Self-Calibration
Why, When, How

Jihyun Kang (KASI)

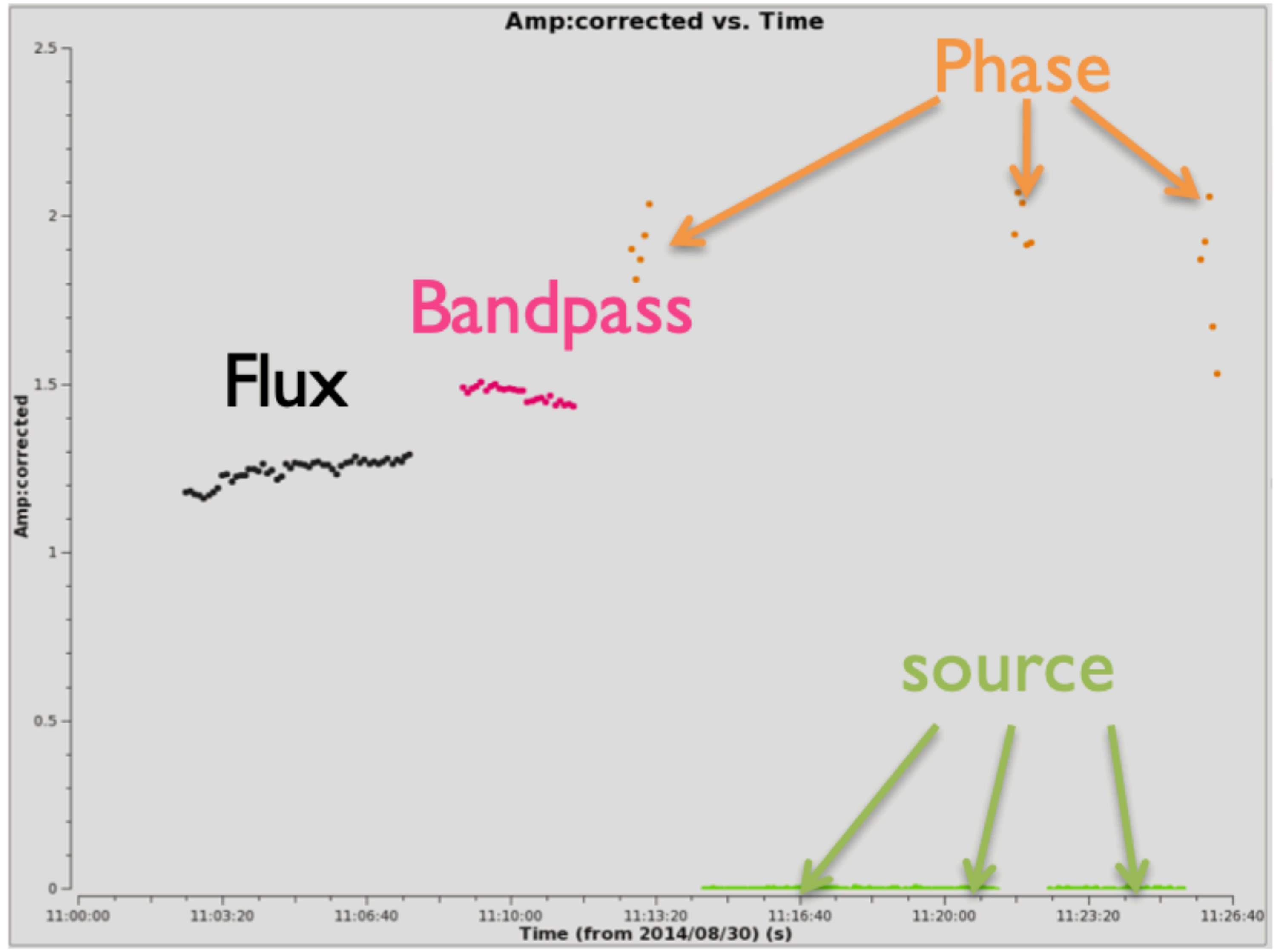
6th ALMA summer school (2022/Aug/23)

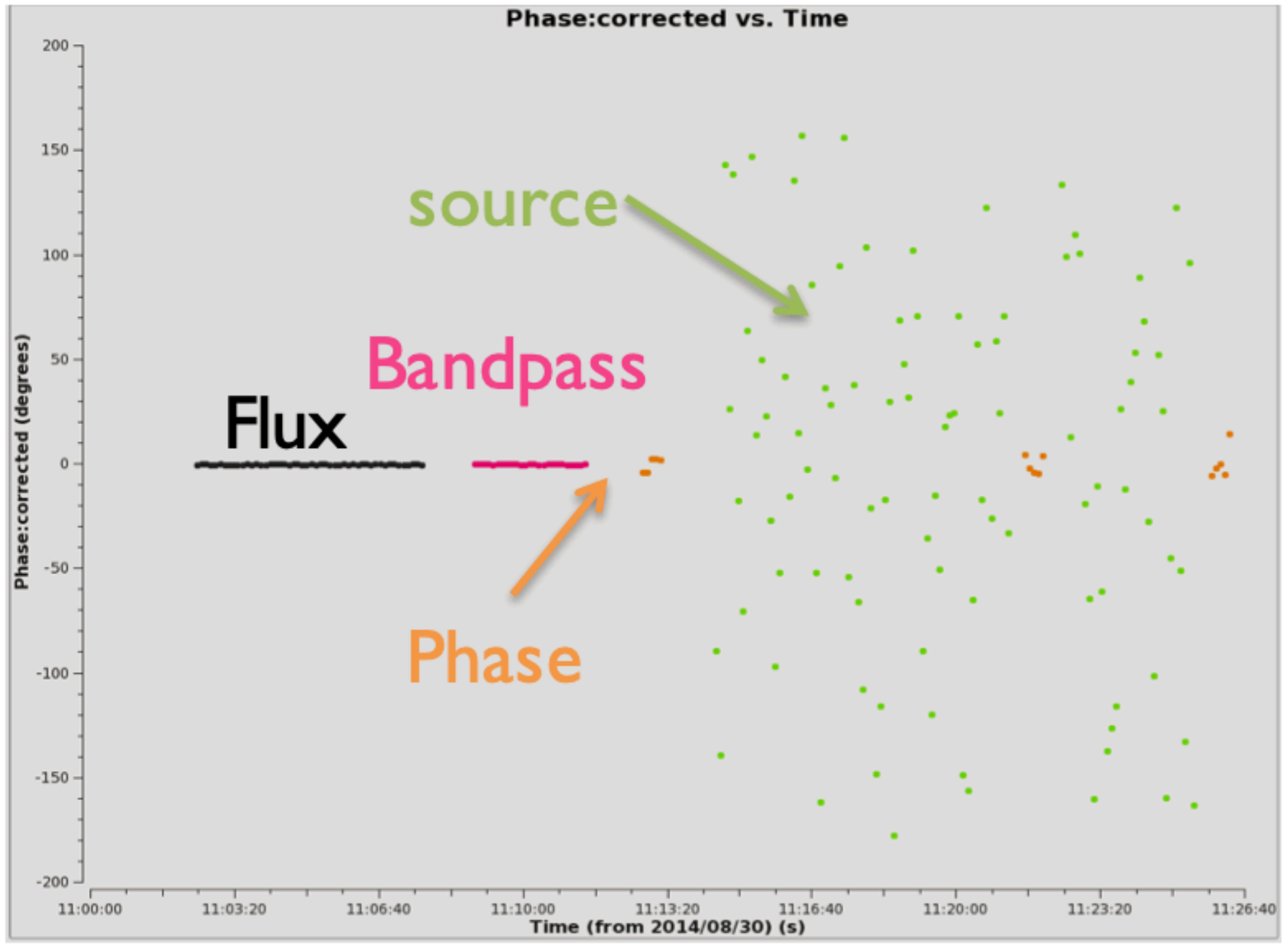
References

- https://science.nrao.edu/science/meetings/2018/16th-synthesis-imaging-workshop/16th-synthesis-imaging-workshop-lectures
- Advanced Gain Calibration Techniques in Radio Interferometry, 2018, Brogan et al.
- https://casaguides.nrao.edu/index.php?title=First_Look_at_Self_Calibration_CASA_6
- https://casaguides.nrao.edu/index.php?title=Image_Line
- https://casaguides.nrao.edu/index.php?title=IRAS16293_Band9_-_Imaging_for_CASA_6.2
- https://casaguides.nrao.edu/index.php?title=Self-Calibration_Template
- https://casadocs.readthedocs.io/en/stable/

A quick review of calibration

- Interferometers measure "visibilities": the amplitude and phase information 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.

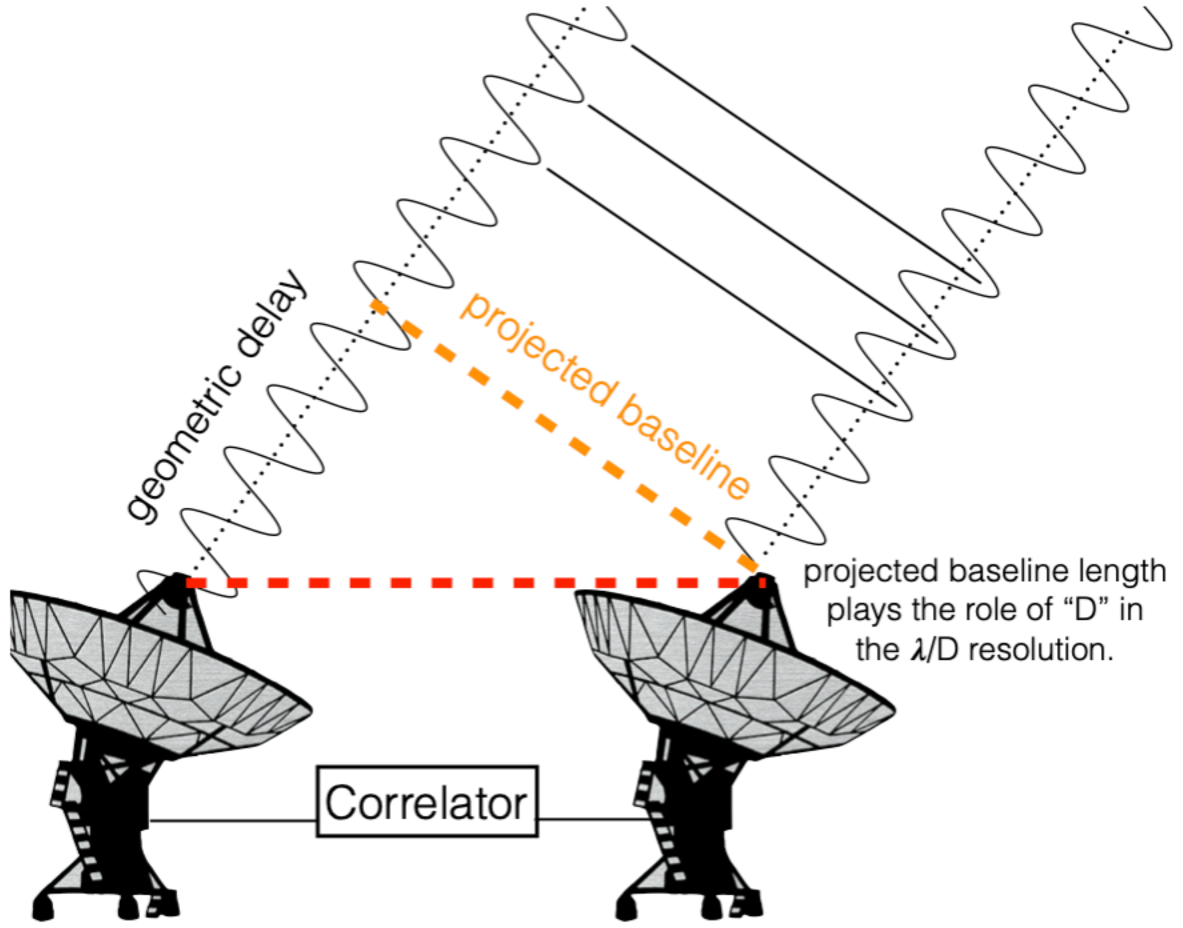




10.2.2 Calibration Simplification

$$\mathcal{V}_o^{i,j}(t_k,\nu_f) = \mathcal{V}_m^{i,j}(t_k,\nu_f) * G^{i,j}(t_k,\nu_f).$$

$$\begin{aligned} G^{i,j}(t_k,\nu_f) = & \\ A^i(t_k,\nu_0) * A^j(t_k,\nu_0) & \text{ Temporal amplitude} \\ +A^i(t_0,\nu_f) * A^j(t_0,\nu_f) & \text{ Bandpass amplitude} \\ \phi^i(t_k,\nu_0) - \phi^j(t_k,\nu_0) & \text{ Temporal phase} \\ +\phi^i(t_0,\nu_f) - \phi^j(t_0,\nu_f) & \text{ Bandpass phase} \\ +\Delta G^{i,j}(t_k,\nu_f) & \text{ Any additional correction} \end{aligned}$$



Self-Calibration: Motivation

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

To surpass this, many objects have enough Signal-to-Noise (S/N) that they can be used to calibrate ***themselves*** to obtain a better image. This is self-calibration.

- 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-calibration may not be included in the data pipelines

…so, it's best you learn how to do it

Data Corruption Types

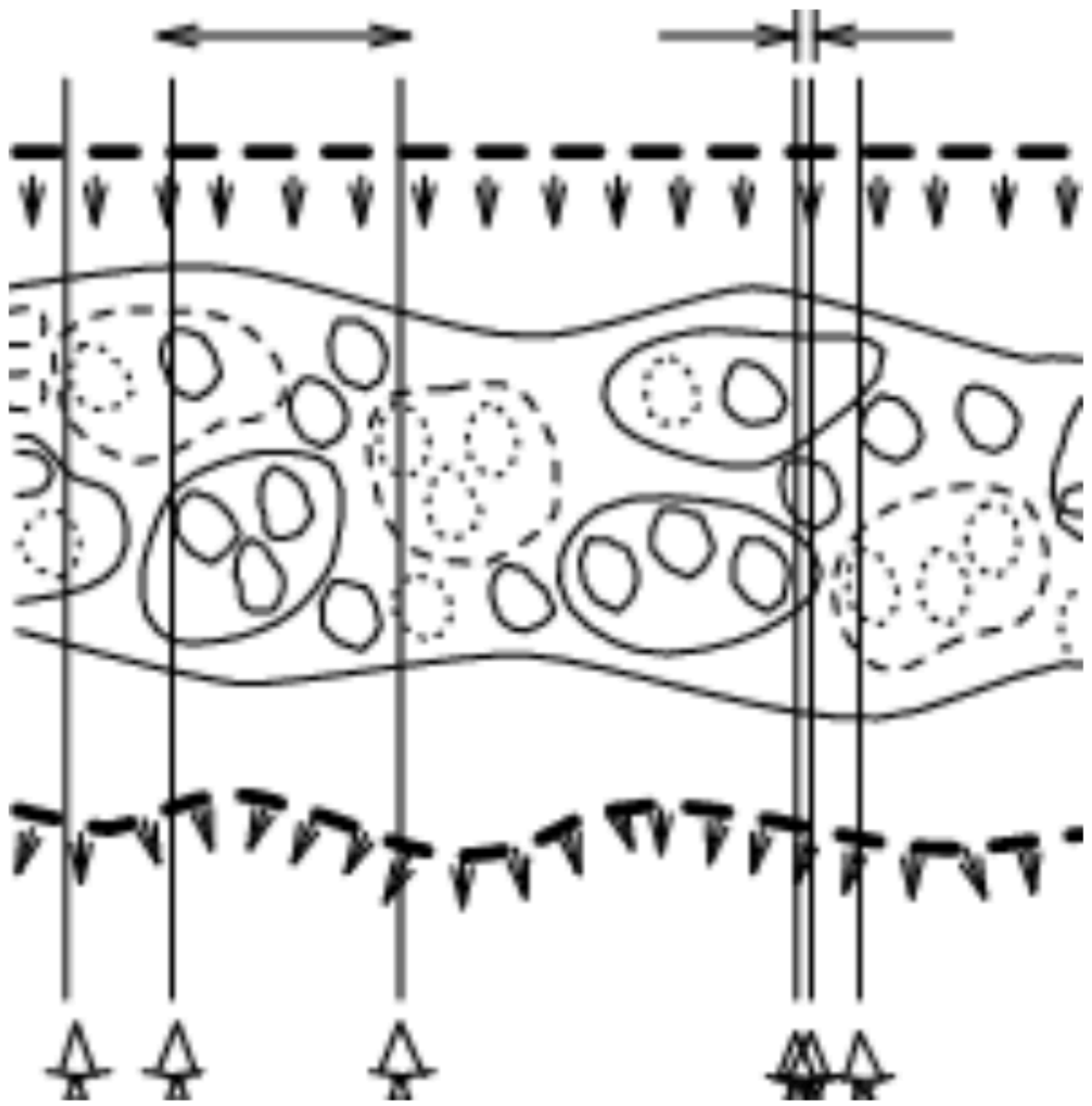
The true visibility is corrupted by many effects:

- Atmospheric attenuation
- Radio "seeing"
- Variable pointing offsets
- Variable delay offsets
- Electronic gain changes
- Electronic delay changes
- Electronic phase changes

Antenna-based

- Radiometer noise
- Correlator malfunctions
- Most Interference signals

baseline-based



2.3.1. Is self-cal possible for my data?

The minimum signal-to-noise required for self-cal solutions depends upon the level of error that you are willing to accept. For example, at a particular solution interval ($t_{\rm solint}$), an antenna-based signal-to-noise ratio $(S/N)_{Ant}(t_{\rm solint}) \gtrsim 3$ ensures that phase solution uncertainties will be less than $< 20°$, and $(S/N)_{Ant}(t_{\rm solint}) \gtrsim 10$ will induce amplitude solution uncertainties $< 10\%$. Henceforth, $(S/N)_{Ant}(t_{\rm solint})$ will be referred to as simply $(S/N)_{self}$. The $t_{\rm solint}$ required to actually substantially improve your data will depend on the dominant sources of error. As described above, errors that are "directional" like those induced by antenna position errors, only require a single $t_{\rm solint}$ averaged over 1 hour (i.e. a typical ALMA execution block length) for correction. Errors that are time dependent require $t_{\rm solint}$ shorter than the time for significant change to occur.

An easy way to determine $(S/N)_{self}$ is to make an initial image using the best practices for your observing mode (see § 2.3.4) below. Measure the rms noise in a representative signal-free region of the image. Convert this noise, σ_{image}, to that appropriate for a single antenna using the formula: $\sigma_{Ant} = \sigma_{image} \times \sqrt{n-3}$ where n is the number of antennas (Cornwell 1981). Then for a given $t_{\rm solint}$,

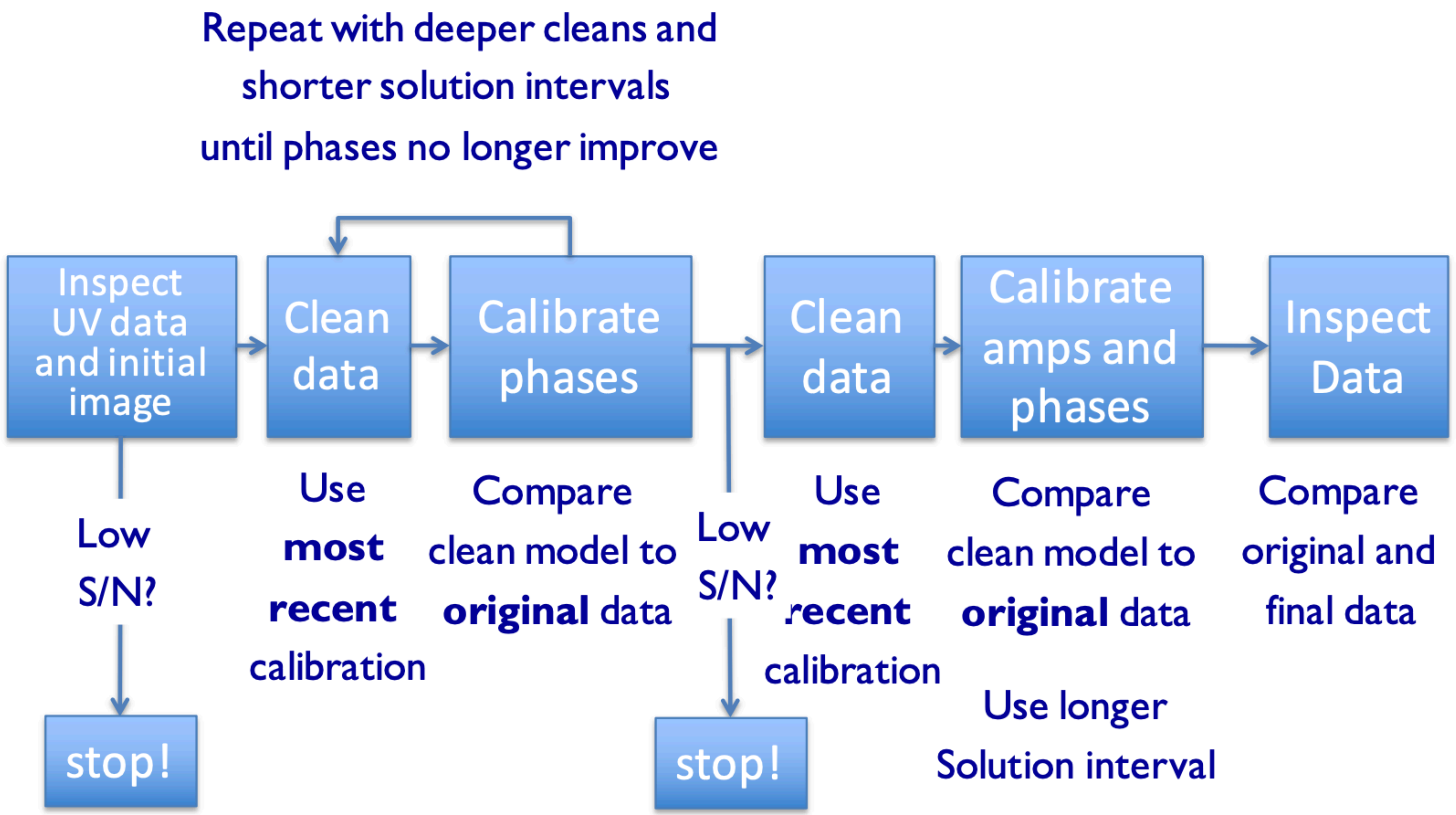
$$\sigma_{self} = \sigma_{image} \times \sqrt{n-3} \times \sqrt{\frac{t_{\rm on\_source}}{t_{\rm solint}}}. \quad (2)$$

In the simplest and most conservative assumption, the "signal" is the peak intensity (I_{peak}) in the image, so that $(S/N)_{self} = I_{peak}/\sigma_{self}$. However, it should be emphasized that σ_{image} will improve with self-cal, thus one should in principle re-evaluate the $(S/N)_{self}$ at each stage of the process. Additionally, there are a number of caveats that add complexity to this simple calculation.

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

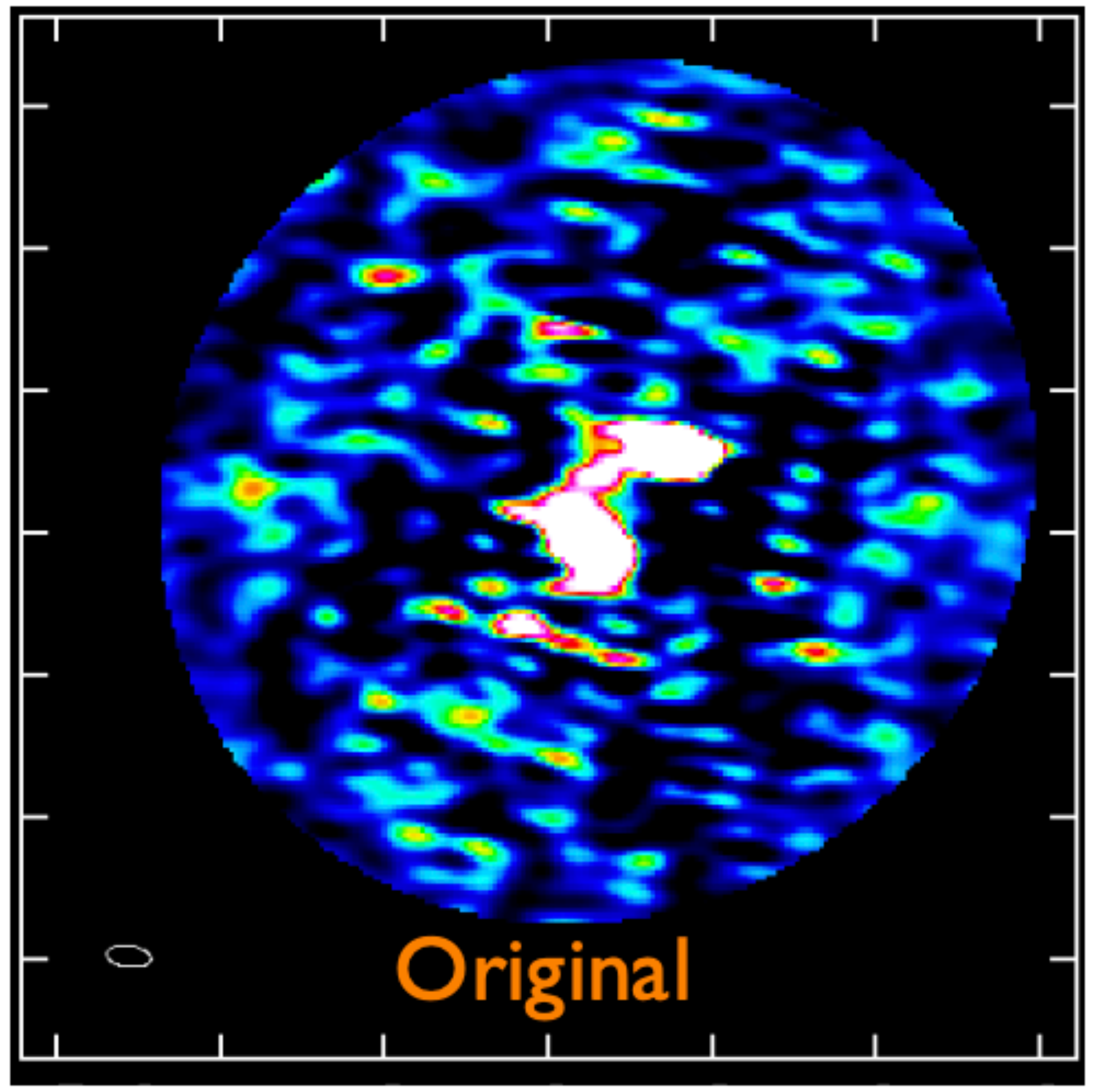
- Increase solint (solution interval)
  - If errors that are directional rather than time dependent, self-calibration solutions can yield surprising improvement even for solints that span the entire observation. Antenna position (aka baseline) errors are a good example.
- gaintype= 'T' to average polarizations
  - 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)

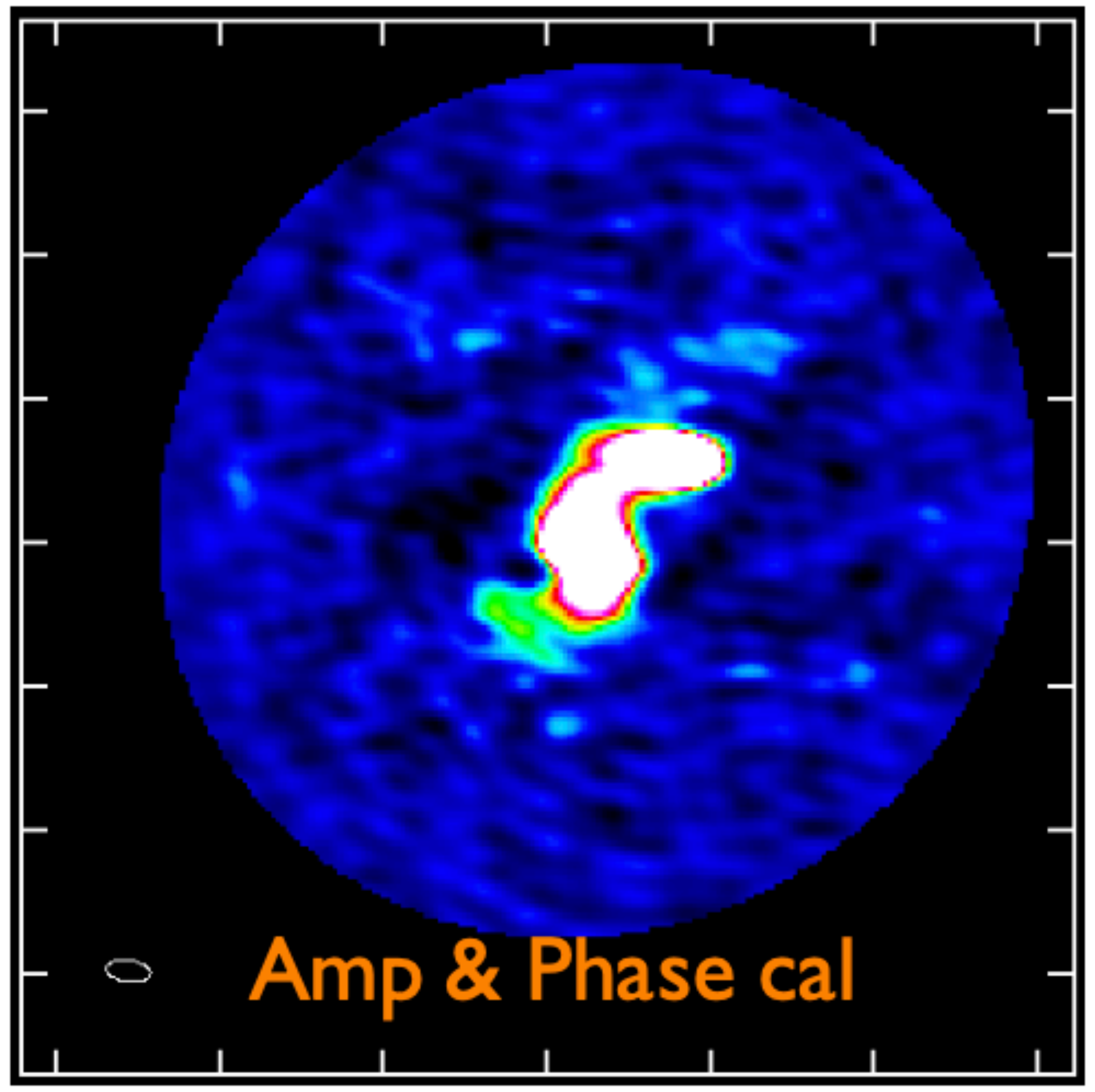
Outline of Self-Calibration Process



Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6



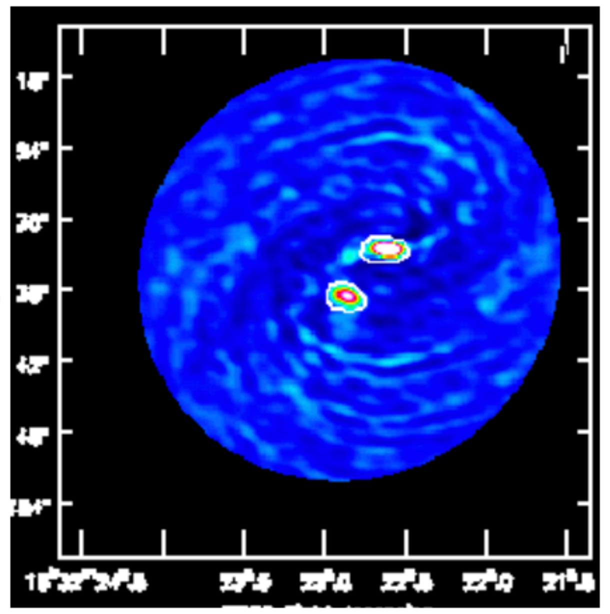


See IRAS16293 CASA guide for detailed commands.

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

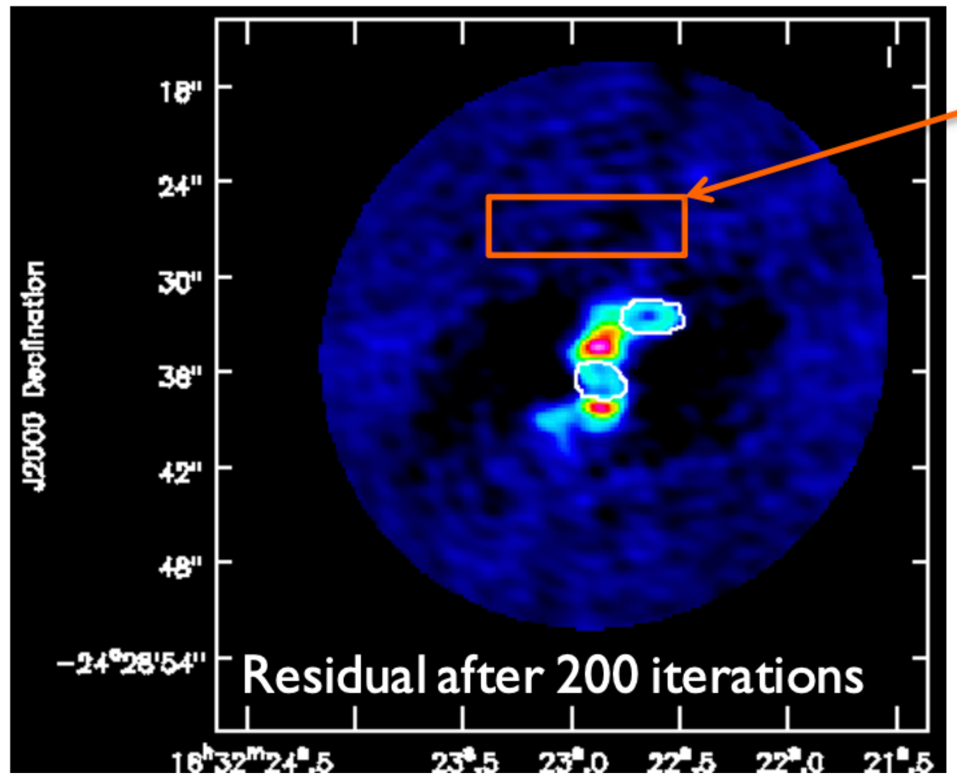
Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

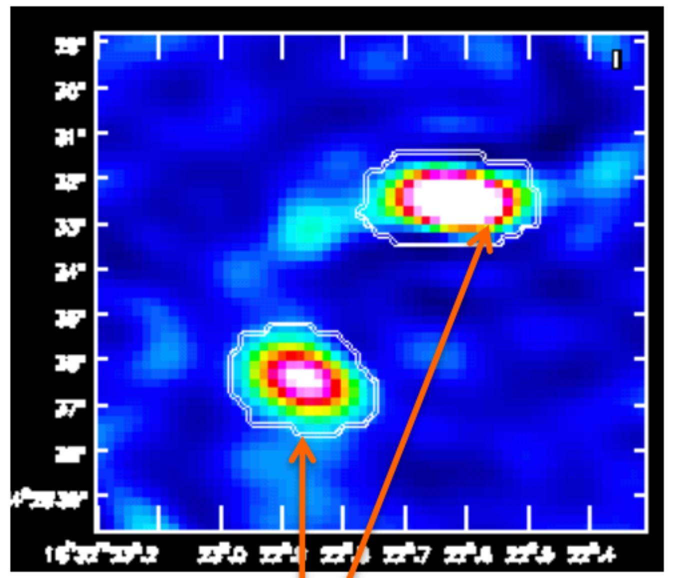


Step 4 – What is the S/N in a conservatively cleaned image?

- What is this "conservative" of which you speak?
- Rms~ 15 mJy/beam; Peak ~ 1 Jy/beam ➔ S/N ~ 67
- Rms > expected and S/N > 20 ➔ self-cal!



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



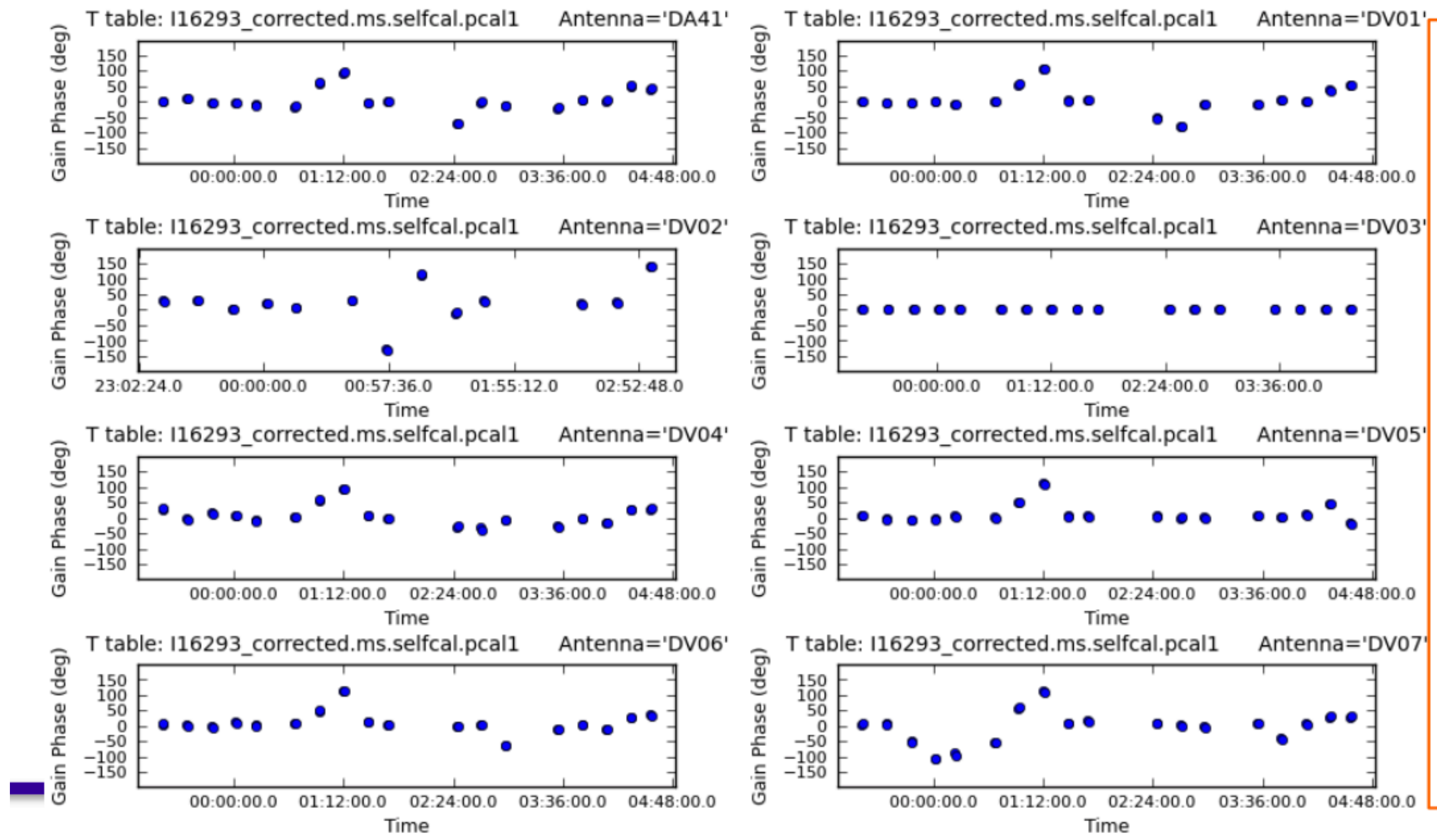
Clean boxes only around emission you are SURE are real at this stage

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 $S/N_{self} \sim 5.4$
- In CASA you can just set solint='inf' (i.e. infinity) and as long as combine ≠ 'scan' AND ≠ 'field' you will get one solution per scan, per field.
- Use 'T' solution to combine polarizations



What to look for:

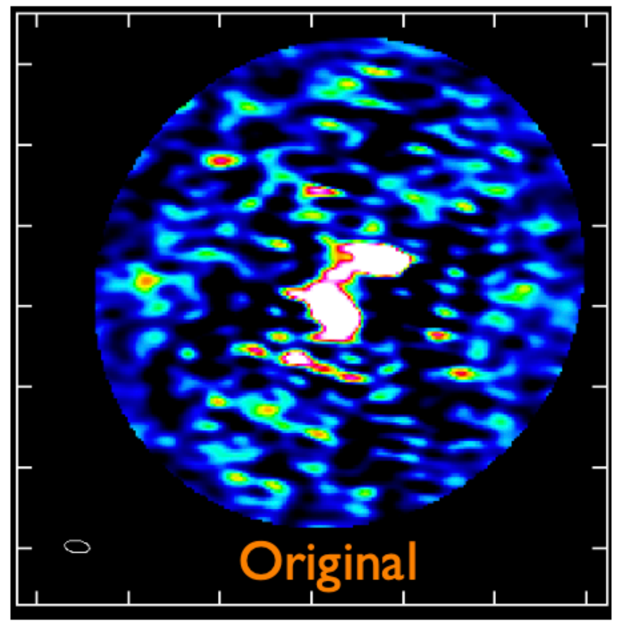
- Lot of failed solutions on most antennas? if so, go back and try to increase S/N of solution = more averaging of some kind
- Do the phases appear smoothly varying with time (as opposed to noise like)

Self-calibration Example:

ALMA SV Data for IRAS16293 Band 6

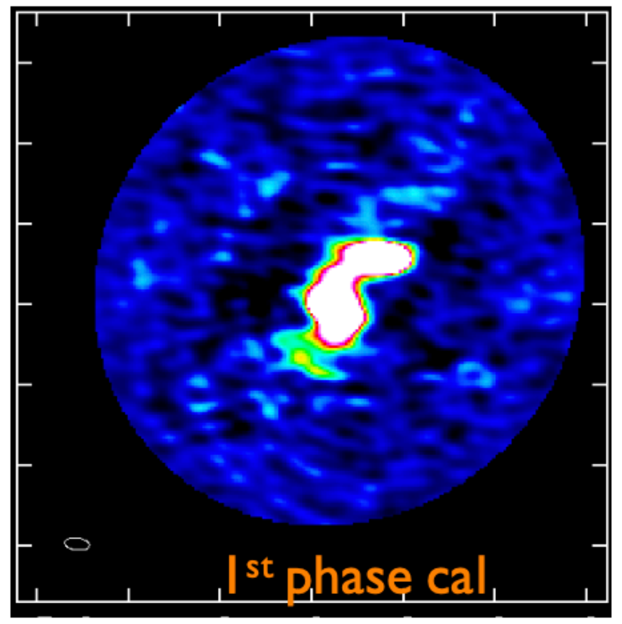
Step 6: Apply solutions and re-clean

- 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



Step 7: Compare Original clean image with 1st phase-only self-cal image

- 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.

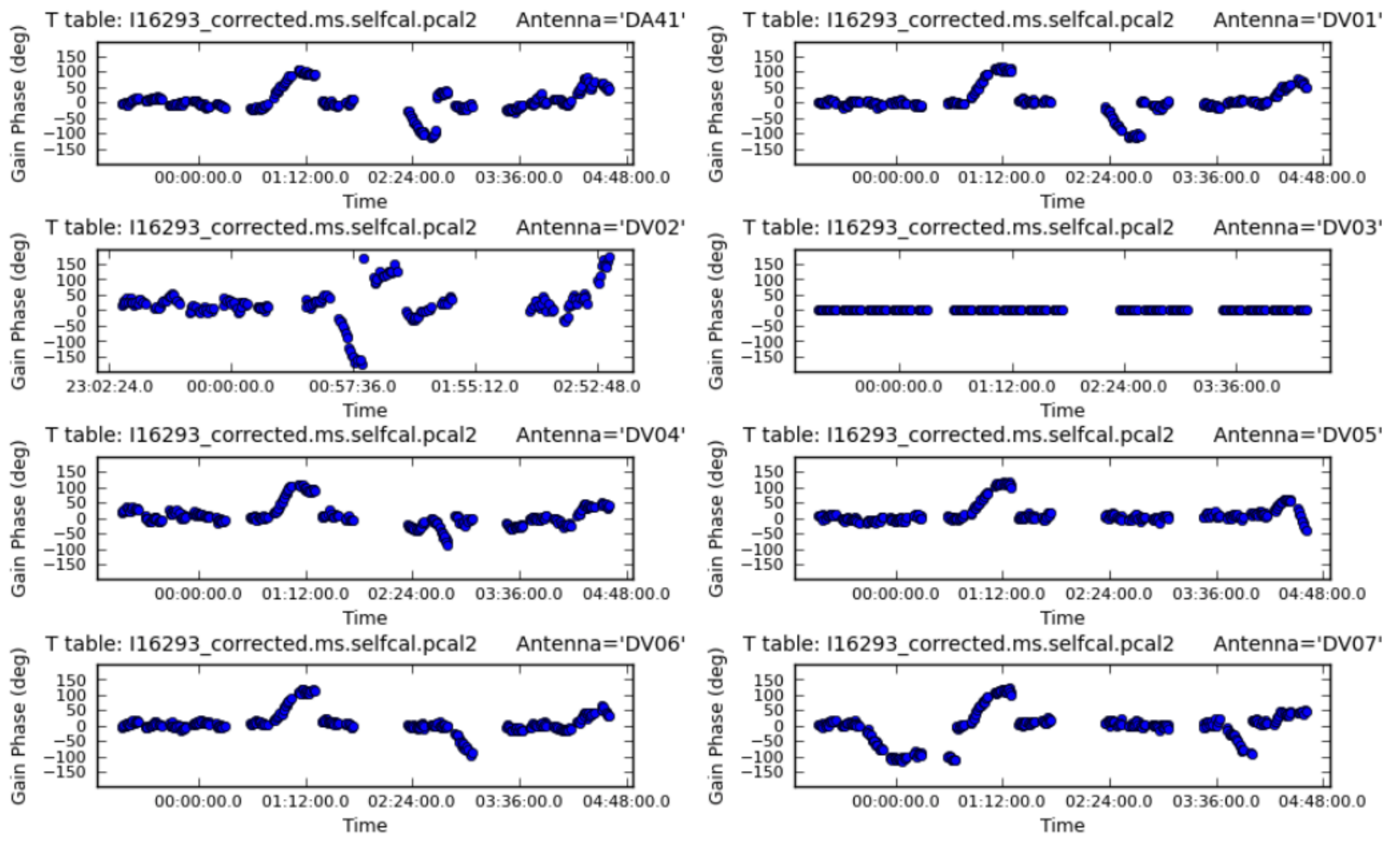


Self-calibration Example:

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



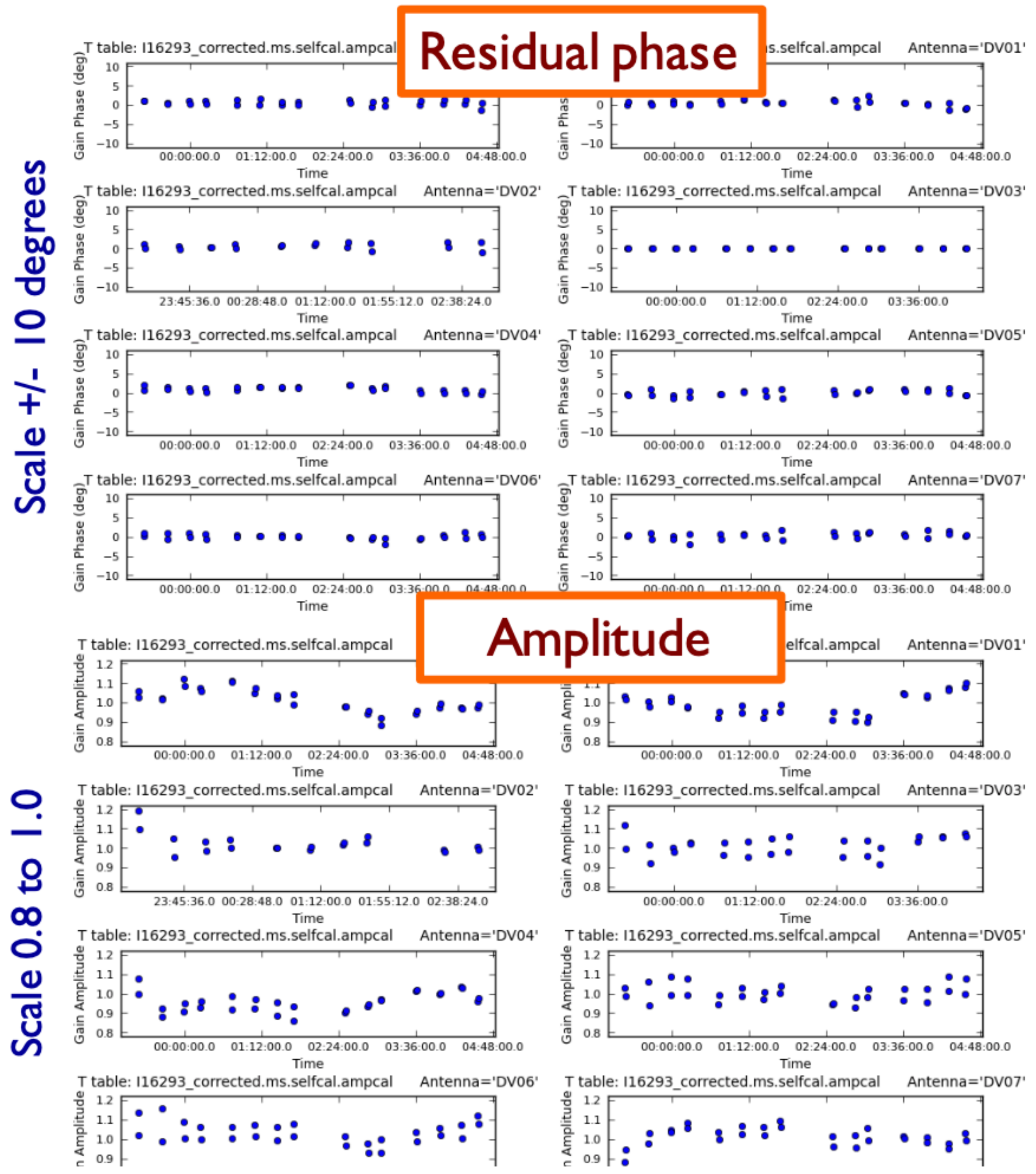
What to look for:

- 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 IRAS16293 Band 6

Step 11: 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' (in gaincal) to check that residual phase solutions are close to zero.
- Again make sure mostly good solutions, and a smoothly varying pattern.

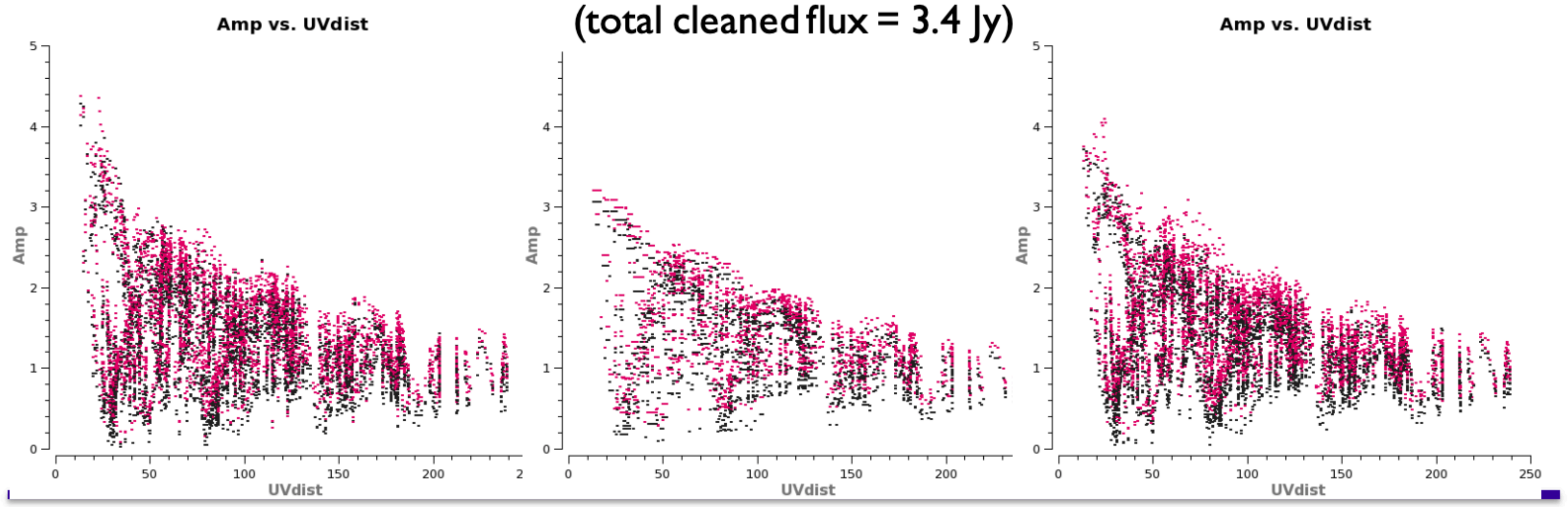


Self-calibration Example:
ALMA SV Data for IRAS16293 Band 6

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 IRAS16293 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