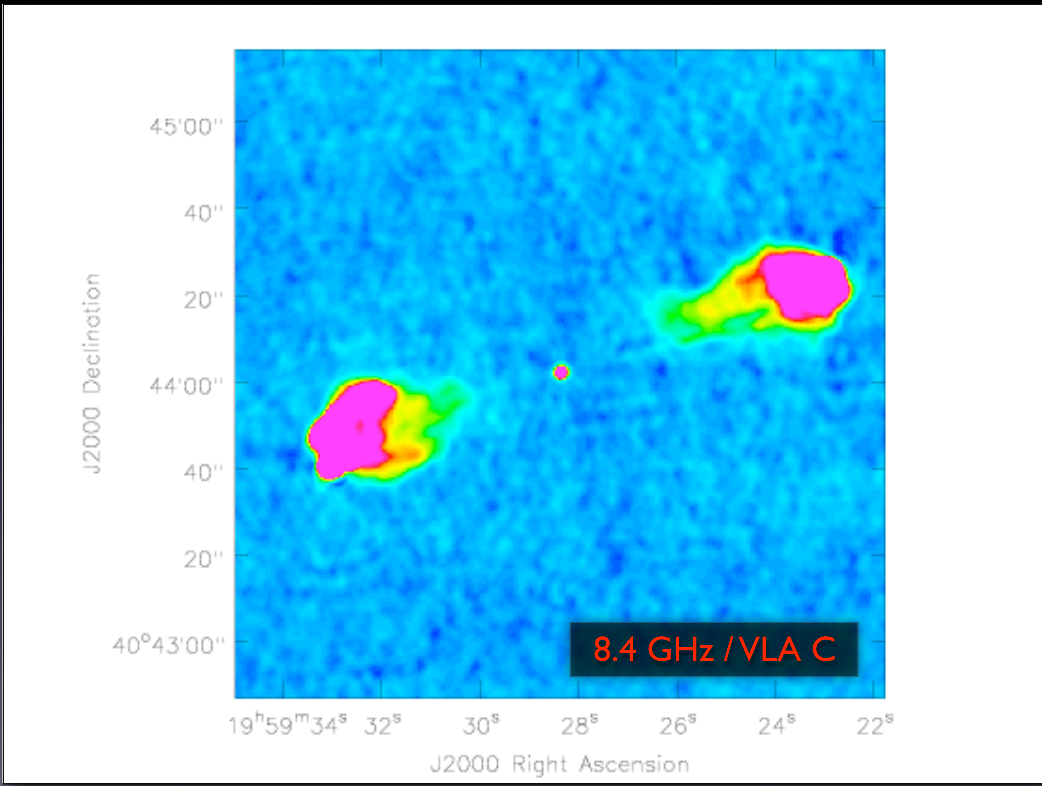
After 1 amplitude and phase calibrations

Self-cal Example: Cygnus A



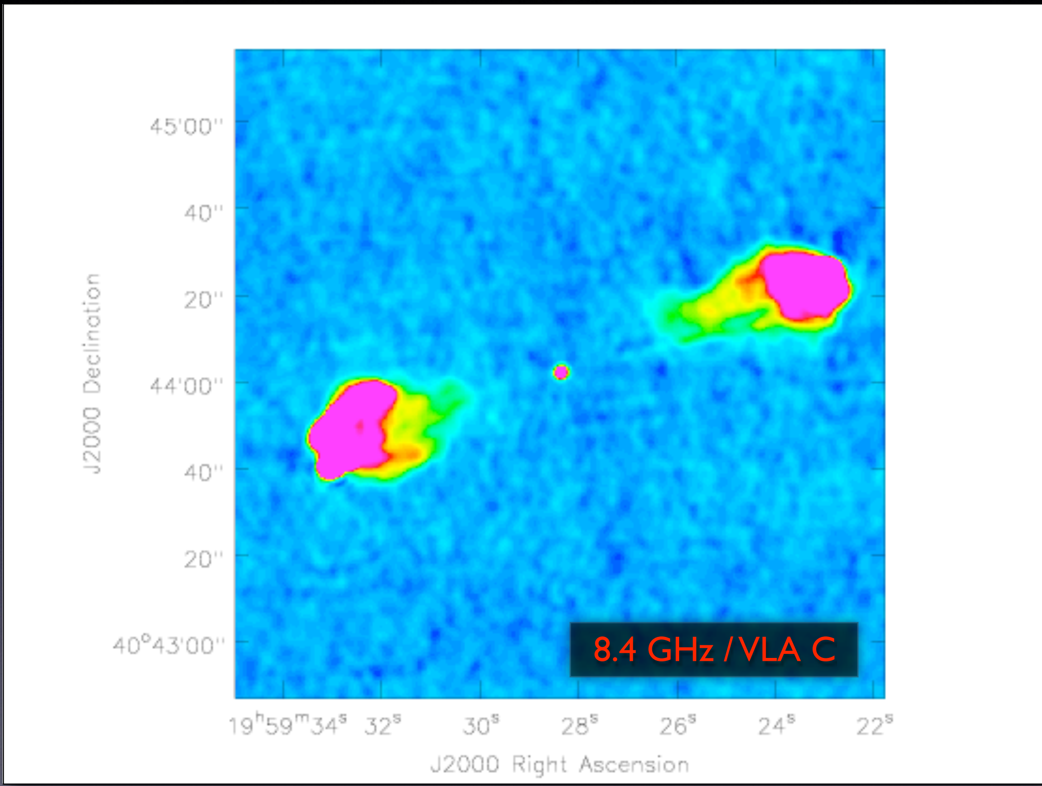
After 2 amplitude and phase calibrations

Self-cal Example: Cygnus A



After 3 amplitude and phase calibrations

Self-cal Example: Cygnus A

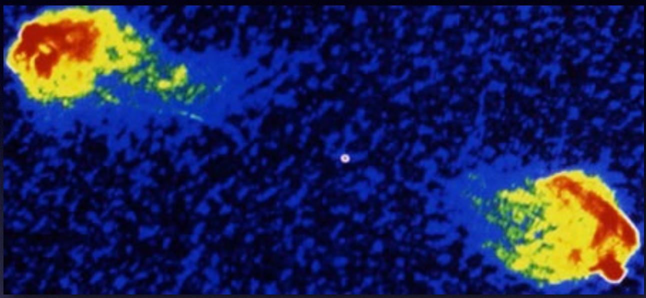


After 4 amplitude and phase calibrations

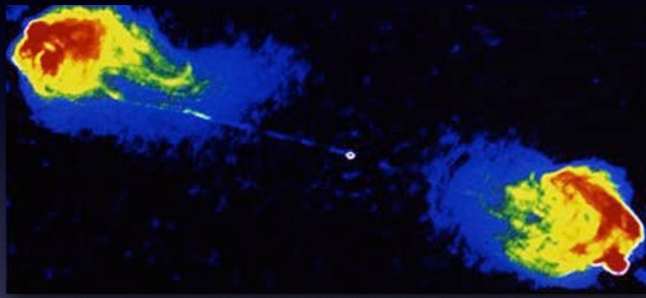
After 4 iterations, the noise properties of the image are basically no longer changing so self-cal stops.

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

After 4 iterations, the noise properties of the image are basically no longer changing so self-cal stops.

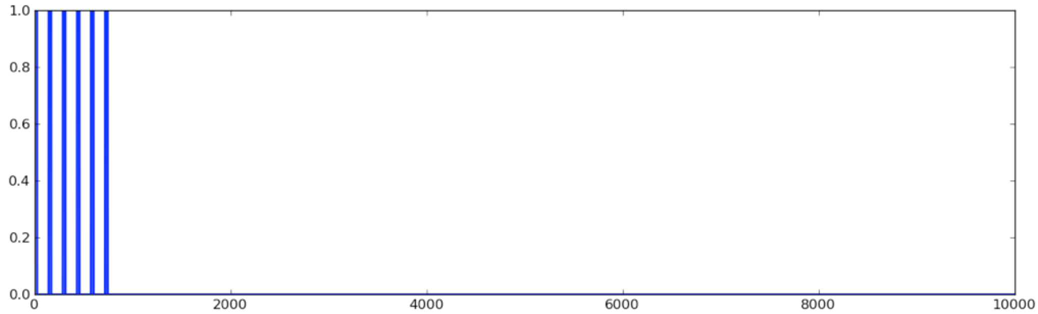
When Self-cal Fails

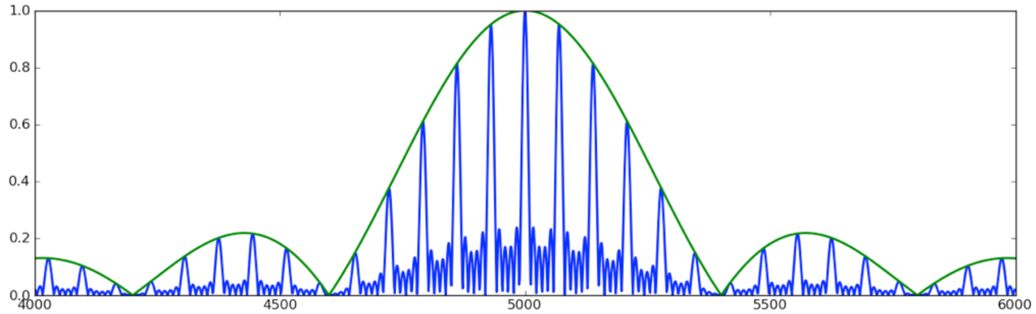
- Astrometry
- Variable sources
- Incorrect model
  - **barely-resolved sources**
  - **self-cal can embed mistakes in the data**
- Bad data
- Images dominated by deconvolution errors
  - **poor boxing**
  - **insufficient uv-coverage**
- Not enough flux density
  - **fast-changing atmosphere**
- Errors 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

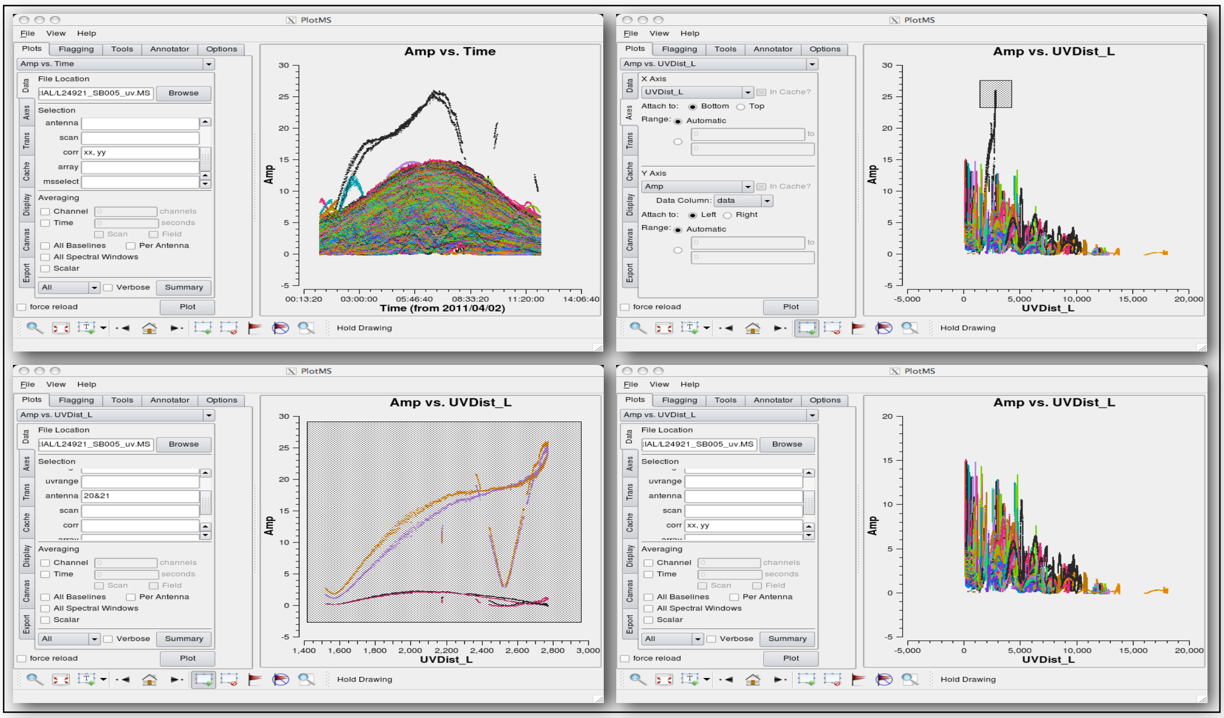
Today's Practicum





- Continue work on your simulated interferometer
- Discuss your proposal ideas

Next Practicum



- Examine, calibrate, and image an actual radio data set

Questions?