

When, why and how to self-cal



Drew Brisbin (NRAO)

Amanda Kepley, Crystal Brogan (NRAO)

Atacama Large Millimeter/submillimeter Array
Expanded Very Large Array
Robert C. Byrd Green Bank Telescope
Very Long Baseline Array



Self-Calibration: Motivation

JVLA and ALMA have impressive sensitivity! But what you achieve is often limited by residual calibration errors

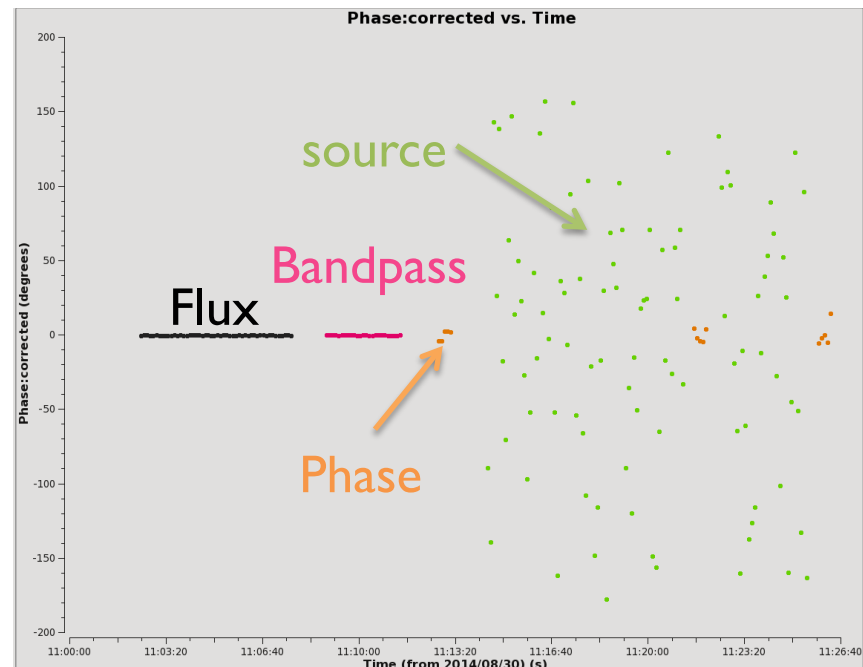
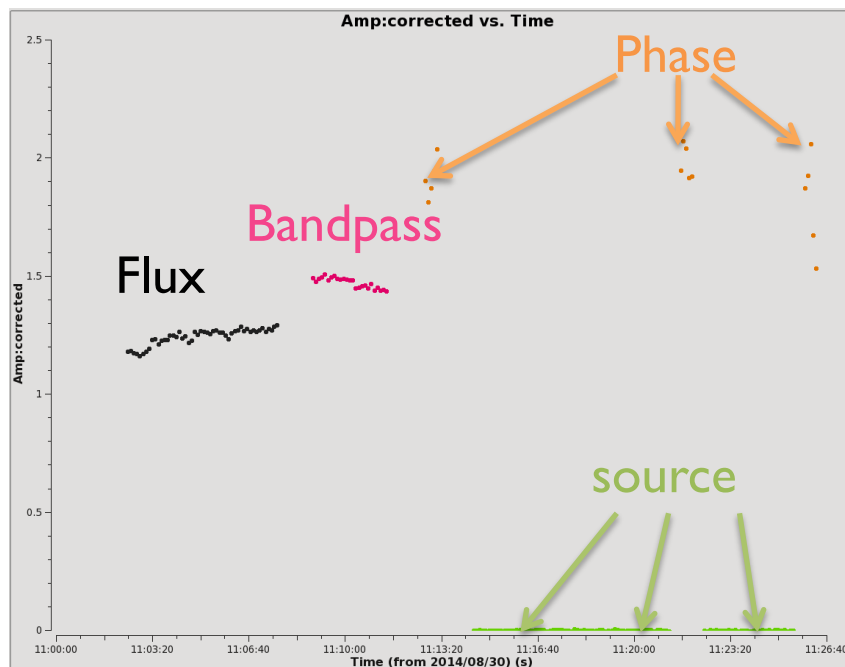
Many objects will have enough Signal-to-Noise (S/N) so they can be used to better calibrate **themselves** to obtain a more accurate image. This is called self-calibration and it really works, if you are careful! Sometimes, the increase in effective sensitivity may be an order of magnitude.

It is not a circular trick to produce the image that you want. It works because the number of baselines is much larger than the number of antennas so that an approximate source image does not stop you from determining a better temporal gain calibration which leads to a better source image.

Self-cal may not be included in the data pipelines. SO, YOU SHOULD LEARN HOW TO DO IT.

A quick review of calibration

- Interferometers measure visibilities, i.e., the amplitude and phase of the cross-correlated signals between pairs of antennas.
- We calibrate these data by determining the complex gains (amplitude and phase) and the frequency response (bandpass) for each antenna.

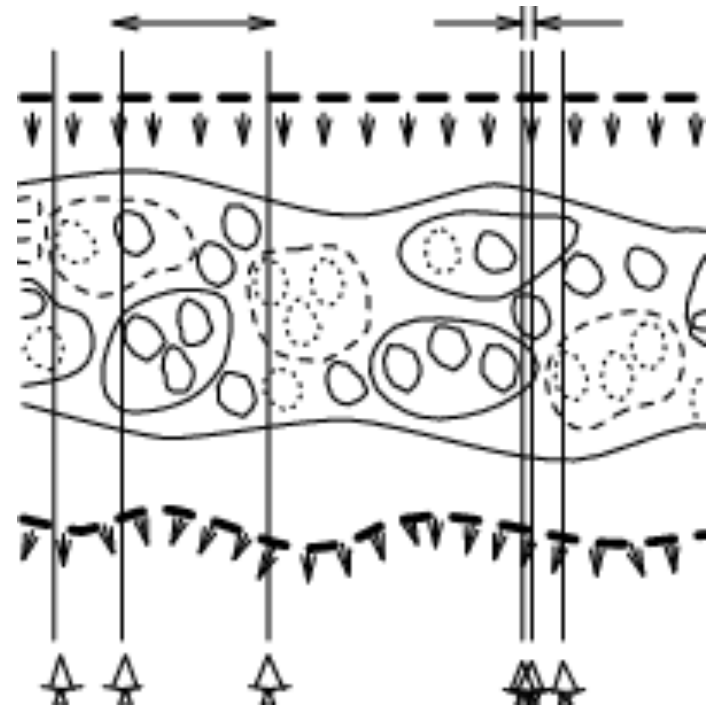


Data Corruption Types

The true visibility is corrupted by many effects:

- Atmospheric attenuation
- Radio “seeing”
- Variable pointing offsets
- Variable delay offsets
- Electronic gain changes
- Electronic delay changes
- Electronic phase changes
- Radiometer noise
- Correlator malfunctions
- Most Interference signals

| Antenna-based
| baseline



Antenna-based Calibration- I

- The most important corruptions are associated with antennas
- Basic Calibration Equation

$$\tilde{V}_{ij}(t) = g_i(t)g_j^*(t)G_{ij}(t)V_{ij}(t) + \varepsilon_{ij}(t) + \epsilon_{ij}(t)$$

- $g_i(t)g_j^*(t)$ Factorable (antenna-based) complex gains
- $G_{ij}(t)$ Non-factorable complex gains (not Antenna based and hopefully small)
- $V_{ij}(t)$ True Visibility
- $\varepsilon_{ij}(t) + \epsilon_{ij}(t)$ Additive offset (not antenna based) and thermal noise, respectively

- Can typically be reduced to approximately

$$\tilde{V}_{ij}(t) = g_i(t)g_j^*(t)V_{ij}(t) + \epsilon_{ij}(t)$$

Antenna-based Calibration-II

- For N antennas, [(N-1)*N]/2 visibilities are measured, but only N amplitude and (N-1) phase gains fully describe the complete Antenna-based calibration. This redundancy is used for antenna gain calibration
- Basic gain (phase and amplitude) calibration involves observing unresolved (point like) “calibrators” of known position with visibility $M_{i,j}(t_k, \nu)$
- Determine gain corrections, g_i , that minimizes S_k for each time stamp t_k where

$$S_k = \sum_k \sum_{i,j}^{i \neq j} w_{i,j} |g_i(t_k) g_j^*(t_k) V_{i,j}^o(t_k) - M_{i,j}(t_k)|^2$$

Data
Weights

Complex
Gains

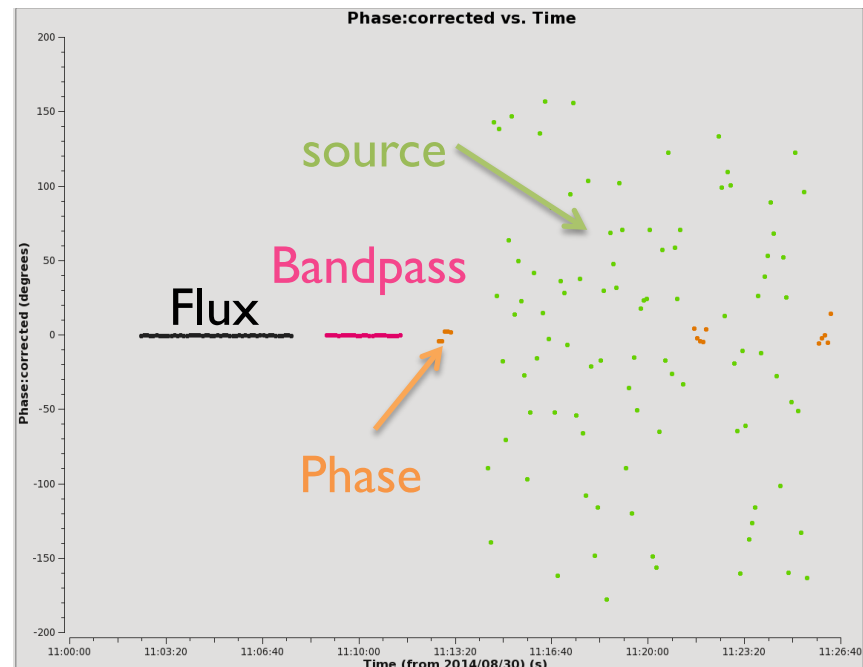
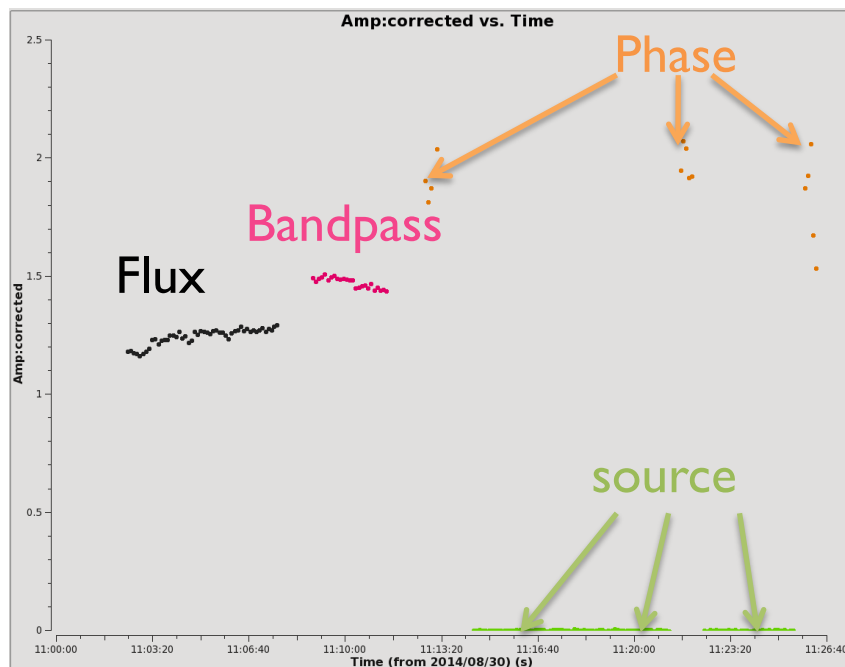
Complex
Visibilities

Fourier transform
of model image

- The solution interval, t_k , is the data averaging time used to obtain the values of g_i , (i.e. solint='int' or 'inf'). The a priori weight of each data point is $w_{i,j}$.
- This IS a form of Self-calibration, only we assume a Model (M_{ij}) that has constant amplitude and zero phase, i.e. a point source
- Transferring these solutions to your science target will be imperfect, but the redundant baselines can also be used with a **model image** for self-calibration of the antenna complex gains.

Self-calibration in a Nutshell

- Self-calibration uses **your source** to better calibrate the antenna-based complex gains (i.e., amplitudes and phases) as a function of time.
- There are two types of self-calibration: phase and amplitude+phase.



Sensitivities for Self-Calibration-I

- **For phase only self-cal:** Need to detect the target with a $S/N > 3$ in a solution time (**solint**) less than the time for significant phase variations for all baselines to a **single antenna**. For 25 antennas, $S/N > 3$ will lead to < 15 deg error.

- Make an initial image, cleaning it conservatively
 - Measure rms in emission free region
 - $\text{rms}_{\text{Ant}} = \text{rms} \times \sqrt{N-3}$ where N is # of antennas
 - $\text{rms}_{\text{self}} = \text{rms}_{\text{Ant}} \times \sqrt{\text{total time}/\text{solint}}$
 - Measure Peak flux density = Signal
 - If $S/N_{\text{self}} = \text{Peak}/\text{rms}_{\text{self}} > 3$ try phase only self-cal

Rule of thumb:

For an array with ~25 antennas, if S/N in image > 20 its worth trying phase-only self-cal

- **CAVEAT 1:** If dominated by extended emission, estimate what the flux will be on the longer baselines (by plotting the uv-data) instead of the image
 - If the majority of the baselines in the array cannot "see" the majority of emission in the target field (i.e. emission is resolved out) at a S/N of about 3, the self-cal will fail in extreme cases (though bootstrapping from short to longer baselines is possible, it can be tricky).
- **CAVEAT 2:** If severely dynamic range limited (poor uv-coverage), it can also be helpful to estimate the rms noise from uv-plots

Sensitivities for Self-Calibration-II

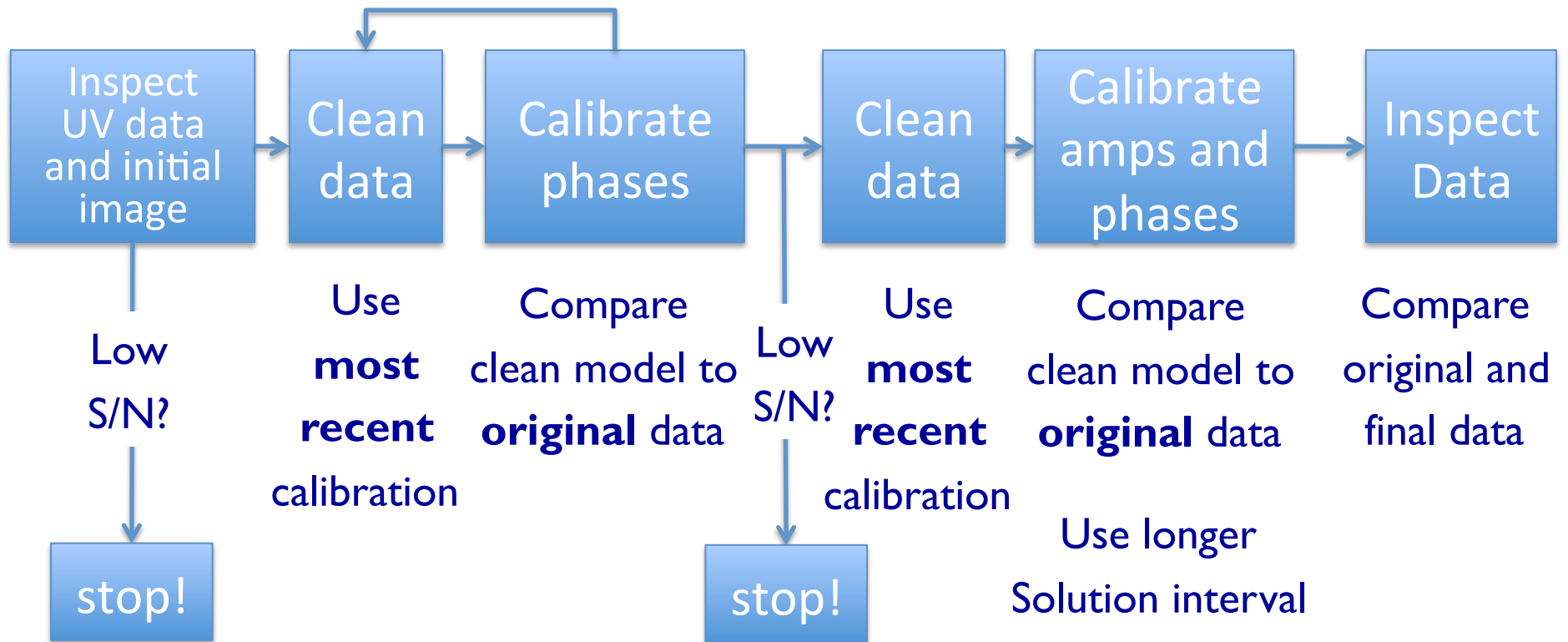
- **For amplitude self-cal:** Need to detect the target with a $S/N > 10$ with only the baselines to a single antenna in a solution time (solint) less than the time for significant amplitude variations. For 25 antennas, an antenna based $S/N > 10$ will lead to a 10% amplitude error.
 - Amplitude corrections are more subject to deficiencies in the model image. Check results carefully!
 - For example, if clean model is missing significant flux compared to uv-data, give uvrange for amplitude solution that excludes short baselines.

Additional S/N for self-cal can be obtained by:

- Increase solint (solution interval)
 - If errors that are directional rather than time dependent, self-calibration solutions can yield surprising improvement even for solints that span the entire observation. Antenna position (aka baseline) errors are a good example.
- gaintype= 'T' to average polarizations
 - Caveat: Only if your source is unpolarized
- Combine = 'spw' to average spw's (assumes prior removal of spw to spw offsets)
 - Caveat: If source spectral index/morphology changes significantly across the band, do not combine spws, especially for amplitude self-cal
- Combine = 'fields' to average fields in a mosaic (use with caution, only fields with strong signal)

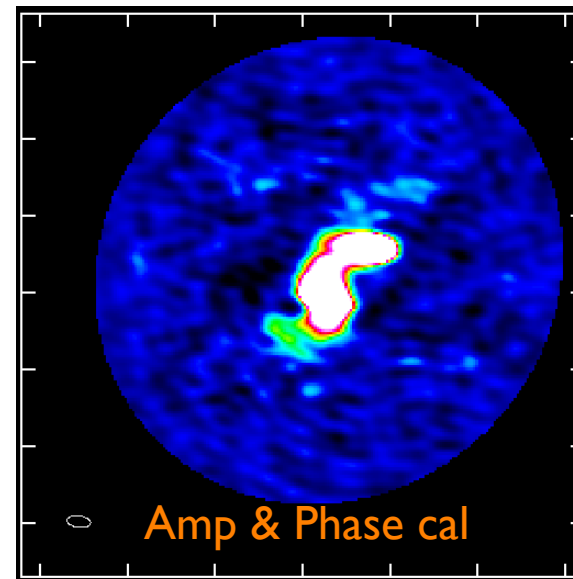
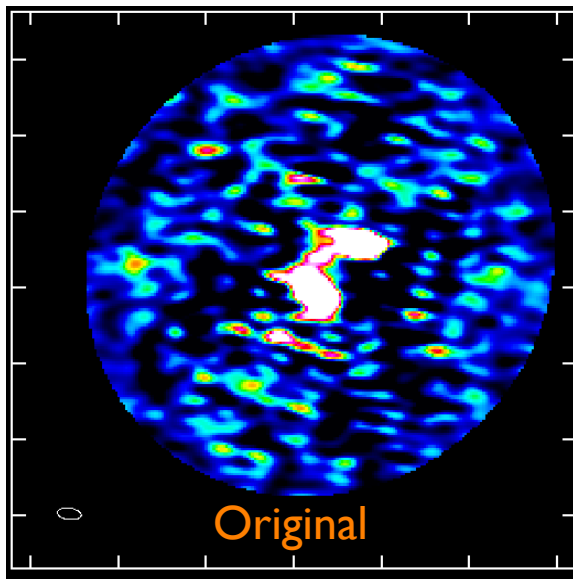
Outline of Self-Calibration Process

Repeat with deeper cleans and shorter solution intervals until phases no longer improve



Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6



See [IRAS16293 CASA guide](#) for detailed commands.

http://casaguides.nrao.edu/index.php?title=IRAS16293_Band9_-_Imaging_for_CASA_4.3

Self-calibration Example:

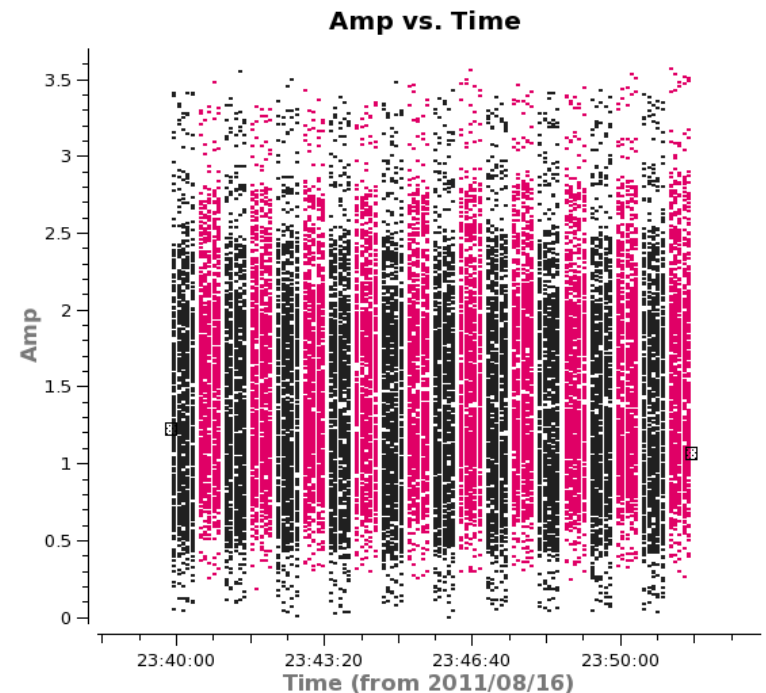
ALMA SV Data for IRAS16293 Band 6

Step 1 – Determine basic setup of data:

- 2 pointing mosaic
- Integration = 6.048 sec; subscans ~ 30sec
- Scan= 11 min 30s (split between two fields)

Step 2 – What is the expected rms noise?

- Use actual final total time and # of antennas on science target(s) from this stage and sensitivity calculator.
- Be sure to include the actual average weather conditions for the observations in question and the bandwidth you plan to make the image from
- 54 min per field with 16 antennas and average $T_{\text{sys}} \sim 80$ K, 9.67 MHz BW; rms= 1 mJy/beam
- Inner part of mosaic will be about 1.6 x better due to overlap of mosaic pointings



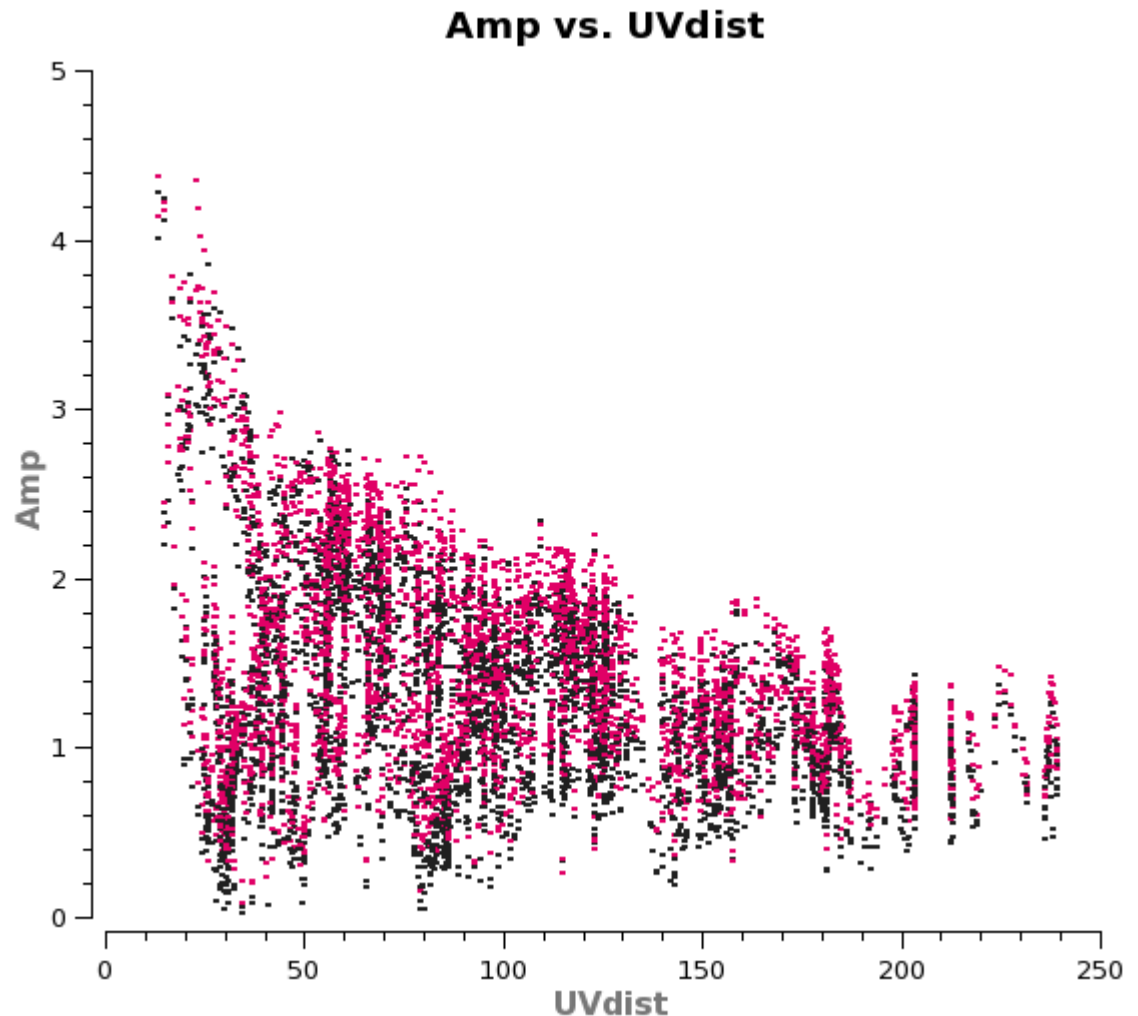
- ALMA mosaic: alternates fields in “subscan” this picture = 1 scan
- EVLA mosaic: alternates fields in scans
- Subscans are transparent to CASA (and AIPS)

Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

Step 3 – What does the amplitude vs uv-distance of your source look like?

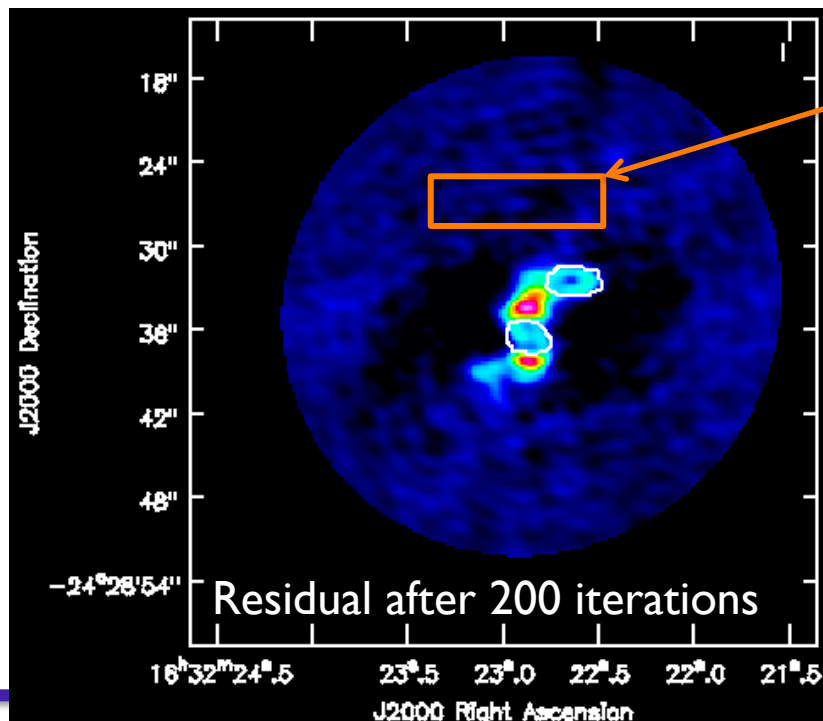
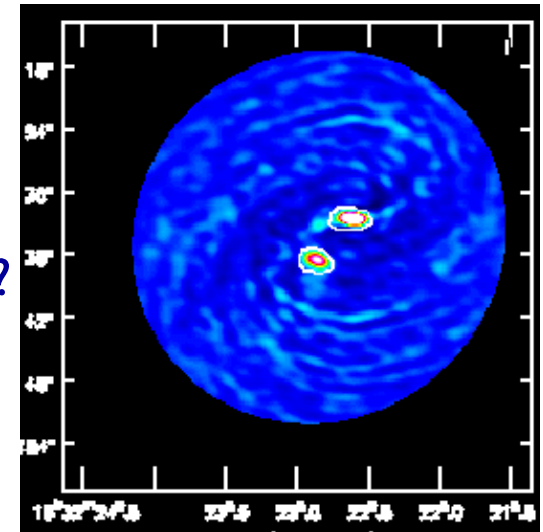
- Does it have large scale structure?
- Here we have increasing flux on short baselines so we need to pay extra attention to how well we are modeling the extended emission.
- What is the flux density on short baselines?
- Keep this 4 Jy peak in mind while cleaning. What is the total cleaned flux you are achieving?



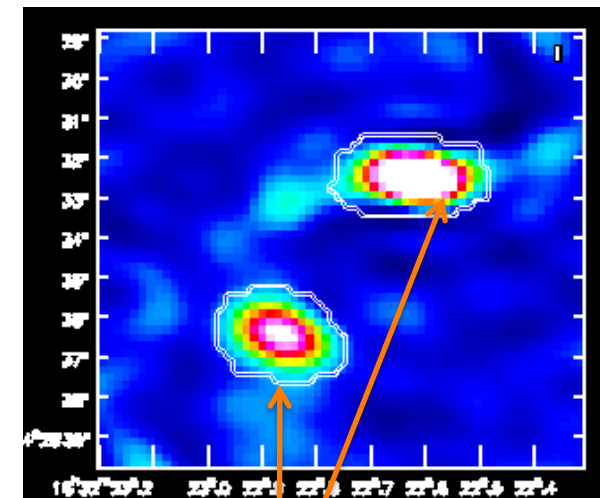
Self-calibration Example: ALMA SV Data for IRAS 16293 Band 6

Step 4 – What is the S/N in a conservatively cleaned image?

- What is this “conservative” of which you speak?
- $\text{Rms} \sim 15 \text{ mJy/beam}$; $\text{Peak} \sim 1 \text{ Jy/beam} \rightarrow \text{S/N} \sim 67$
- $\text{Rms} > \text{expected}$ and $\text{S/N} > 20 \rightarrow \text{self-cal!}$



Stop clean,
and get rms
and peak
from image,
avoiding
negative
bowls and
emission



Clean boxes only around
emission you are SURE are
real at this stage

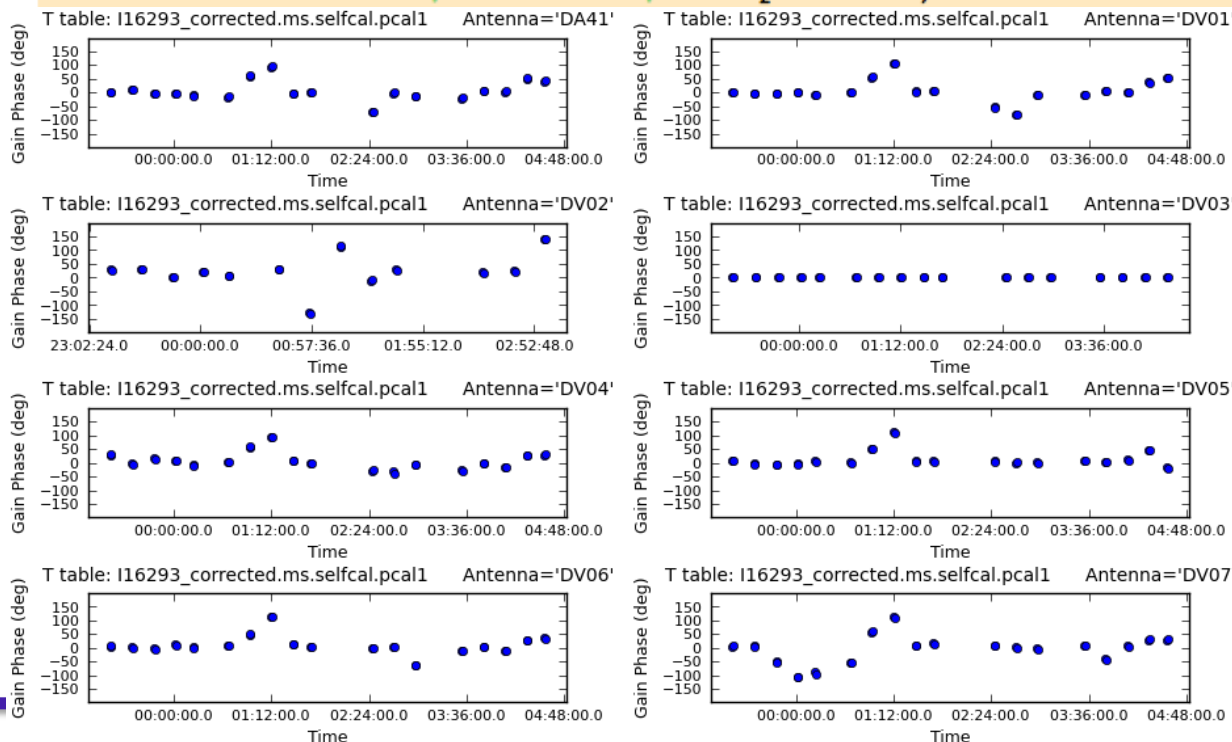
Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

Step 5: Decide on a time interval for initial phase-only self-cal

- A good choice is often the scan length (in this case about 5 minutes per field)
 - Exercise for reader: show using equations on slide 8 that $S/N_{\text{self}} \sim 5.4$
- In CASA you can just set solint='inf' (i.e. infinity) and as long as combine \neq 'scan' AND \neq 'field' you will get one solution per scan, per field.

```
gaincal(vis='IRAS16293_Band9.fixed.CONT.msAVG', caltable='cont_pcal1', gaintype='T',
        refant='DV14', calmode='p', combine='', spw='', field='',
        solint='inf', minsnr=3.0, minblperant=4)
```



What to look for:

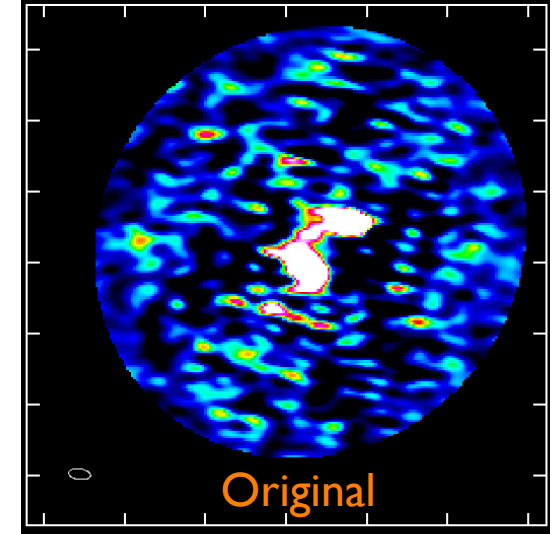
- Lot of failed solutions on most antennas? if so, go back and try to increase S/N of solution = more averaging of some kind
- Do the phases appear smoothly varying with time (as opposed to noise like)

Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

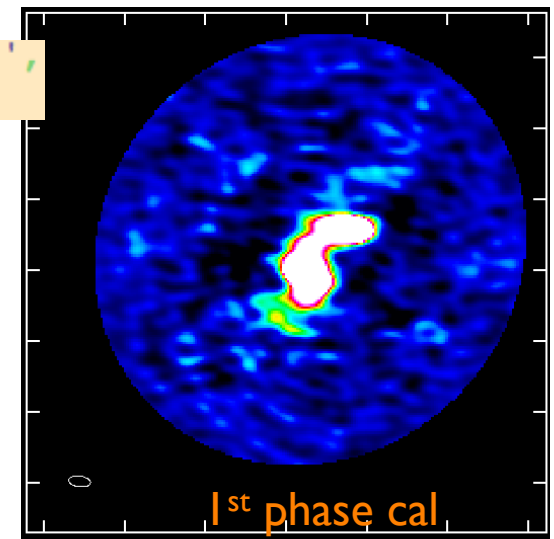
Step 6: Apply solutions and re-clean

- Incorporate more emission into clean box if it looks real
- Stop when residuals become noise-like but still be a bit conservative, ESPECIALLY for weak features that you are very interested in
 - You **cannot** get rid of real emission by not boxing it
 - You can create features by boxing noise



```
applycal(vis='IRAS16293_Band9.fixed.CONT.msAVG', spwmap=[], spw='',  
         gaintable=['cont_pcal1'], calwt=F, flagbackup=F)
```

- Original:
Rms~ 15 mJy/beam; Peak ~ 1 Jy/beam → S/N ~ 67
- 1st phase-only:
Rms~ 6 mJy/beam; Peak ~ 1.25 Jy/beam → S/N ~ 208
- Did it improve? If, yes, continue. If no, something has gone wrong or you need a shorter solint to make a difference, go back to Step 4 or stop.

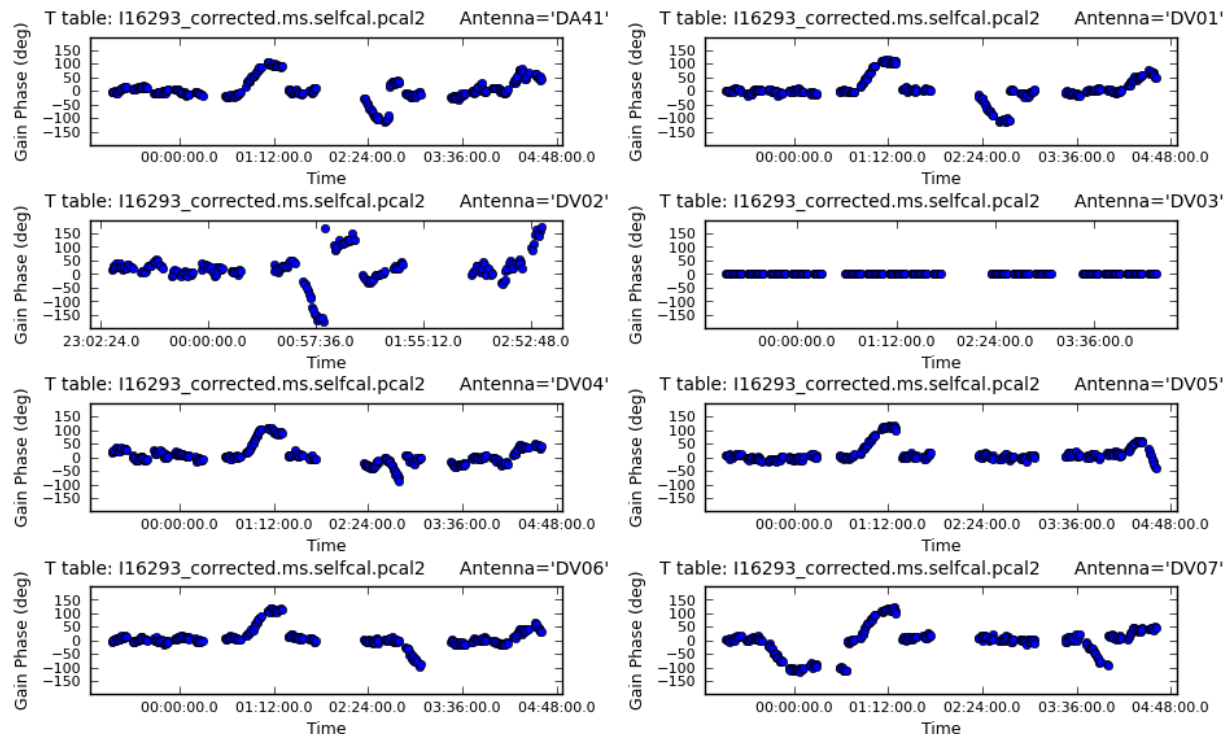


Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

Step 8: Try shorter solint for 2nd phase-only self-cal

- In this case we'll try the subscan length of 30sec
- It is best NOT to apply the 1st self-cal while solving for the 2nd. i.e. incremental tables can be easier to interpret but you can also “build in” errors in first model by doing this



What to look for:

- Still smoothly varying?
- If this looks noisy, go back and stick with longer solint solution
- If this improves things a lot, could try going to even shorter solint

Self-calibration Example:

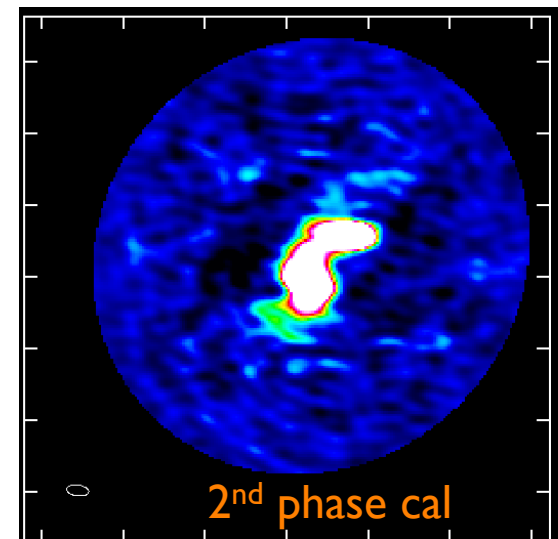
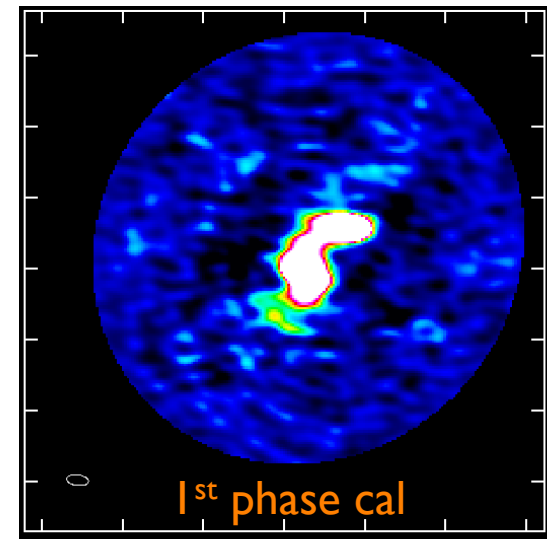
ALMA SV Data for IRAS 16293 Band 6

Step 9: Apply solutions and re-clean

- Incorporate more emission into clean box if it looks real
- Stop when residuals become noise-like but still be a bit conservative, ESPECIALLY for weak features that you are very interested in
 - You **cannot** get rid of real emission by not boxing it
 - You can create features by boxing noise

Step 10: Compare 1st and 2nd phase-only self-cal images

- 1st phase-only:
Rms~ 6 mJy/beam; Peak ~ 1.25 Jy/beam → S/N ~ 208
- 2nd phase-only:
Rms~ 5.6 mJy/beam; Peak ~ 1.30 Jy/beam → S/N ~ 228
- Did it improve? Not much, so going to shorter solint probably won't either, so we'll try an amplitude self-cal next

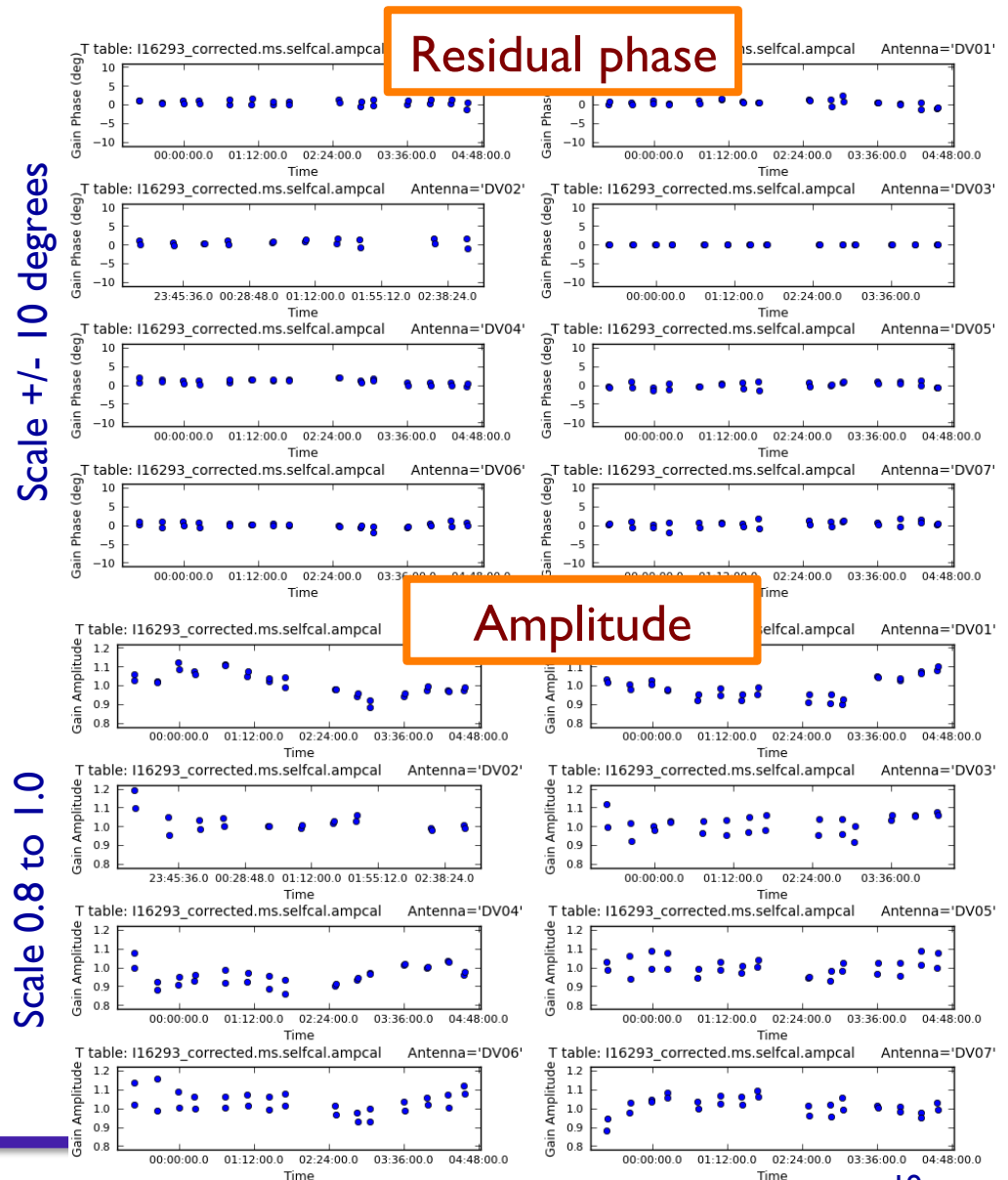


Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

Step 1: Try amplitude self-cal

- Amplitude tends to vary more slowly than phase. It's also less constrained, so solints are typically longer. Lets try two scans worth or 23 minutes
- Essential to apply the best phase only self-cal before solving for amplitude. Also a good idea to use mode='ap' rather than just 'a' to check that residual phase solutions are close to zero.
- Again make sure mostly good solutions, and a smoothly varying pattern.



Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

Step 12: Apply solutions

- Apply both 2nd phase and amp cal tables
- Inspect uv-plot of corrected data to
 - Check for any new outliers, if so flag and go back to Step 9.
 - Make sure model is good match to data.
 - Confirm that flux hasn't decreased significantly after applying solutions

Original

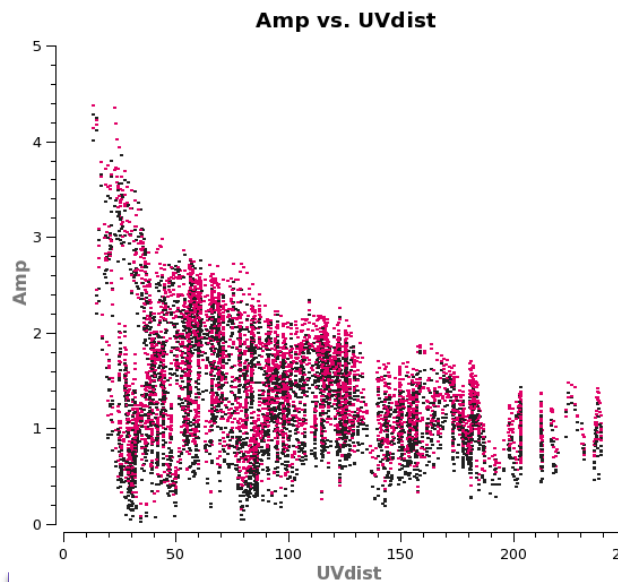
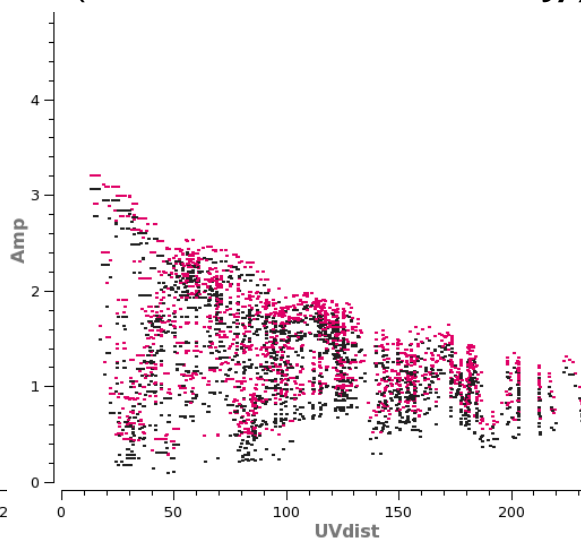
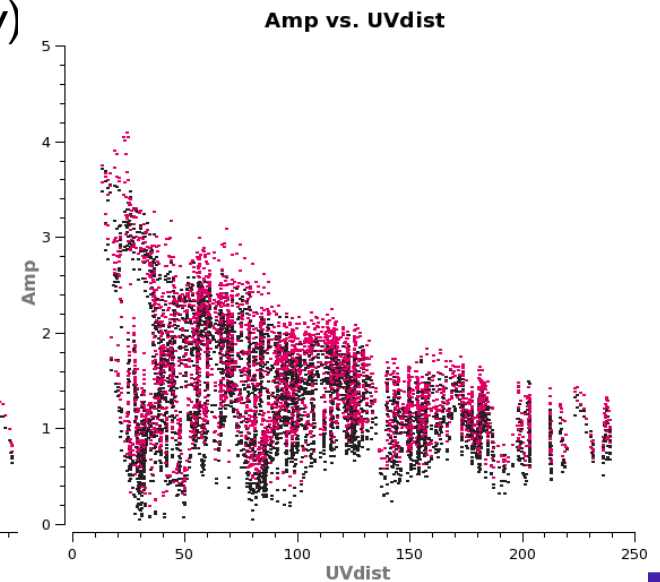


Image Model

(total cleaned flux = 3.4 Jy)



Amp & Phase applied



Self-calibration Example:

ALMA SV Data for IRAS 16293 Band 6

Step 13: Re-clean

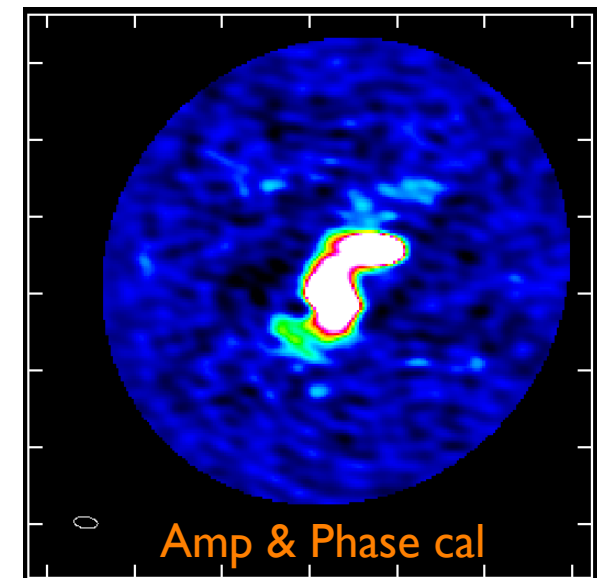
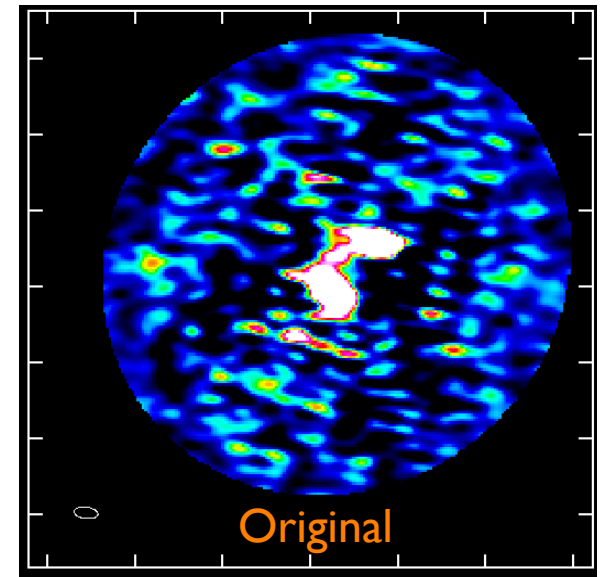
- Incorporate more emission into clean box
- Stop when residuals become noise-like – clean everything you think is real

Step 14: Compare 2nd phase-only and amp+phase self-cal images

- 2nd phase-only:
Rms~ 5.6 mJy/beam; Peak ~ 1.30 Jy/beam → S/N ~ 228
- Amp & Phase:
Rms~4.6 mJy/beam; Peak~1.30 Jy/beam → S/N ~283
- Did it improve? → Done!

Final: S/N=67 vs 283!

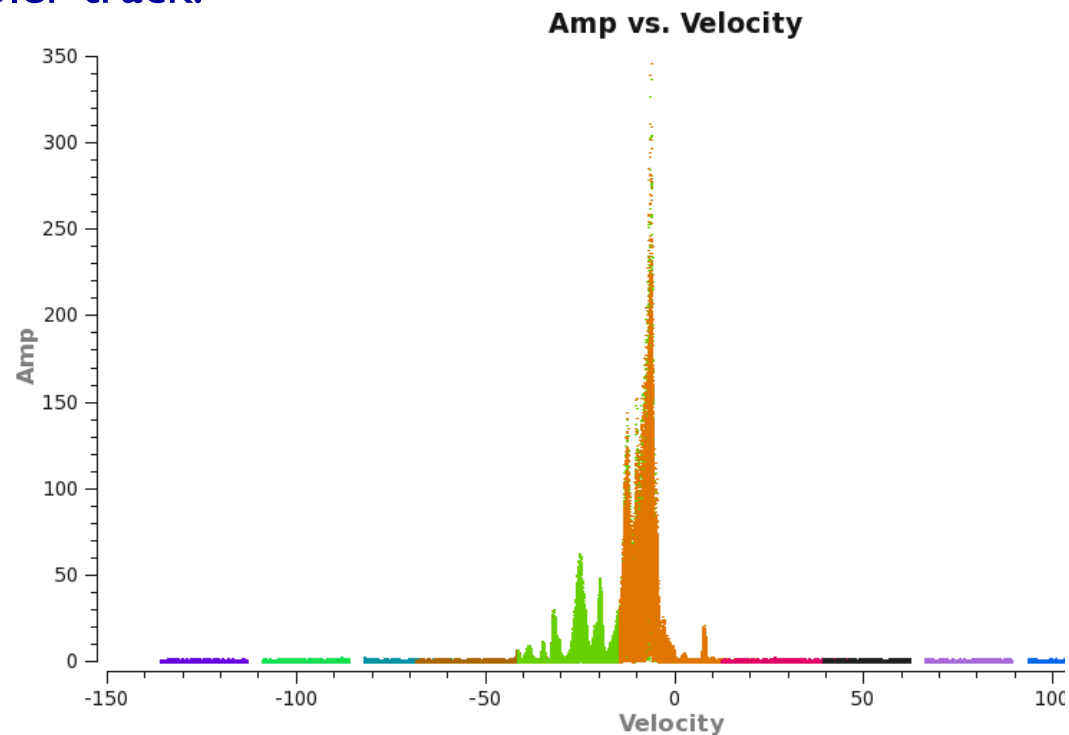
But not as good as theoretical
= dynamic range limit



You can also use bright line emission to self-calibrate!

Similar process to continuum self-calibration, only you want to use an image of the line channels with bright emission to create your model.

NOTE: In general unless you've already run CVEL, DATA CHANNEL NUMBER \neq IMAGE CHANNEL NUMBER due to Doppler Shift. In other words, ALMA does not Doppler track.



Self-Calibration Summary

- Self-calibration isn't magic. It works because we have over-constrained data.
- Try phase-only selfcal if S/N is $> \sim 20$ (for ~ 25 antennas).
- Clean conservatively! It's hard to get rid of real sources, but easy to add new ones.
- Always use your original data set to determine the new complex gains. However, you should always create new images from the data with the latest calibration applied.
- Amplitude self-calibration has more free parameters than phase calibration, so you need to exercise a bit more caution, especially if there's extended emission.
- Make sure to check your solutions and compare the original visibility values to the new ones.