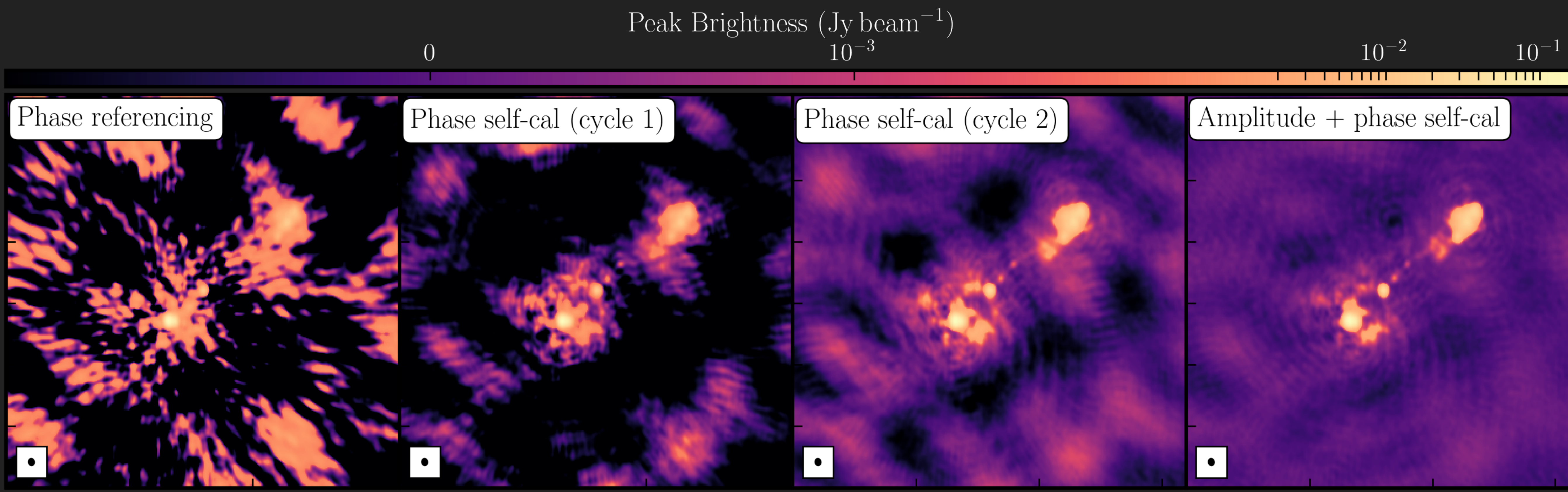


EUROPEAN RADIO INTERFEROMETRY SCHOOL 2019

SELF-CALIBRATION

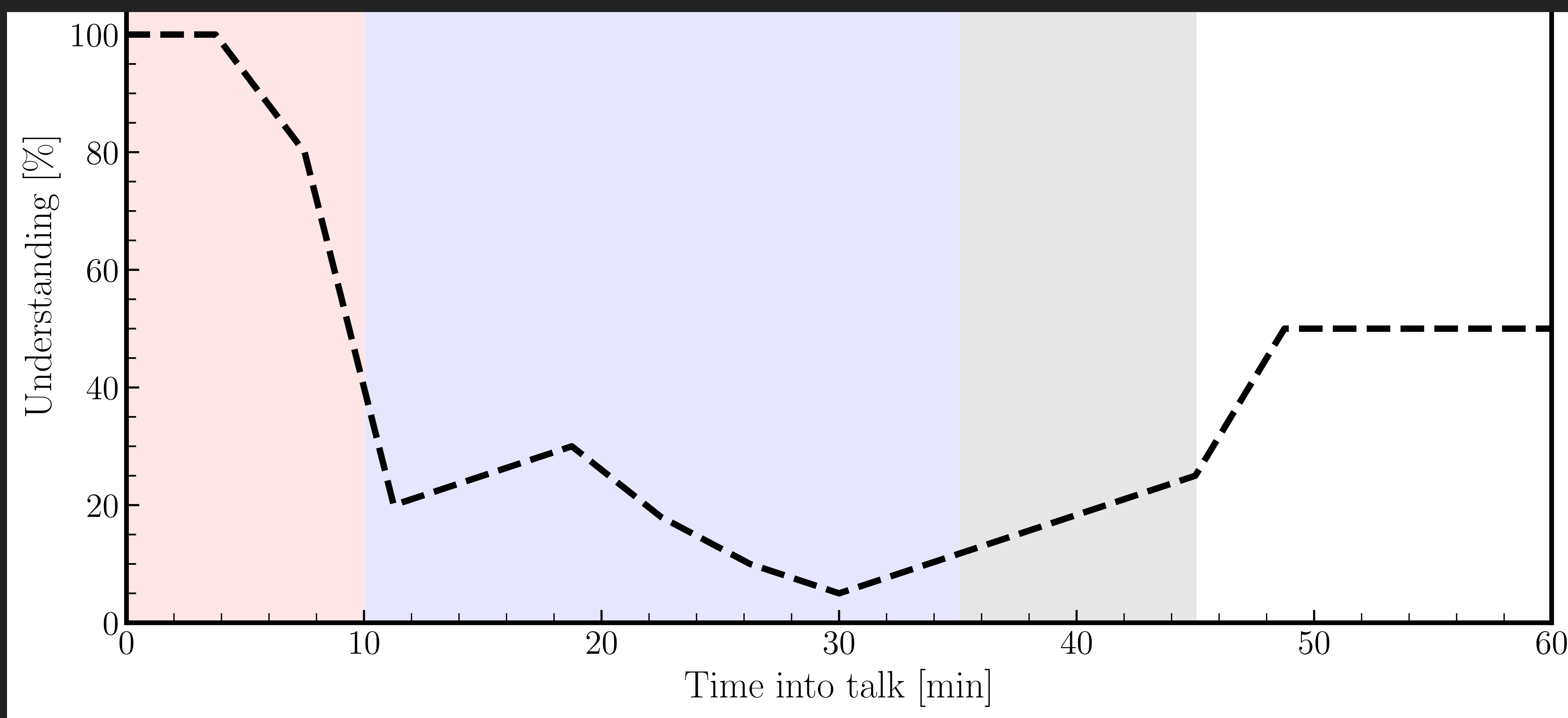
Jack Radcliffe & Anita Richards

Universiteit van Pretoria
University of Pretoria
Yunibesithi ya Pretoria
RadioNet
MANCHESTER 1824
The University of Manchester
NRF National Research Foundation
SARAO South African Radio Astronomy Observatory

A DISCLAIMER…

1. Some of this talk is technically orientated so… If you don't understand some jargon → **interrupt me**
2. I have a "Northern" (actually Midlands…) British accent so if you miss something → **also tell me please!**

There is nothing worse than a talk where:



OUTLINE

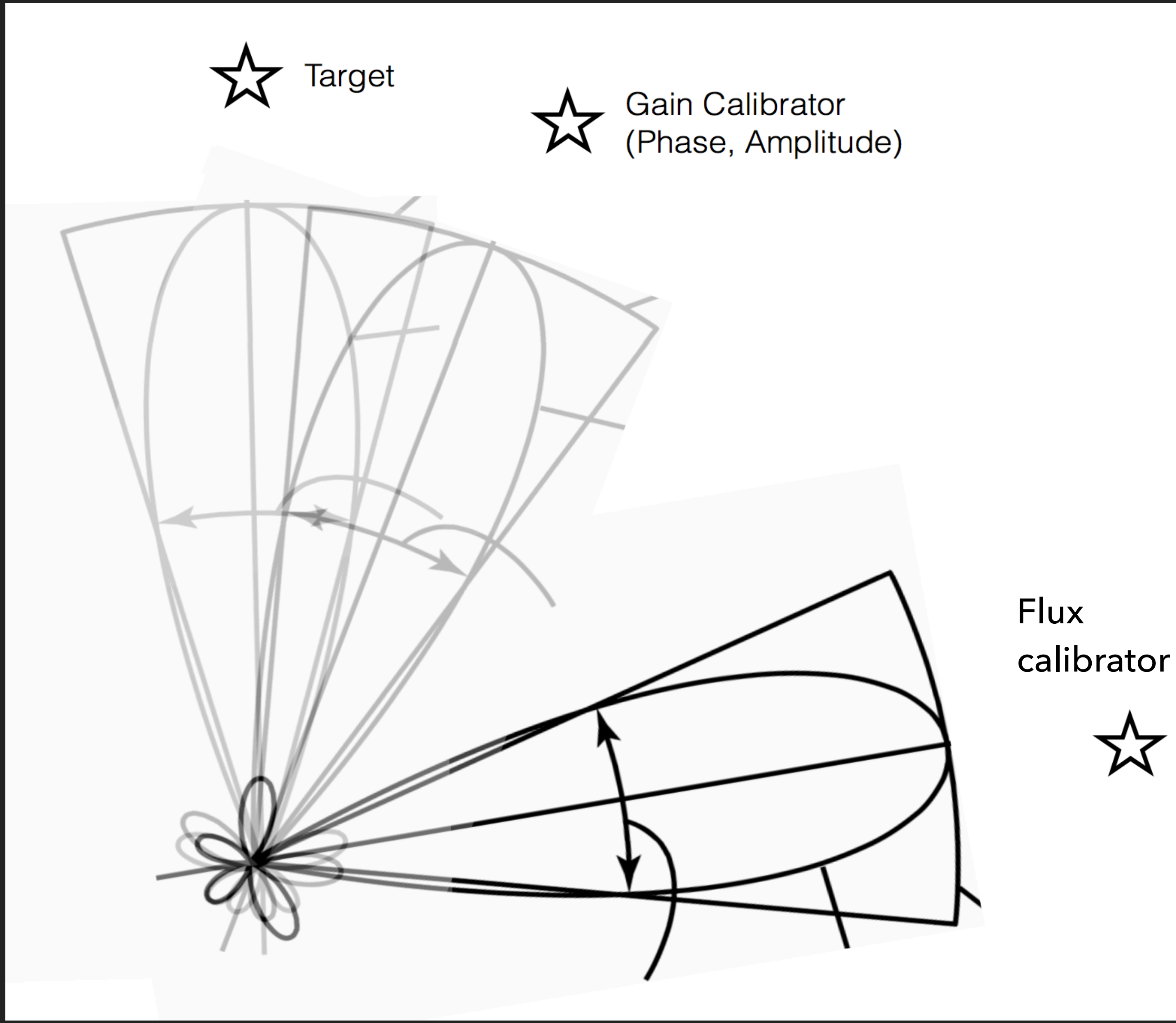
Welcome to the wonderful world of self-calibration!

1. Re-cap of phase-referencing
2. Why phase-referencing is insufficient
3. Self-calibration theory & principles
4. Self-calibration in practice (CASA)
5. Choices when self-calibrating (i.e. knowing when you do a good job!)

And thanks to Anita Richards (for helping too!), Joe Callingham and Andre Offringa for slides & ideas

PHASE REFERENCING

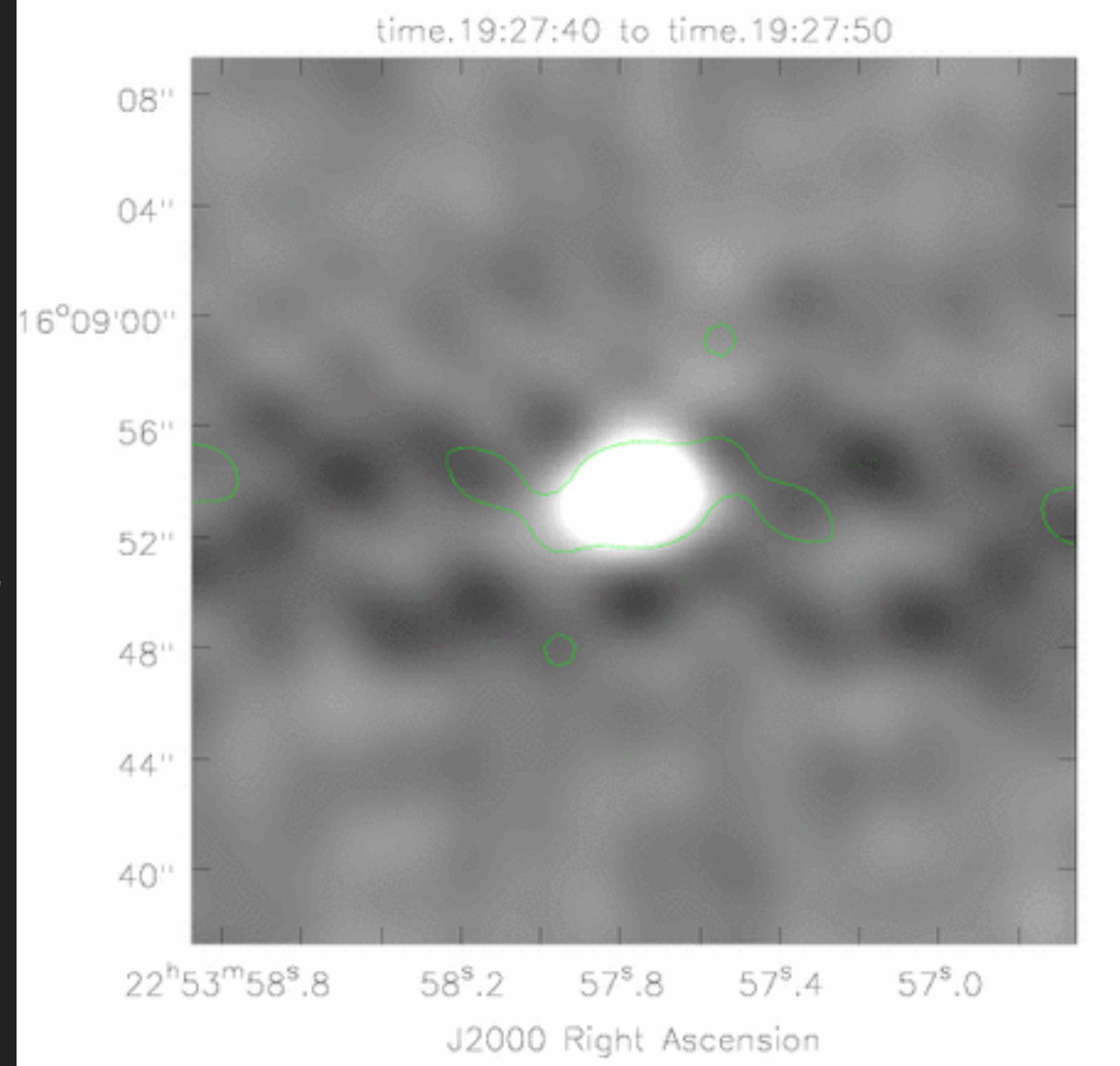
- Apply instrumental corrections: e.g.
- Edit the data as required
- Apply bandpass, (polarisation) corrections
- Apply phase and (flux-scaled) amplitude corrections derived from phase reference
- Close enough to see similar atmosphere
- Nod on suitable timescale e.g. 10:2min
- Derive time-dependent corrections to make phase-ref visibilities match model
- Apply same corrections to target



THE ATMOSPHERIC CONTRIBUTION

- The atmosphere is similar, not identical, above the target and above the phase-reference.
- There are offsets in distance and in time.
- Neutral atmosphere contains water vapour
- Index of refraction differs from "dry" air
- Variety of moving spatial structures in the atmosphere.
- Typically worse for low frequencies $\sim 100\text{s MHz}$ (ionosphere) & high frequencies ≥ 20 GHz (water vapour)

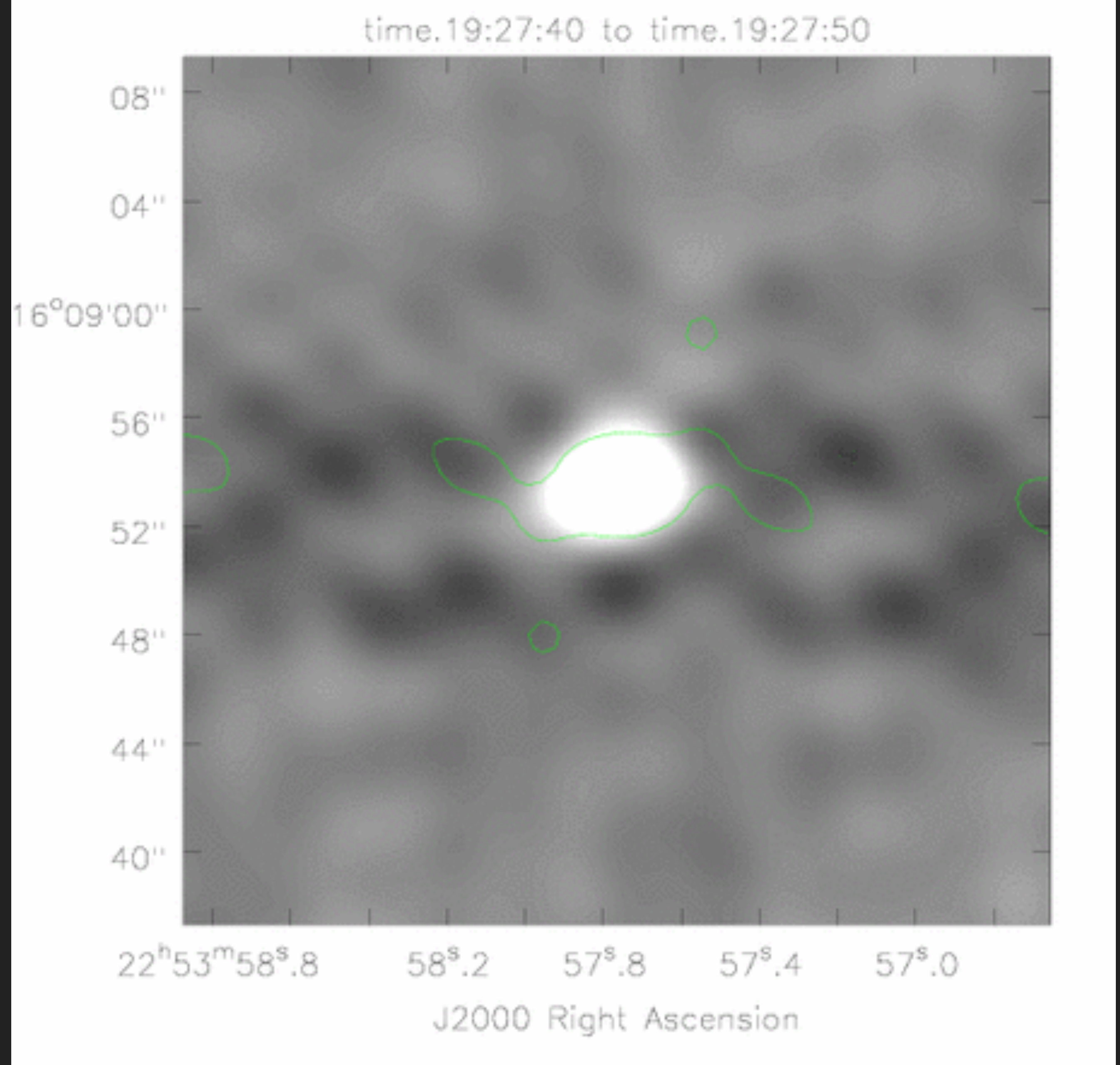
Example 22GHz point source observed with the EVLA



THE ATMOSPHERIC CONTRIBUTION

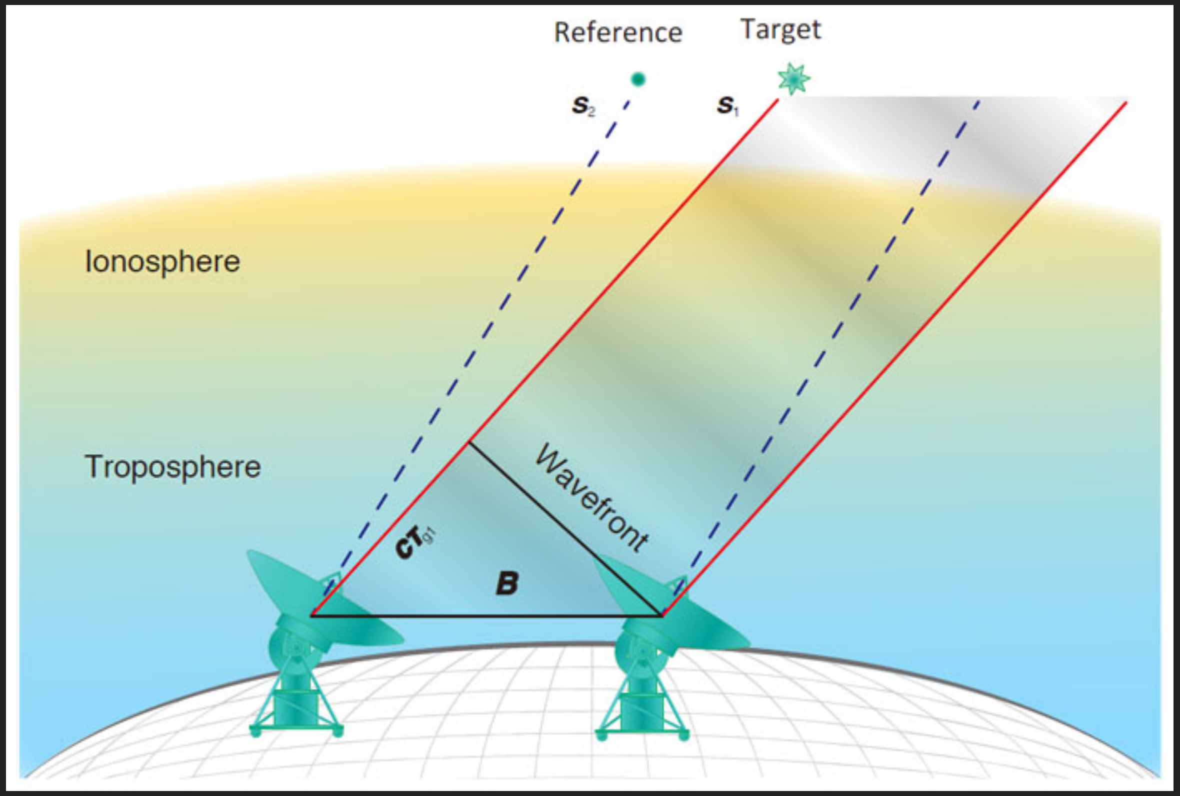
- The atmosphere is similar, not identical, above the target and above the phase-reference.
- There are offsets in distance and in time.
- Neutral atmosphere contains water vapour
- Index of refraction differs from "dry" air
- Variety of moving spatial structures in the atmosphere.
- Typically worse for low frequencies $\sim 100\mathrm{s\ MHz}$ (ionosphere) & high frequencies $\geq 20\ \mathrm{GHz}$ (water vapour)

Example 22GHz point source observed with the EVLA



SELF-CALIBRATION

- The process of self calibration is to **primarily correct for complex gain variations between target and phase calibrator positions,** typically induced by the atmosphere!
- Only useful for targets where the expected S/N is high - typically >20 for arrays with 6-10 elements or >100 for 20-50 elements
- If you don't expect high enough S/N to self-cal (how to estimate this to be explained slide 45) then the only way to improve the dynamic range would be by improving the calibration or removing bad data, confusion etc.



THE CALIBRATION ERROR ZOO

The atmosphere is not our only worry, our true visibilities are corrupted by lots of other effects. Self-calibration can help with the antenna-based errors.

Atmospheric attenuation

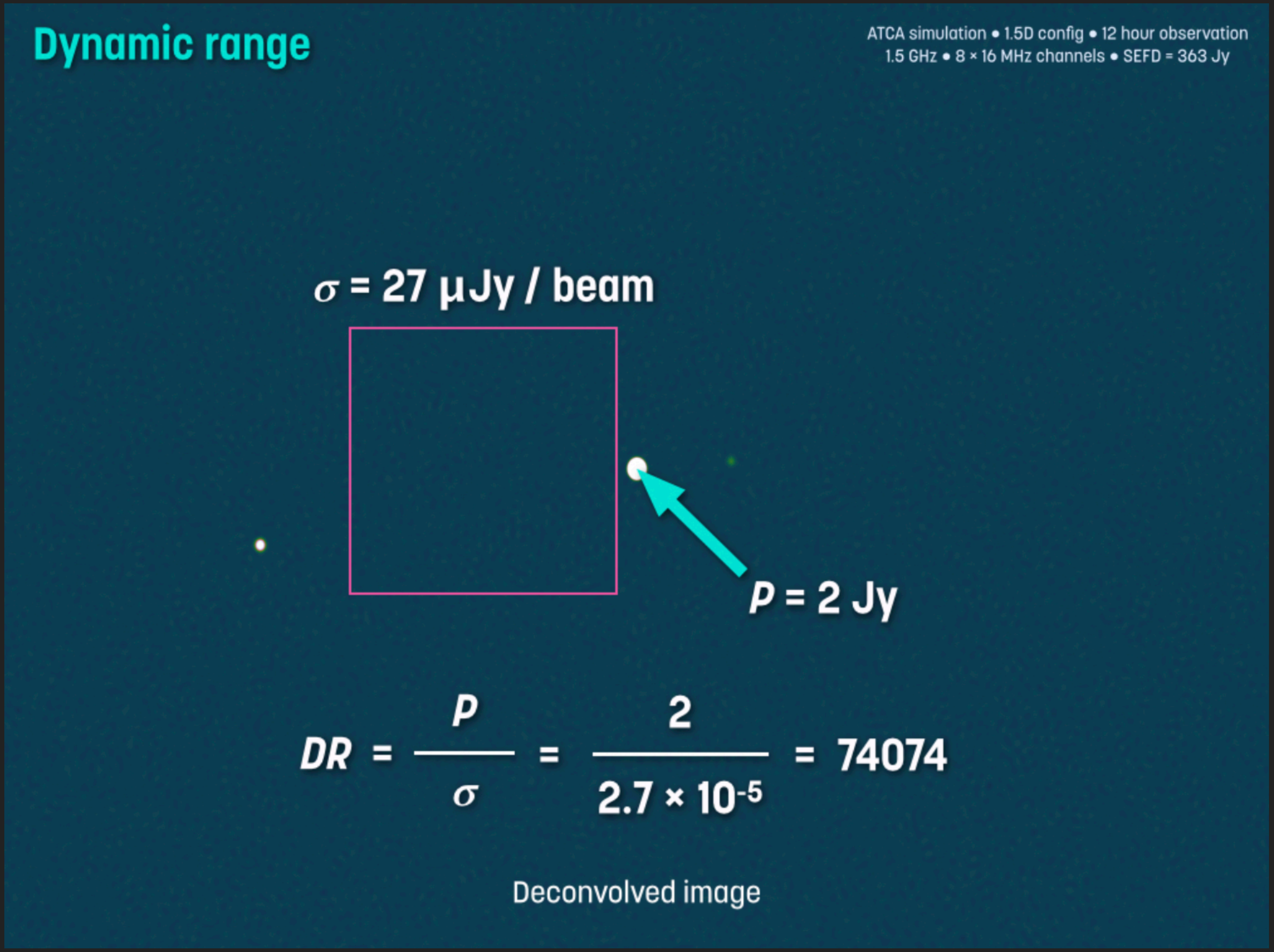
Antenna

1. Radio “seeing”
2. Variable pointing offsets
3. Variable delay offsets
4. Electronic gain changes
5. Electronic delay changes
6. Electronic phase changes

Baseline

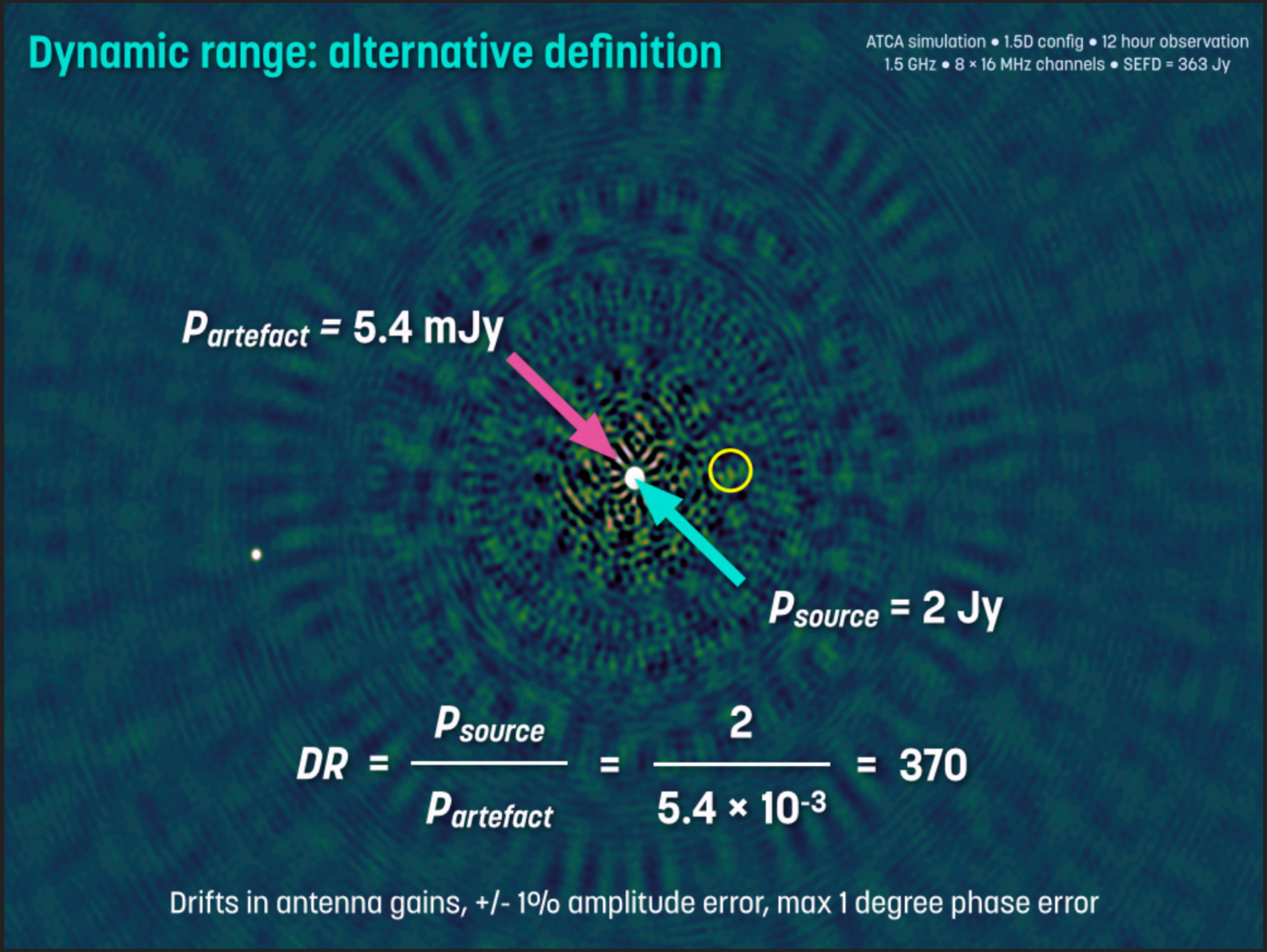
7. Heterogenous PB corrections
8. Correlator accuracy limitations
9. Most interference signals

AN ASIDE: DYNAMIC RANGE



Credit: I Heywood

AN ASIDE: DYNAMIC RANGE – ALTERNATIVE DEFINITION

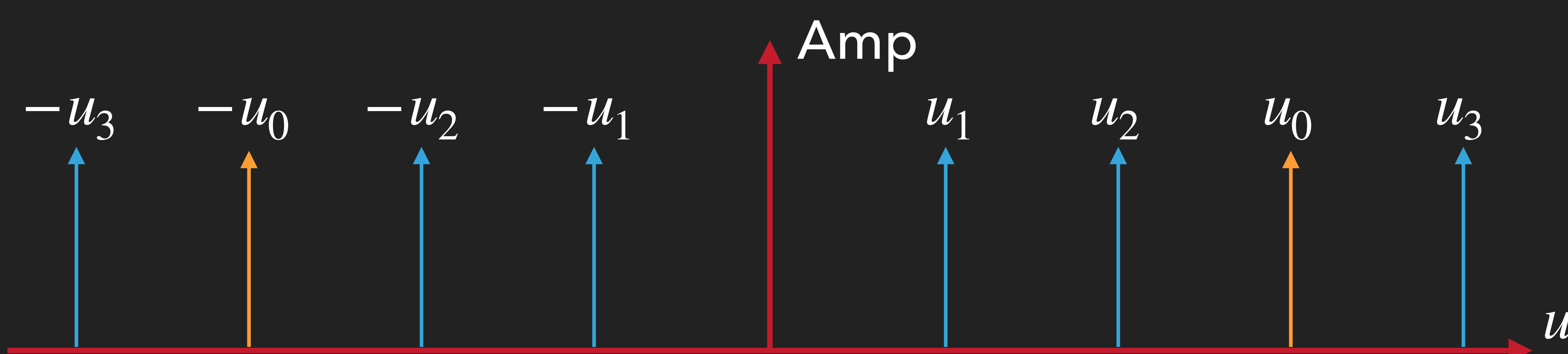


Credit: I Heywood

NB: Noise extrema can be +ve or -ve

WHAT HAPPENS IF YOU LEAVE THE ERRORS IN?

- Flat, linear array, N antennas (only need to consider u axis)
- Single integration observation of a point source - $N(N-1)/2$ visibilities
- Each baseline visibility is a δ spike in the *uv* plane
- All but one are 'perfect' (unit amplitude, zero phase)
- These have $V(u) = \delta(u - u_k)$ for the kth baseline.
- Phase error on baseline length u_0 of ϕ_ϵ radians so $V(u) = \delta(u - u_0)\exp(-i\phi_\epsilon)$



PHASE ERRORS AND DYNAMIC RANGE

- Image is formed via FT: $I(l) = \int V(u)\exp(i2\pi ul)\mathrm{d}u$
- Each baseline contributes at position u_k and complex conjugate $-u_k$ in the visibility plane
- Evaluating term for every $\left(N(N-1)/2\right) - 1$ good baselines gives $2\cos\left(2\pi u_k l\right)$
- Single bad baseline contributes - $2\cos\left(2\pi u_0 l - \phi_\epsilon\right)$
- Assuming small ϕ_ϵ gives: $\approx 2\left[\cos\left(2\pi u_0 l\right) + \phi_\epsilon \sin\left(2\pi u_0 l\right)\right]$
- The image integral becomes:

* $\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$

$$I(l) = 2\phi_\epsilon \sin\left(2\pi u_0 l\right) + 2\sum_{k=1}^{N(N-1)/2} \cos\left(2\pi u_k l\right)$$

PHASE ERRORS AND DYNAMIC RANGE

- The synthesised beam in this case is

$$B(l) = 2 \sum_{k=1}^{N(N-1)/2} \cos\left(2\pi u_k l\right)$$

- Deconvolution is the subtraction of the beam from the image which leaves the residual error:

$$R(l) = I(l) - B(l) = \left[2\phi_\epsilon \sin\left(2\pi u_0 l\right) + 2 \sum_{k=1}^{N(N-1)/2} \cos\left(2\pi u_k l\right)\right] - 2 \sum_{k=1}^{N(N-1)/2} \cos\left(2\pi u_k l\right)$$

$$= 2\phi_\epsilon \sin\left(2\pi u_0 l\right)$$

- This is an ‘odd’ sinusoidal with amplitude $2\phi_\epsilon$ and period $1/u_0$

PHASE ERRORS AND DYNAMIC RANGE

- For a small phase error, ϕ_ϵ, and large N, the ratio of peak to noise residual is:

$$D_{\mathrm{B}}\left(\phi_\epsilon\right) \sim I(l)/R(l) \sim N^2/\sqrt{2}\phi_\epsilon$$

- An amplitude error on a single baseline has the effect of $V(u) = (1+\epsilon)\delta\left(u - u_0\right)$ leading (via a cos function) to:

$$D_{\mathrm{B}}(\epsilon) \sim I(l)/R(l) \sim N^2/\sqrt{2}\epsilon$$

- Therefore:
  - **A phase error of 10° is as bad as a $20\,\%$ amplitude error**
  - **Phase errors are sin (odd) and amplitudes are cos (even)**

LIMITING DYNAMIC RANGE

- Only considered one integration + one baseline
- Antenna based errors (all baselines to one antenna affected by same error)
  - $(N-1)$ bad baselines ($\sim N$ for lots of baselines)
  - $D_{\mathrm{ant}} = D_B/(N-1) = \left[N^2/(N-1)\right]/\sqrt{2}\phi_\epsilon \sim \mathbf{N}/\sqrt{\mathbf{2}}\phi_\epsilon$
- **All** baselines affected by random noise:
  - $D_{\mathrm{all}} = D_{\mathrm{B}}/\sqrt{N(N-1)/2} = \sqrt{N(N-1)/2)}/\phi_\epsilon \sim \mathbf{N}/\phi_\epsilon$
- Expressions valid if errors are correlated in time i.e. single phase ref. scan or not much change in u, v, w

DYNAMIC RANGE LIMITATIONS

- For M periods (scans?) between which noise is uncorrelated:

  Dynamic range is increased - $D_{\text{all}} \sim \sqrt{M}N/\phi_\epsilon$

- For **10** antennas & **12 independent scans -** all phase referencing applied + RFI edited

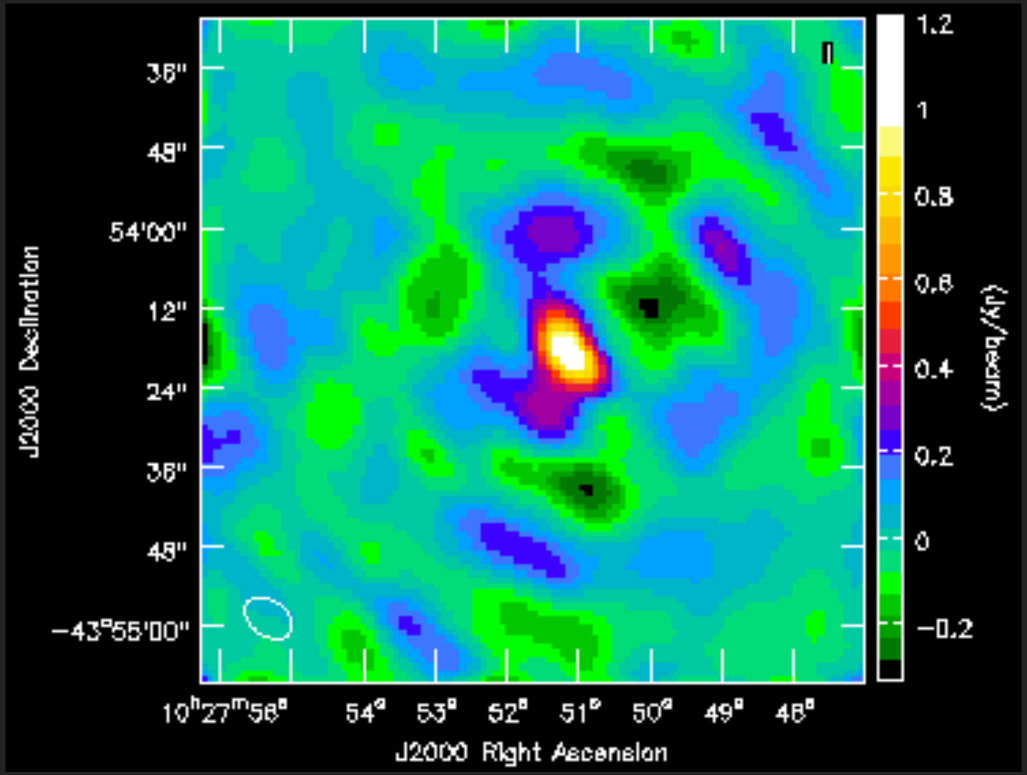
  Typical residual phase scatter $\sim 20^\circ$ so $D_{\text{all}} \sim \sqrt{M}N/\phi_\epsilon \sim 100$

Can we improve on this??

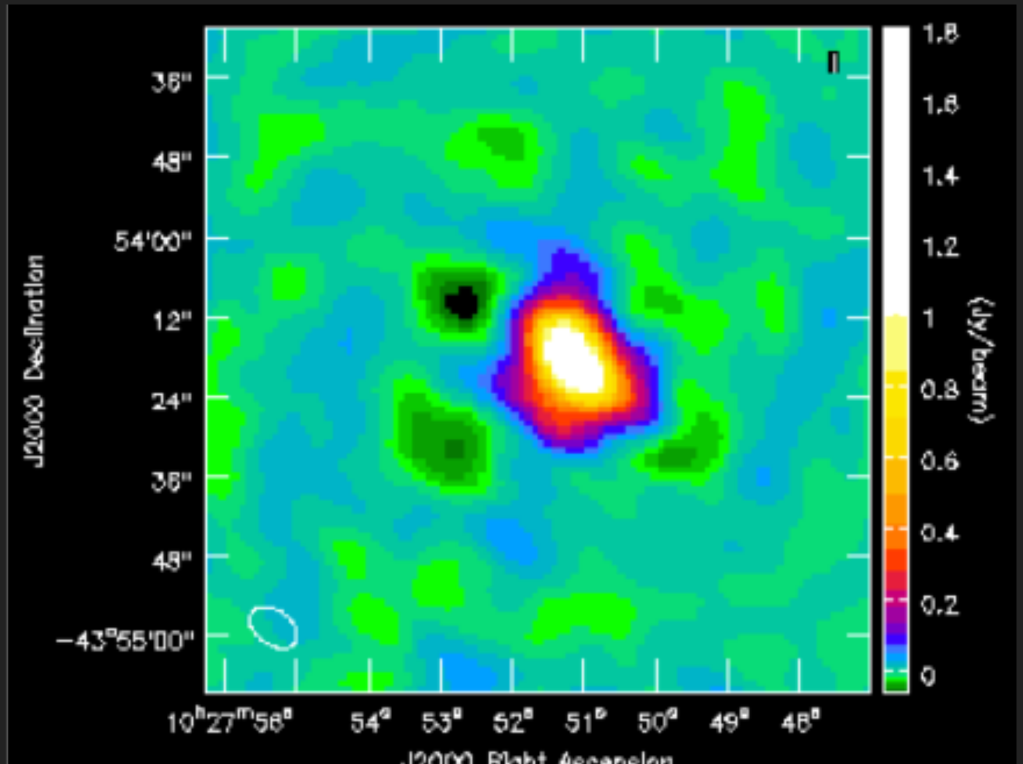
- If map noise near thermal, T_{sys}, noise level then no!
- If noise is non-Gaussian + shows errors and well above T_{sys} limit then there's calibration errors

So we need to use **self-calibration** to reduce phase errors below $\sim 20^\circ$ provided target is bright enough!

IDENTIFYING IMAGE ERRORS

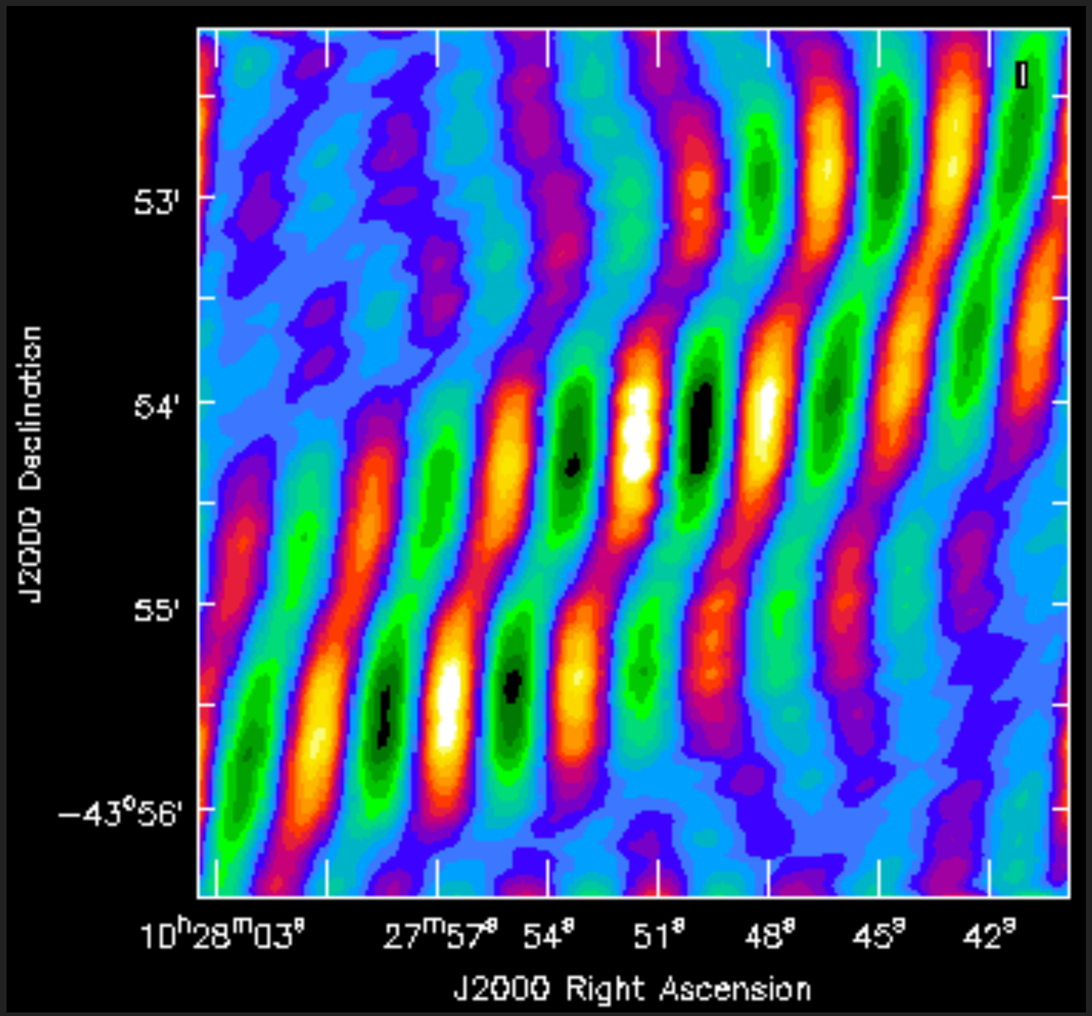


Phase referencing solutions only
Anti-symmetric phase errors dominate
Remember sin function?



Phase errors removed via self-calibration
Symmetric amplitude errors dominate

IDENTIFYING IMAGE ERRORS



A few high amps (easily flagged) cause symmetric stripes

Asymmetric stripes are usually delay errors! (cannot self-calibrate these out easily…)

Can see why phase referencing is insufficient now!
We need to use self-calibration to fix these errors that limit our image fidelity!

SELF-CALIBRATION – RECAP OF THE RIME

- Remember the radio measurement equation

$$\overrightarrow{V}_{ij}^{\text{obs}} = M_{ij}B_{ij}F_{ij}G_{ij}D_{ij}E_{ij}P_{ij}T_{ij}\overrightarrow{V}_{ij}^{\text{true}}$$

  - $\overrightarrow{V}_{ij}^{\text{obs}}$ - visibility measured between antennas i and j
  - $G_{ij}D_{ij}E_{ij}$ etc. - Jones matrices representing corrupting phenomena on baseline i, j
  - $\overrightarrow{V}_{ij}^{\text{true}}$ - 'true visibility i.e. $\iint I(l, m)e^{-2\pi i(ul+vm)}dldm$

This equation is represented in matrix form & the complex gain terms decomposed into Jones matrices which correspond to various sources of error.

SELF-CALIBRATION

- For self-calibration, we assume that most other effects apart from residual complex gain errors are calibrated therefore:

$$\overrightarrow{V}_{ij}^{\text{obs}} = G_{ij}\overrightarrow{V}_{ij}^{\text{true}}$$

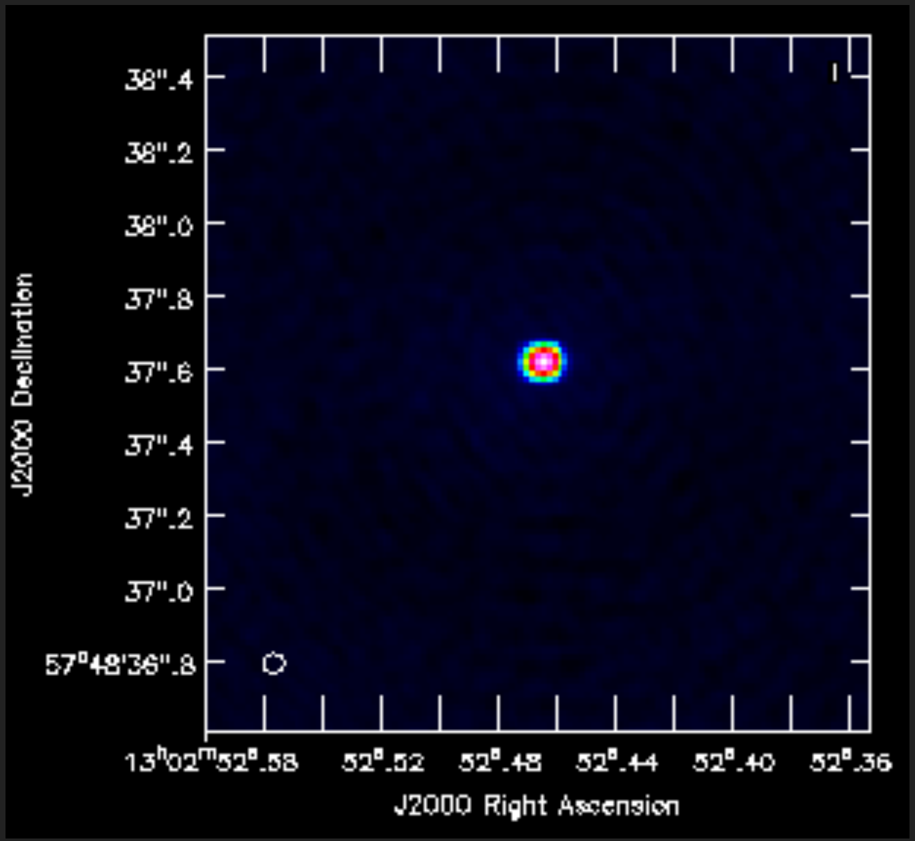
- And we assume that the complex gains are antenna-based (as with standard calibration)
  - Phases - atmosphere above telescope
  - Amplitudes - variable receiver gains

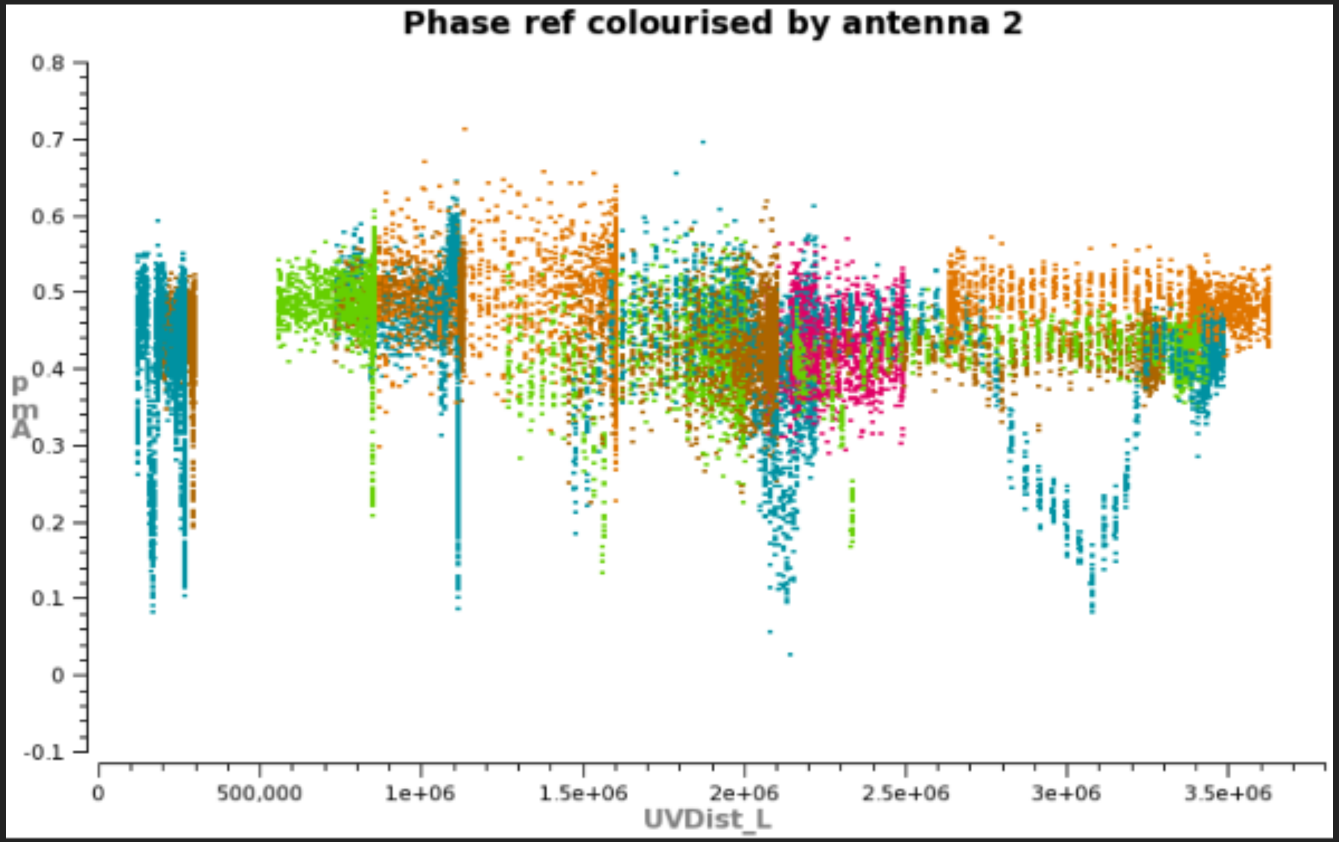
$$\overrightarrow{V}_{ij}^{\text{obs}} = G_i G_j \overrightarrow{V}_{ij}^{\text{true}}$$

SELF-CALIBRATION

- Calibration methodology is same as standard calibration **BUT** rather than a point source model, we use a model of the target visibilities i.e.

Standard calibration - point source

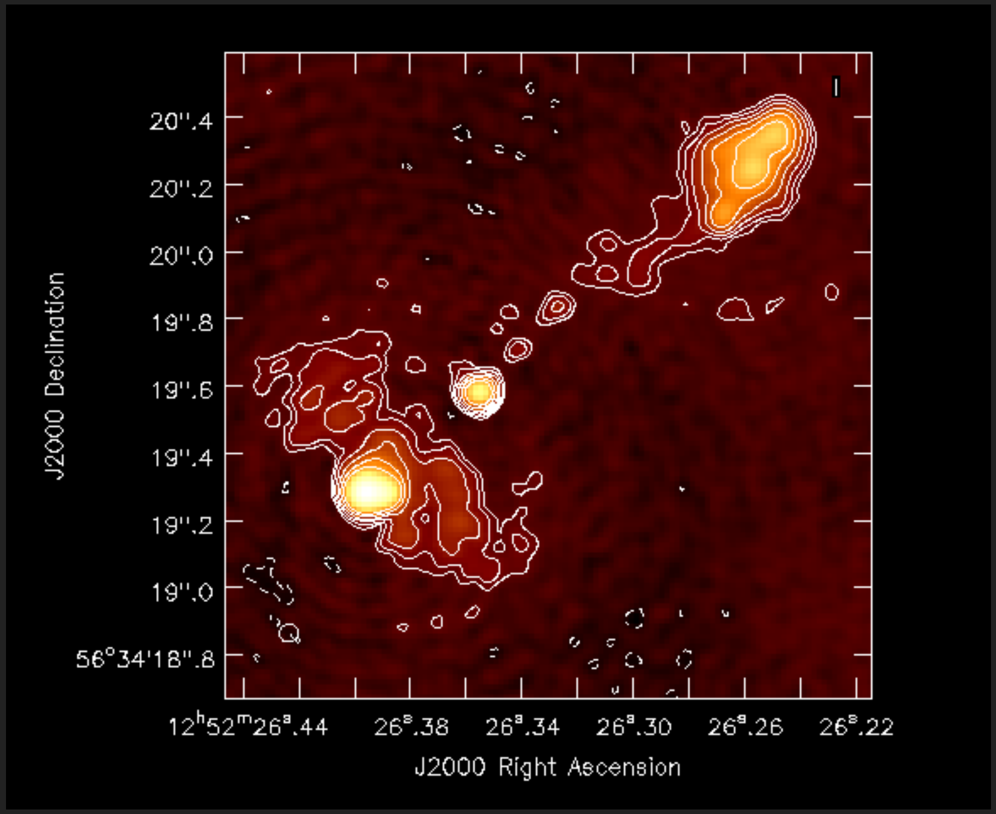


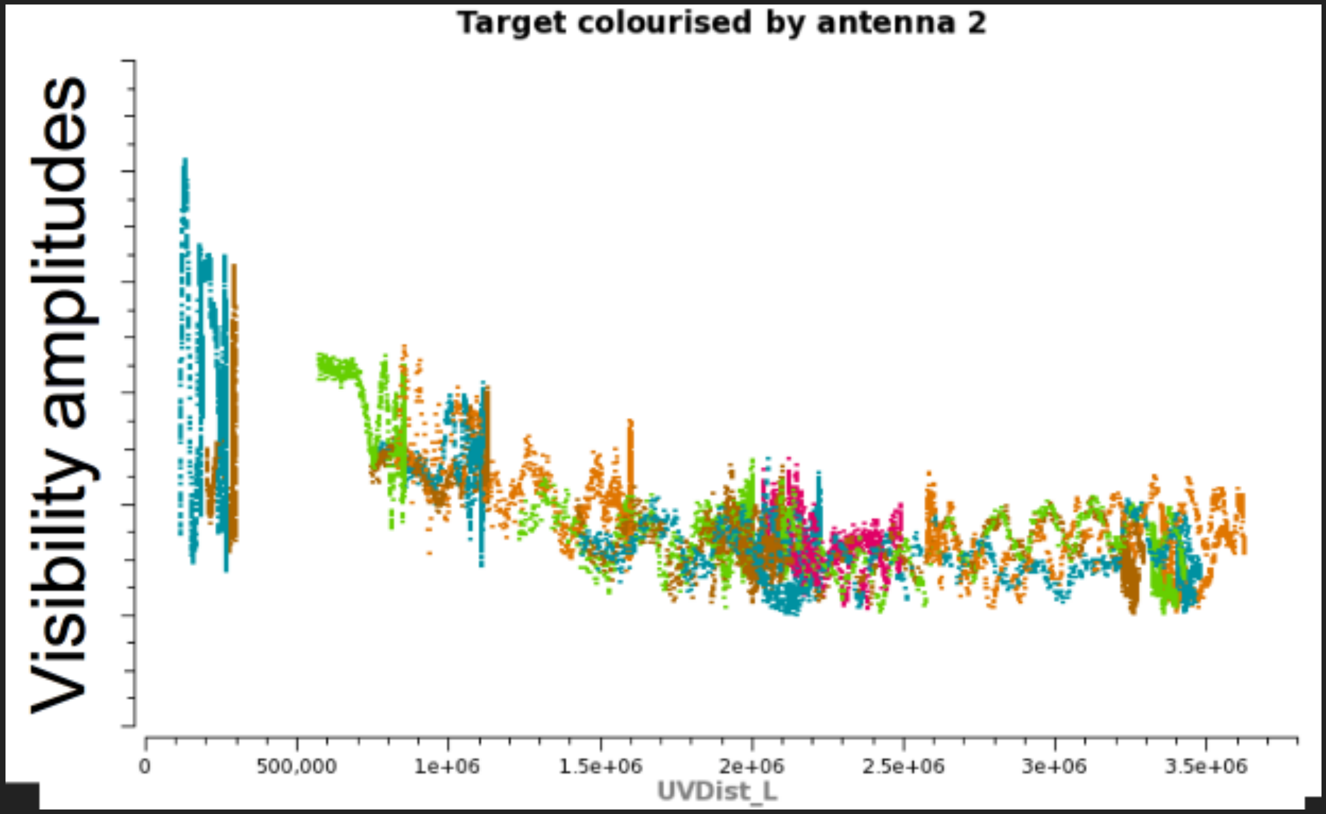


SELF-CALIBRATION

- Calibration methodology is same as standard calibration **BUT** rather than a point source model, we use a model of the target visibilities i.e.

Self-calibration - extended/multiple source(s) (often!)

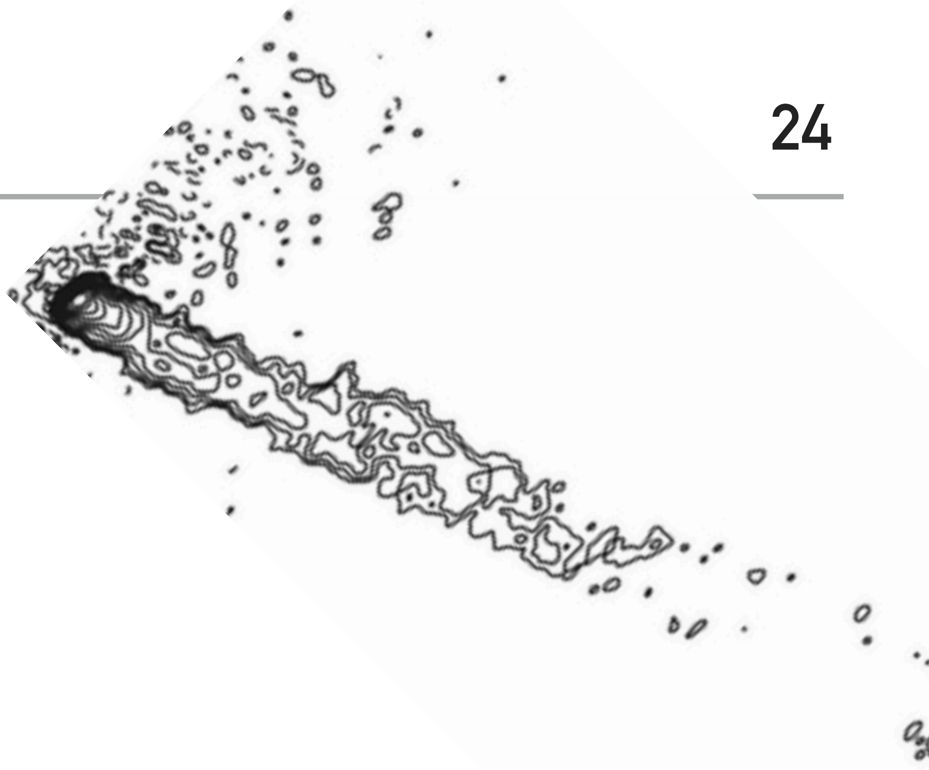


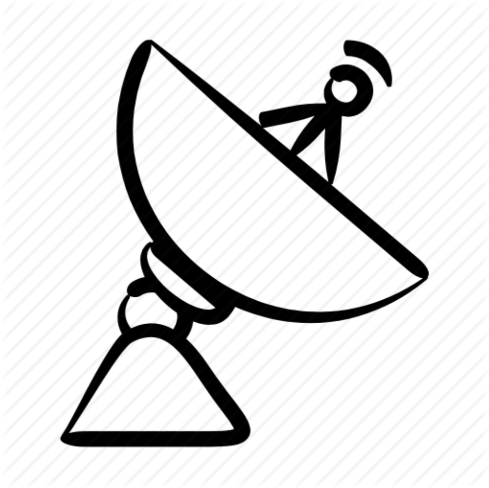


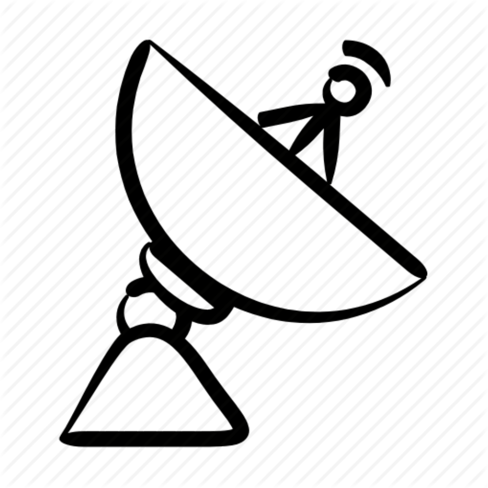
SELF-CALIBRATION PRINCIPLES

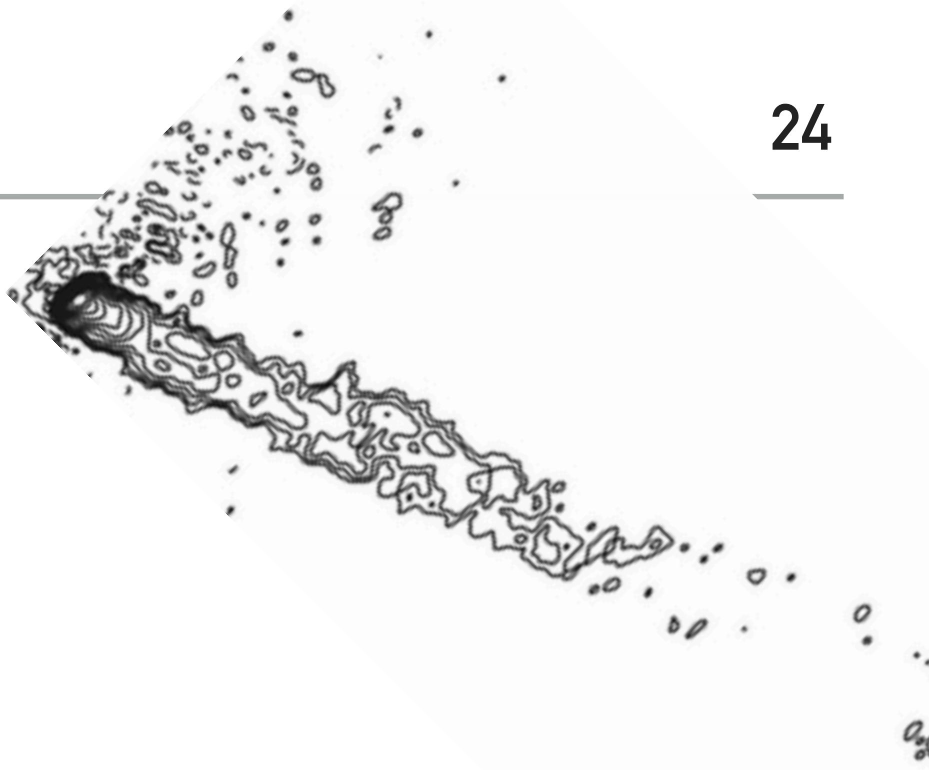
- Use *target* visibilities and allow the antenna gains to be free parameters.
- If all baselines correlated, there are N complex gain errors corrupting the $N(N-1)/2$ complex visibility measurements for a given time & frequency.
- Therefore there are $(N(N-1)/2) - N$ complex numbers that can be used to constrain the true sky brightness distribution.
- Even after adding the degrees of freedom from the antenna gains, the estimation of an adequate model of the target brightness is still overdetermined - need a model to constrain!
- The improved model created from constraining these parameters can then be used to constrain the visibilities & remove residual phase & amp errors

A LESS WORDY ANALOGY



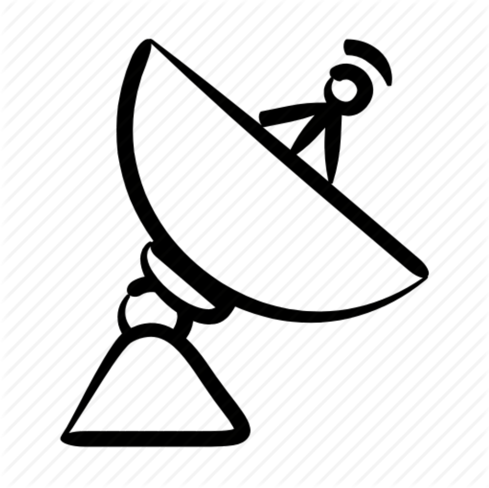


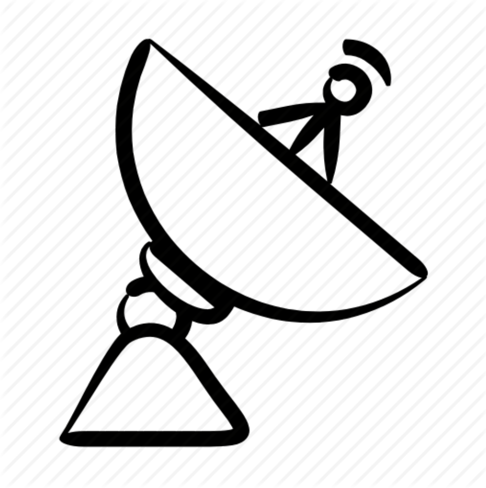


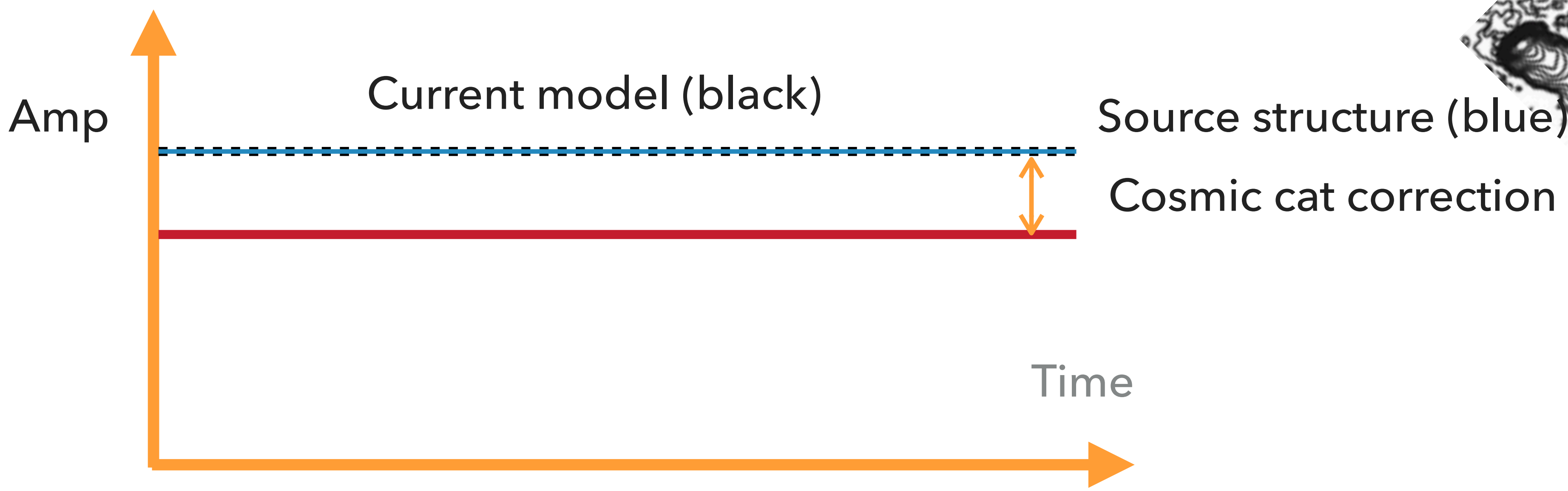


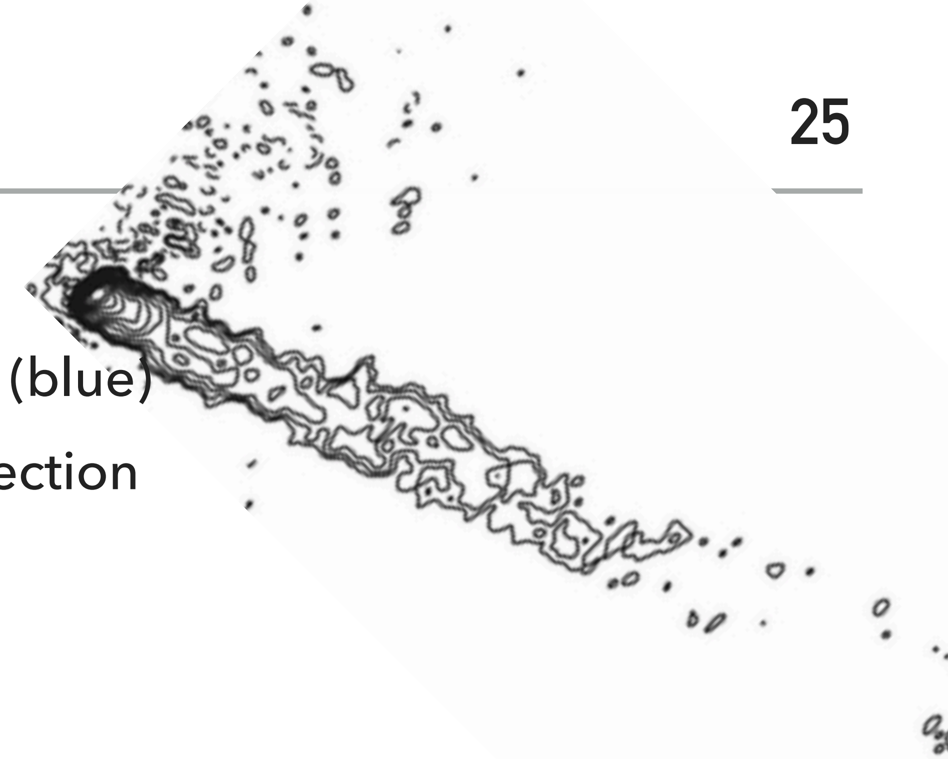
A LESS WORDY ANALOGY

Cosmic cat changes your amplitudes :(

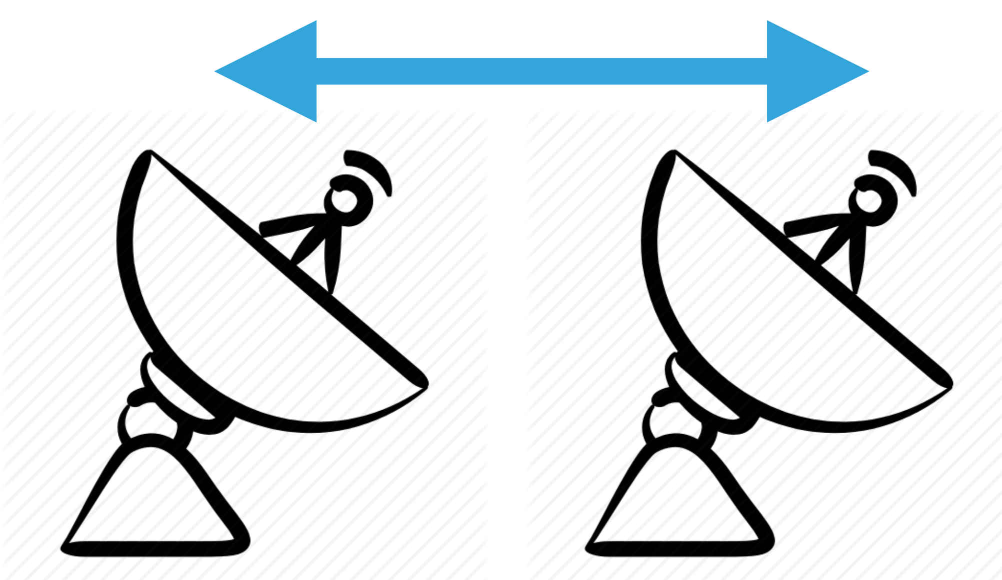


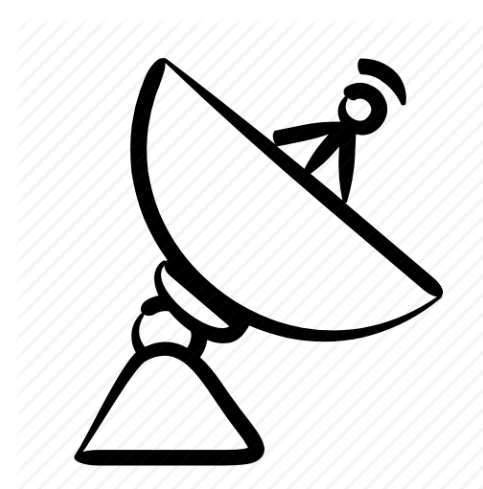


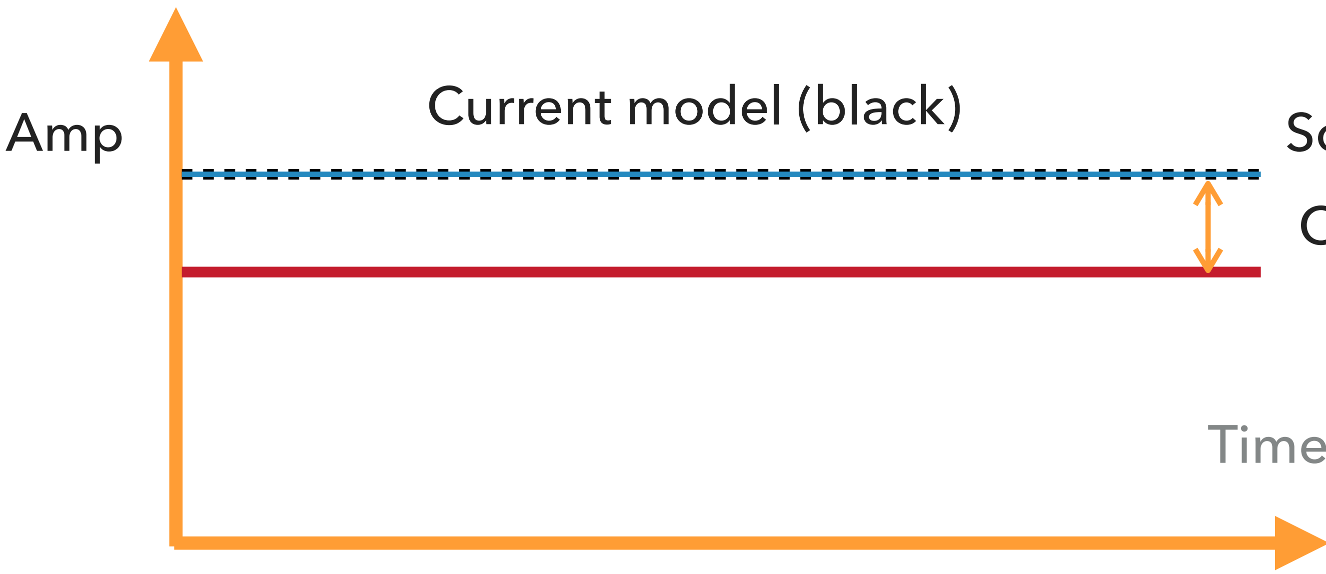


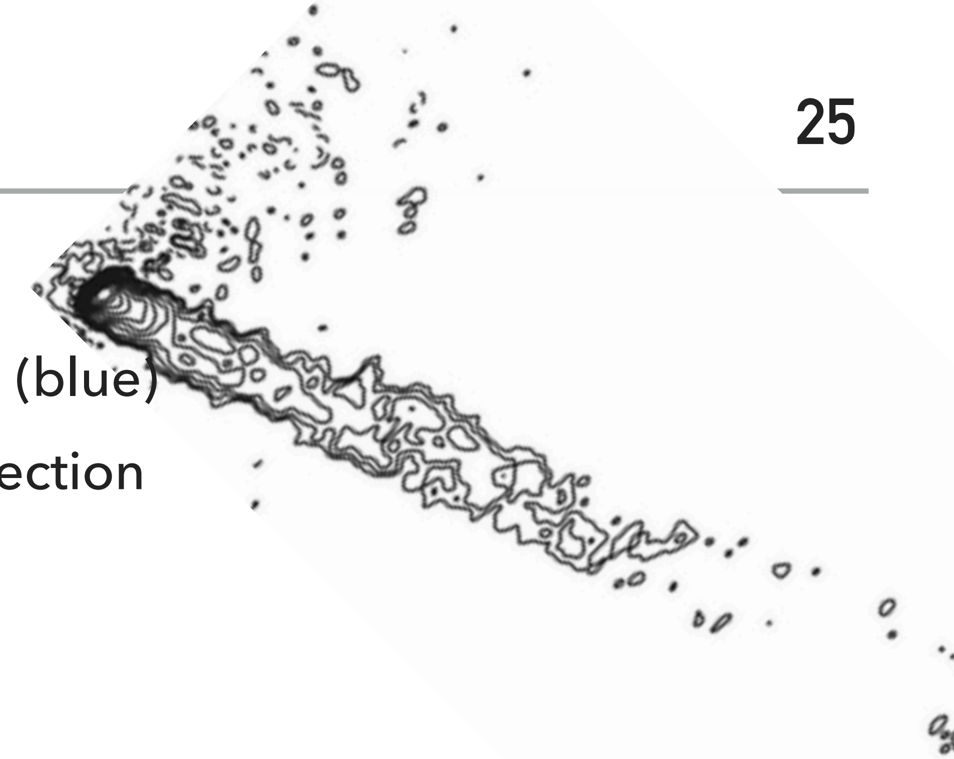


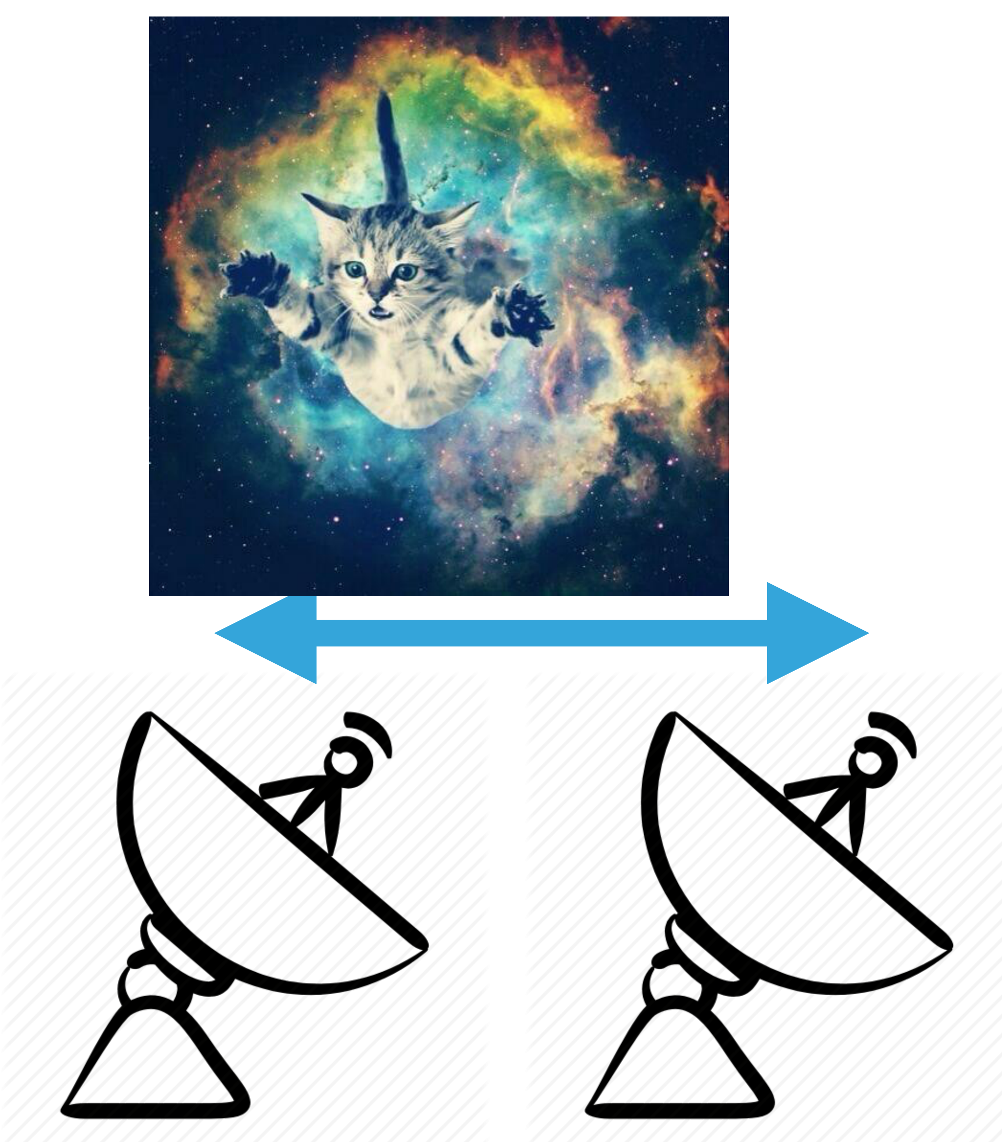
Cosmic cat
affects all
baselines to
antenna 2



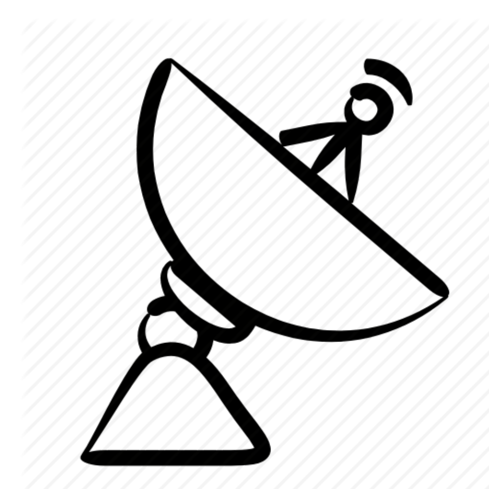


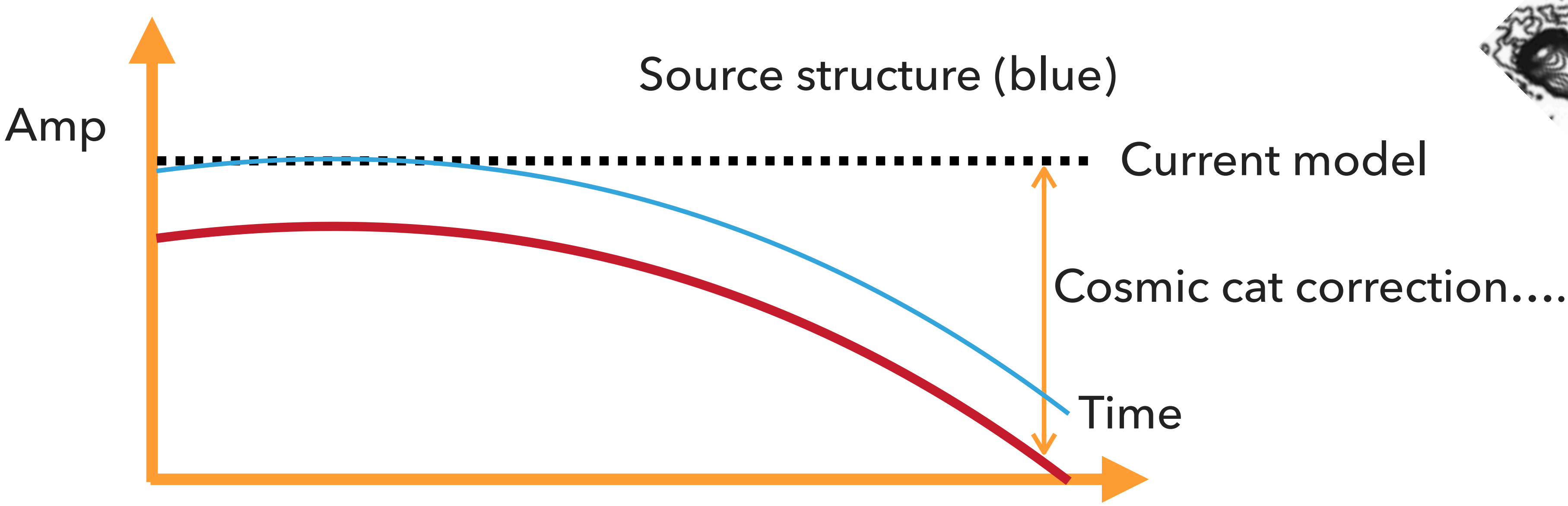


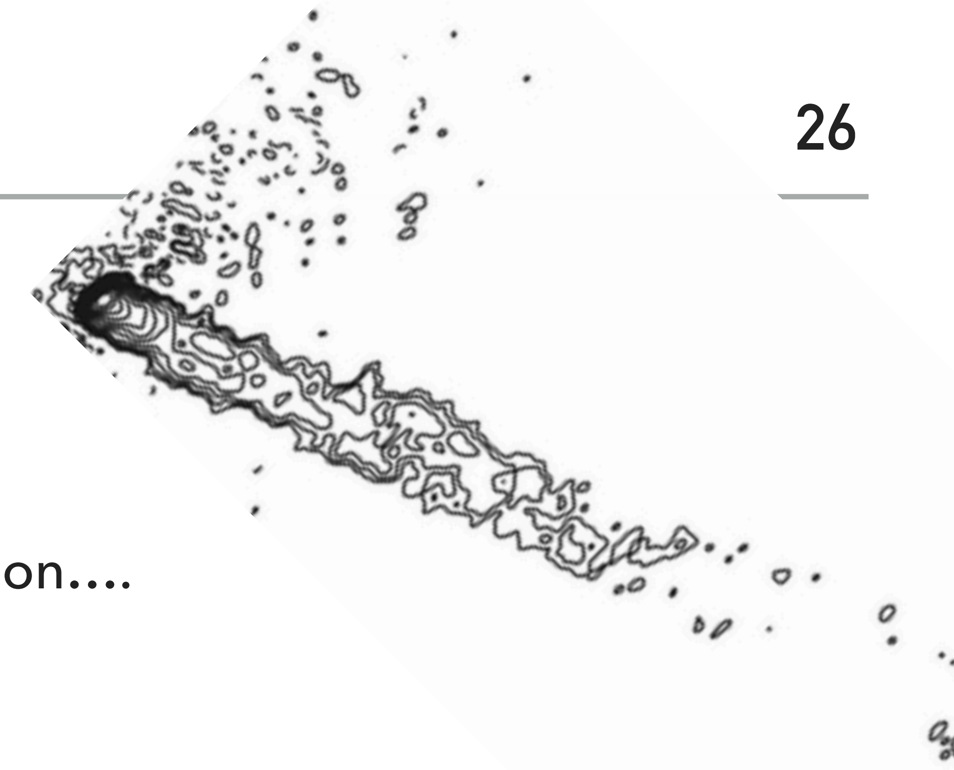




Cosmic cat
affects all
baselines to
antenna 2

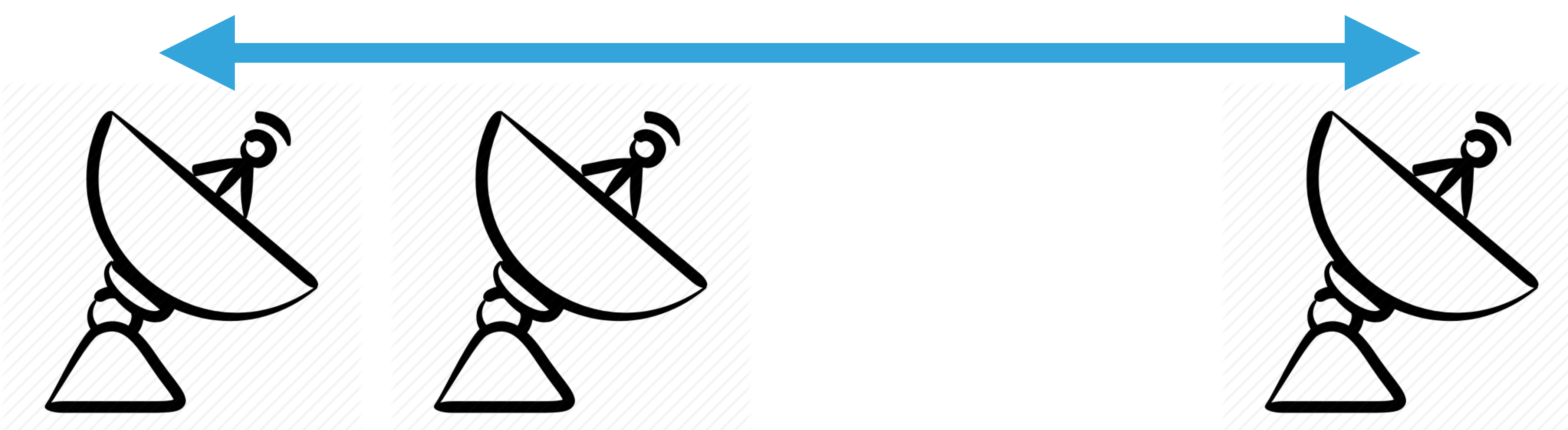




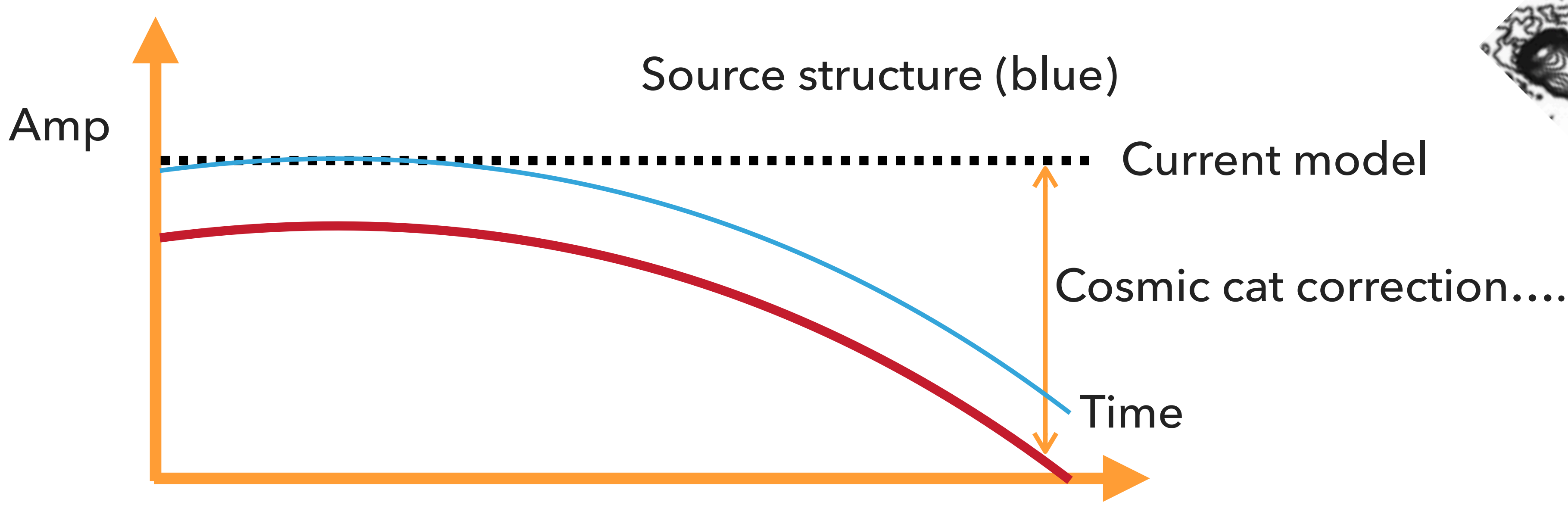


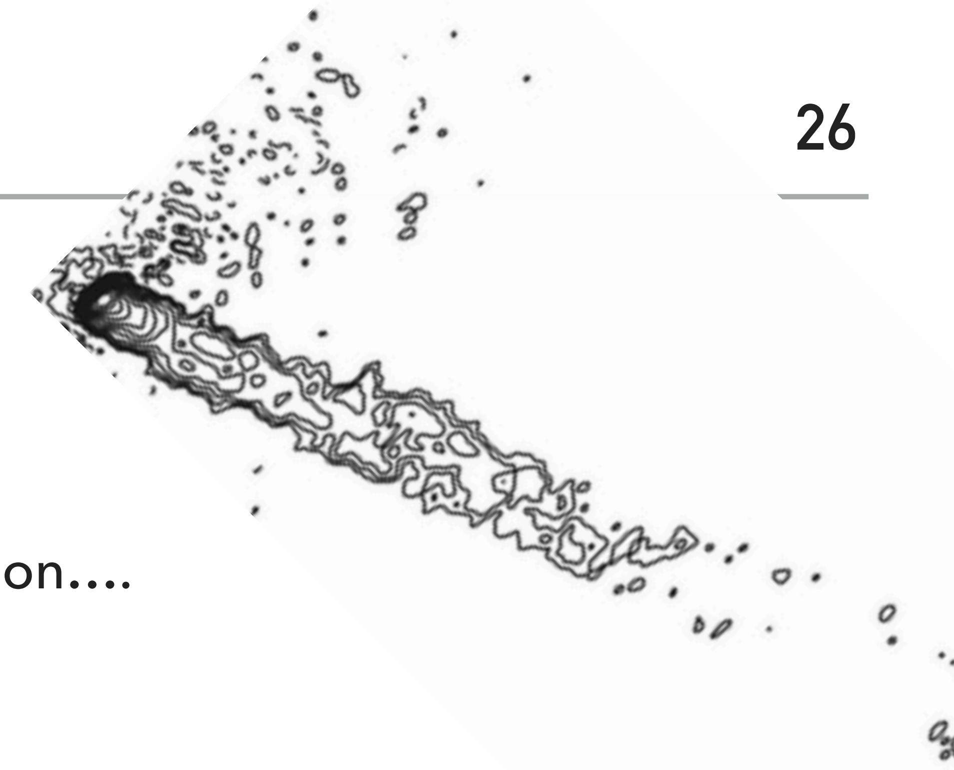
Correction factors

Get different values for different baselines to same antenna so solutions will FAIL!



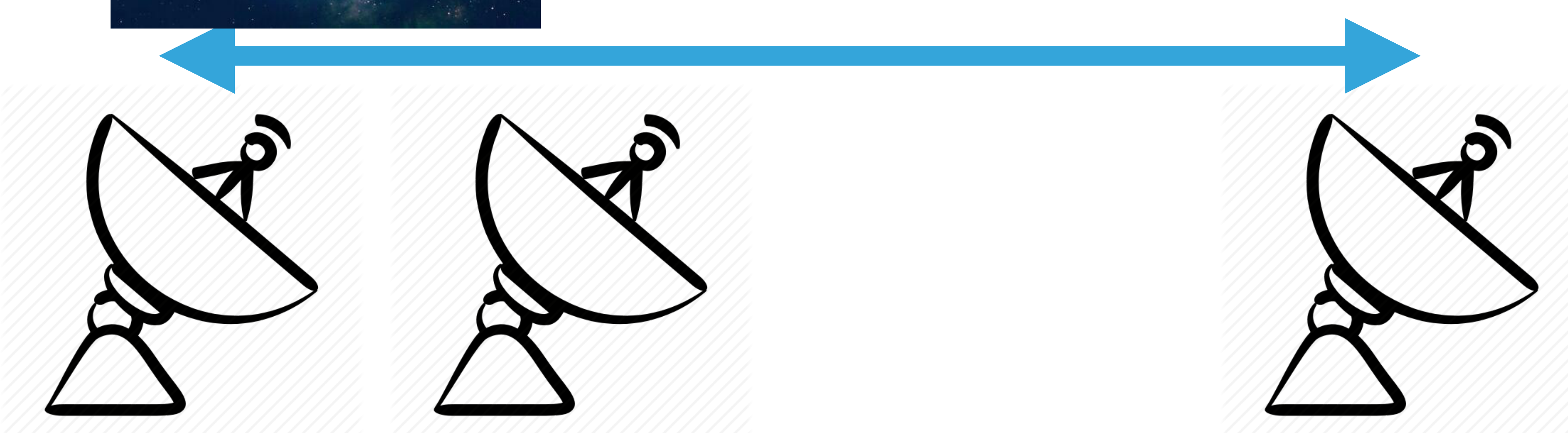
3. SELF CALIBRATION THEORY & PRINCIPLES

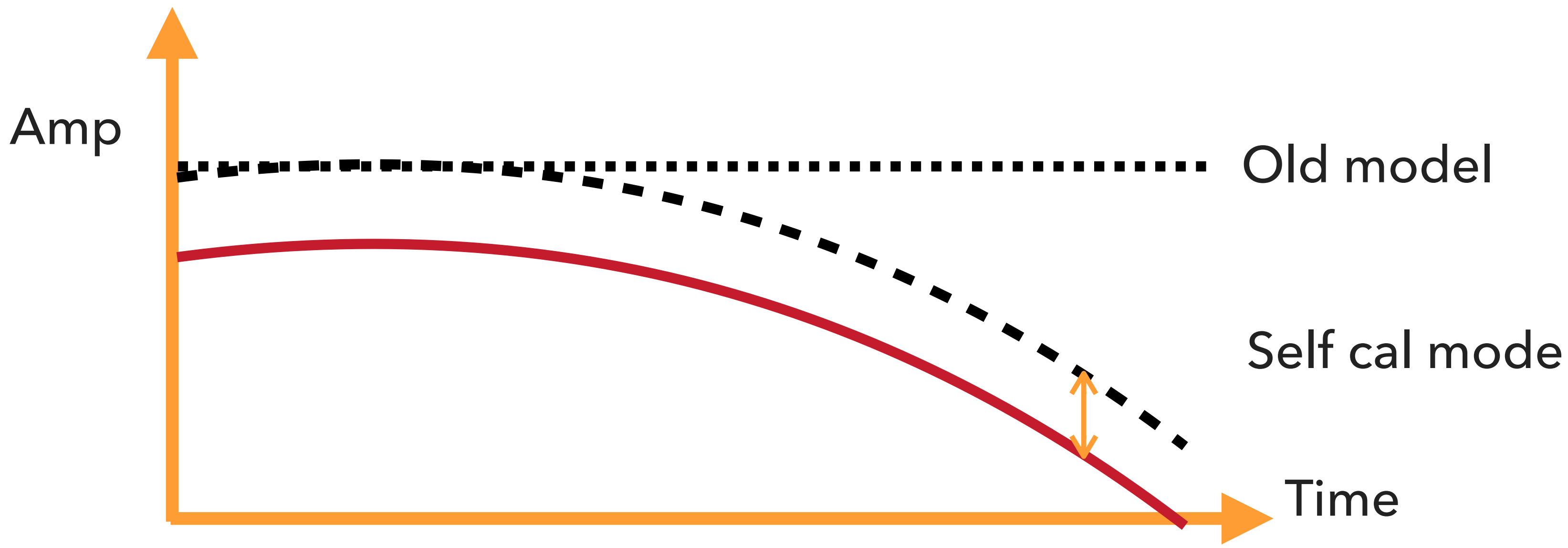


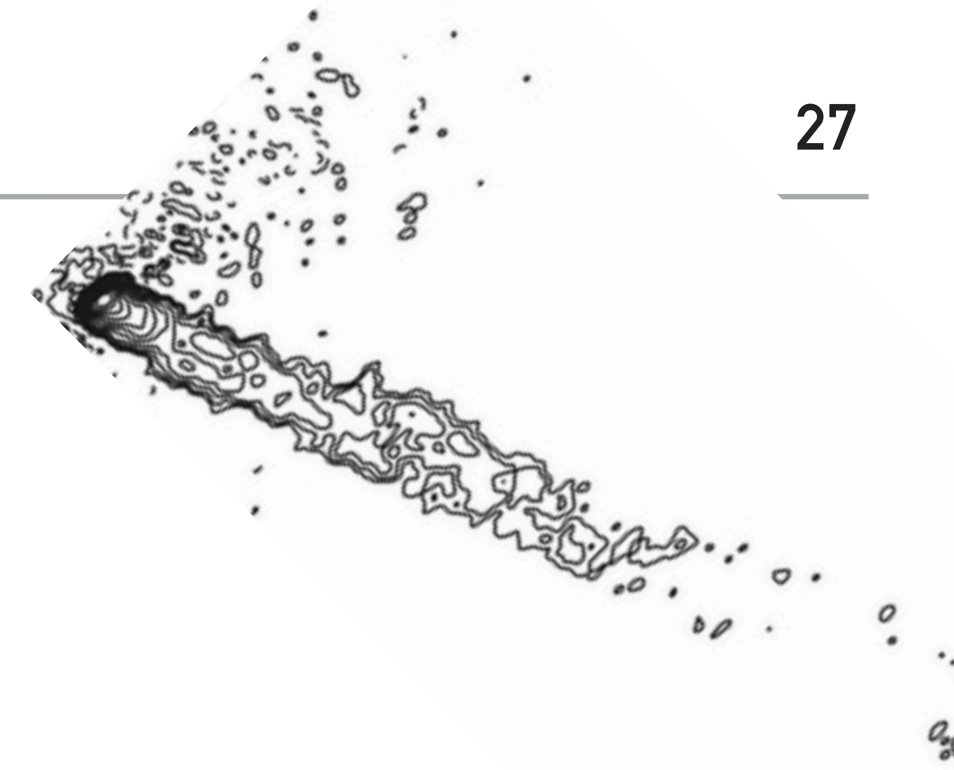


Correction factors

Get different values for different baselines to same antenna so solutions will FAIL!

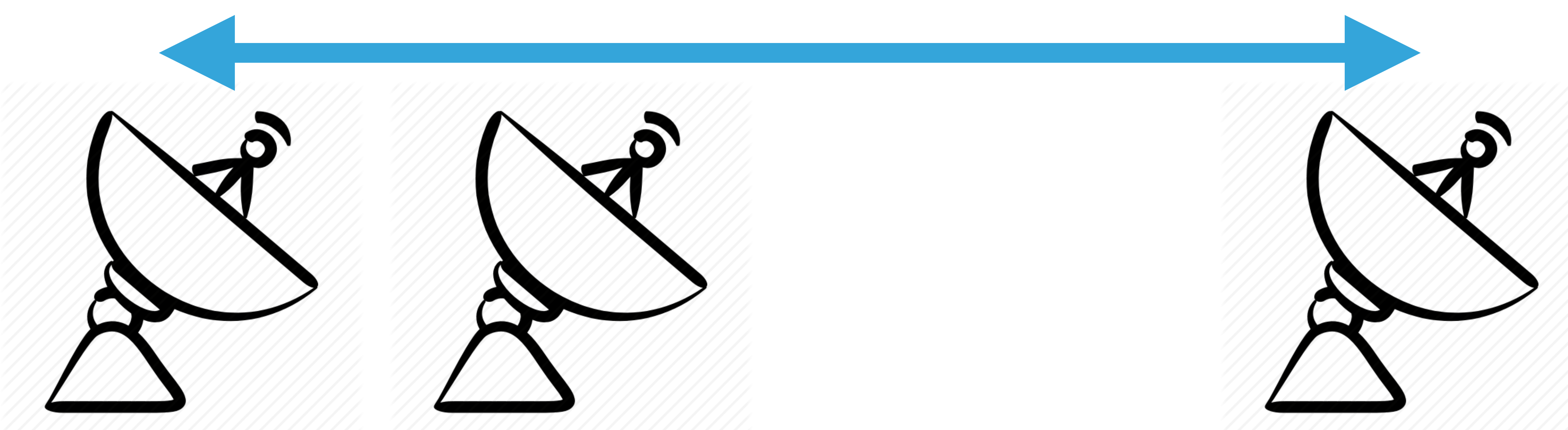


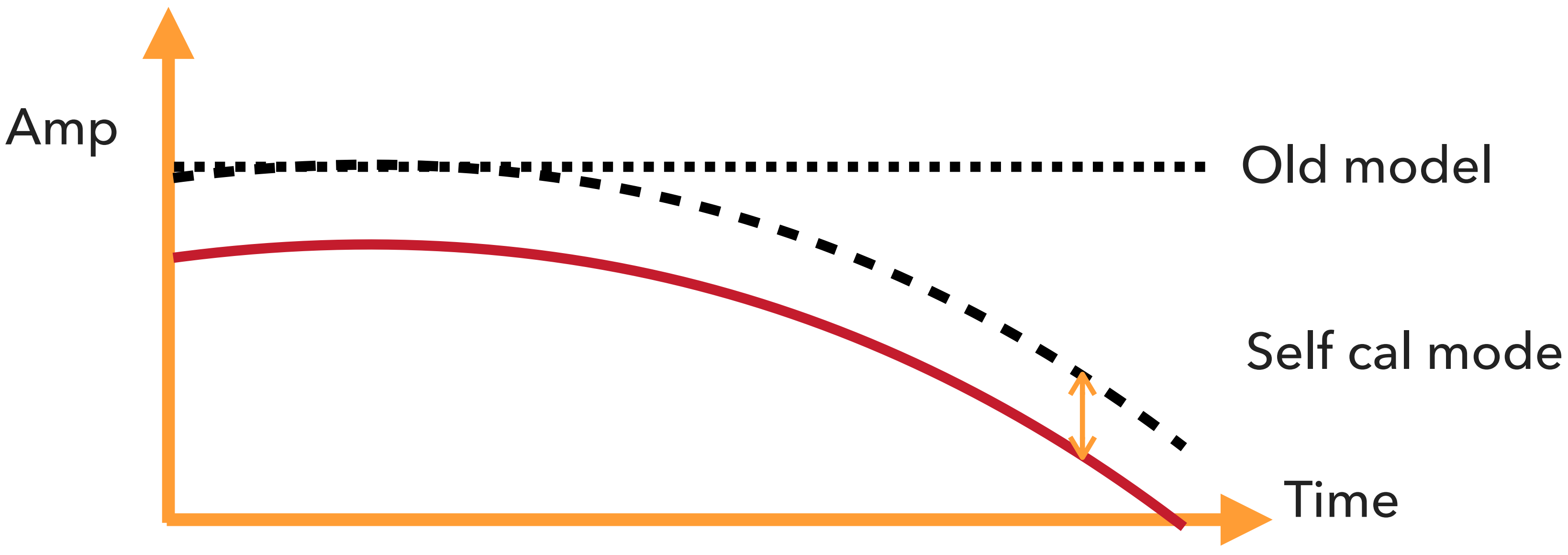


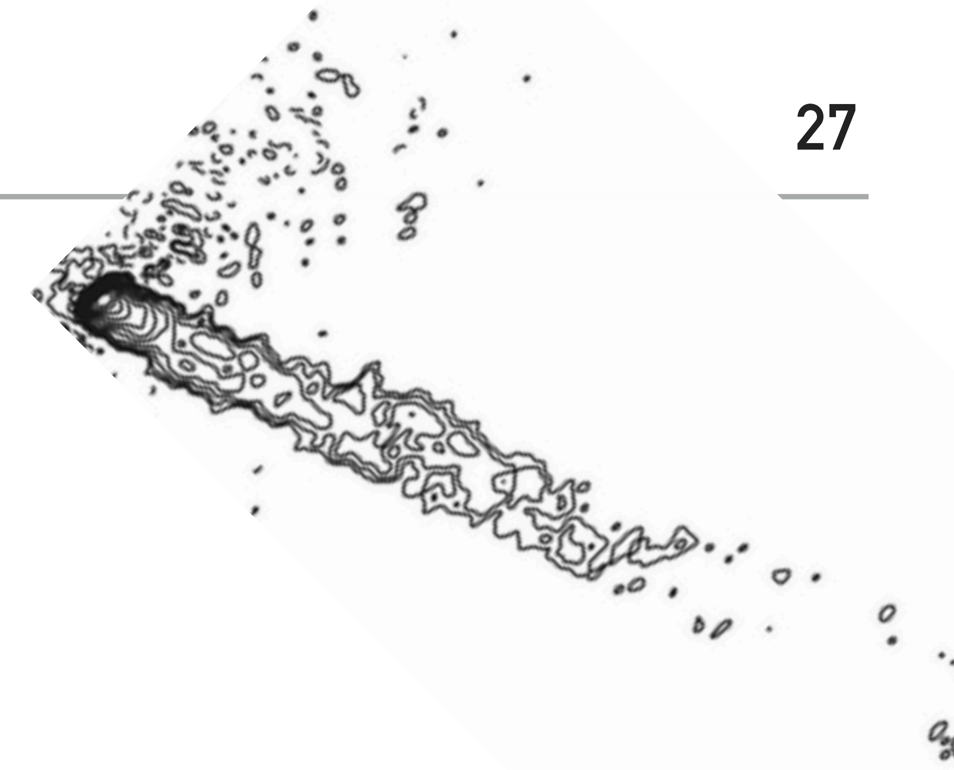


Correction factors

Correction factors same for both baselines to same antenna! Solutions will succeed

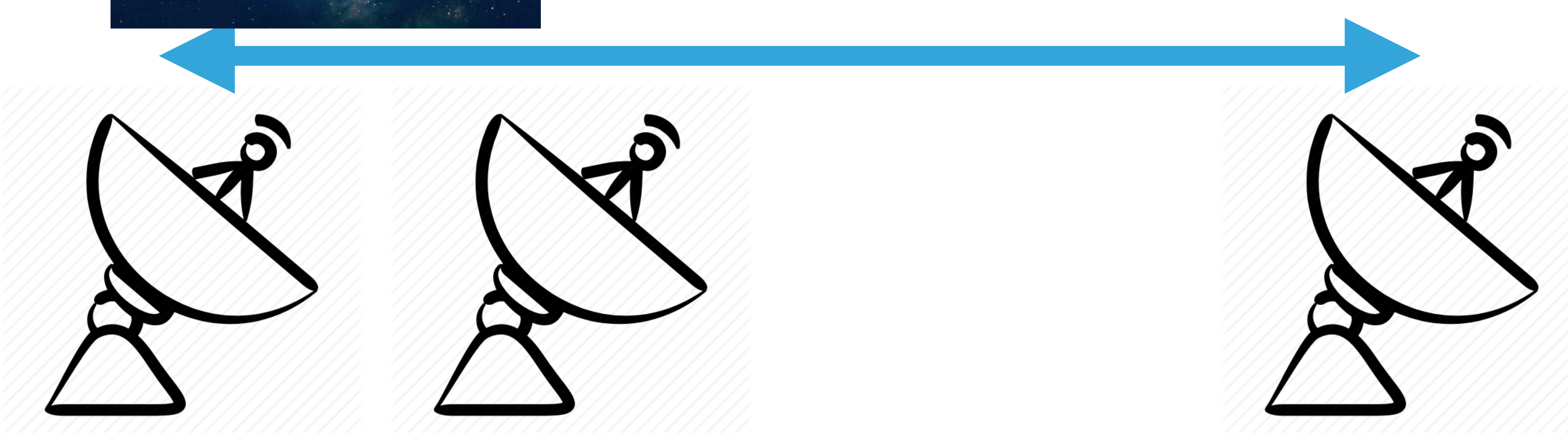






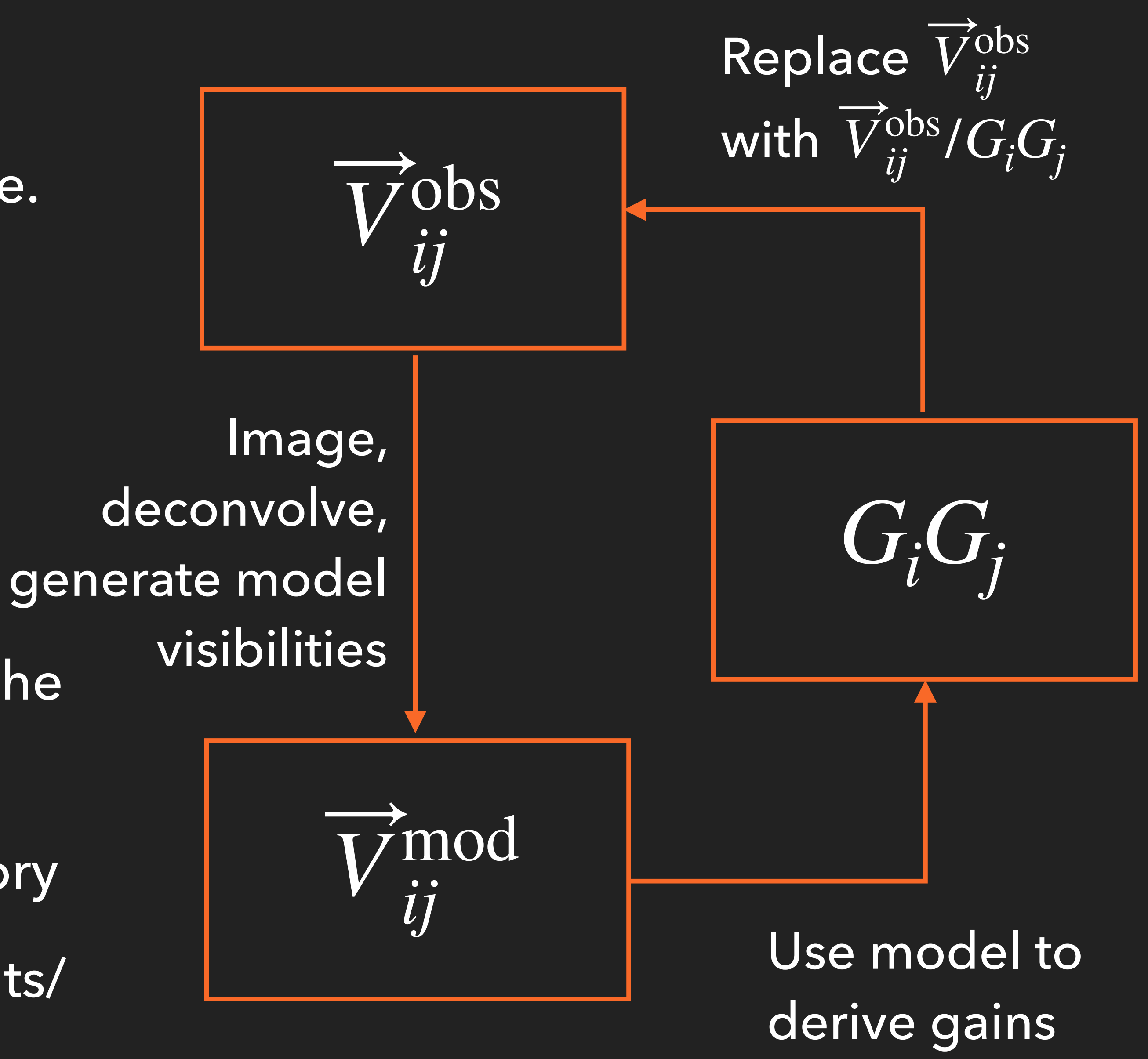
Correction factors

Correction factors same for both baselines to same antenna! Solutions will succeed



SELF-CALIBRATION METHODOLOGY

1. Create an initial source model, typically from an initial image (or else a point source)
   - Use full resolution information from the model image NOT the restored image (ie. CLEAN +residuals)
2. Find antenna gains
   - Using "least squares" fit to visibility data
3. Apply gains to correct the observed data
4. Create new model visibilities ($\vec{V}_{ij}^{\mathrm{mod}}$) from the corrected data
5. Go to (2), unless current model is satisfactory
   - shorter solution interval, different uv limits/ weighting
   - phase → amplitude & phase



SELF-CALIBRATION IN PRACTICE

1. Apply phase referencing corrections
2. Image target
   - Model image generated when imaging using tclean (in *.model file)
3. Fourier Transform model into measurement set using ft (or set tclean parameter savemodel='modelcolumn' when imaging) to create model visibilities ($\vec{V}_{ij}^{\text{mod}}$)
4. *(optional)* Compare model with visibilities (using plotms). Differences due to either/both:
   - Deficiencies in the model
   - Atmospheric or other errors affecting visibilities

SELF-CALIBRATION IN PRACTICE

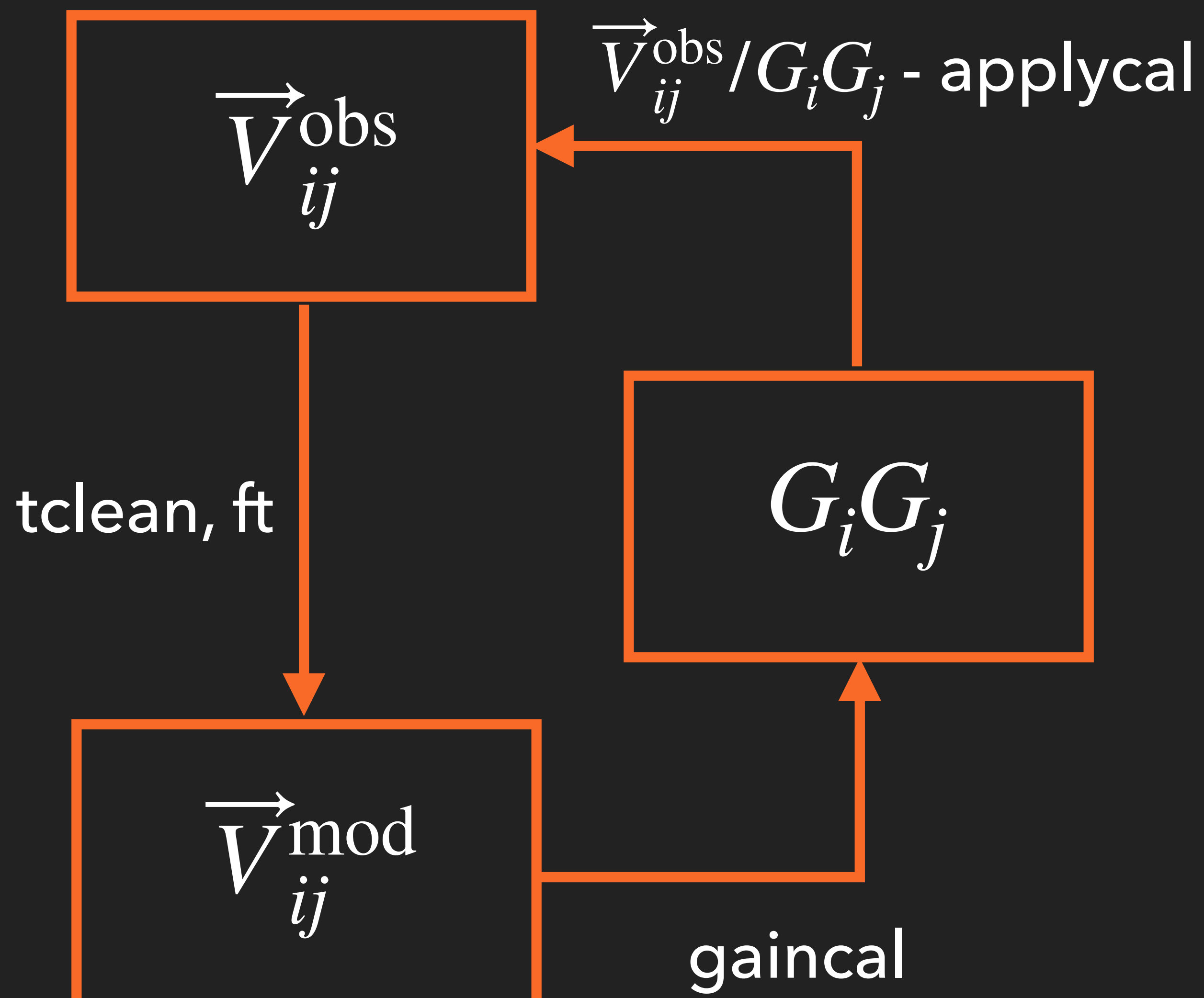
5. Estimate corrections for phase visibility errors
   - Use CASA task gaincal with calmode='p'.
6. Apply and make another image
   - This should provide an improved model
7. Repeat until phase corrections converge then do an amplitude & phase self calibration (calmode='ap').

Remember:

- Phase correcting first allows longer averaging times for amplitudes
- Apply your phase solutions too when using calmode='ap' as residual phase solutions should be small.

SELF-CALIBRATION IN PRACTICE

- The CASA cycle

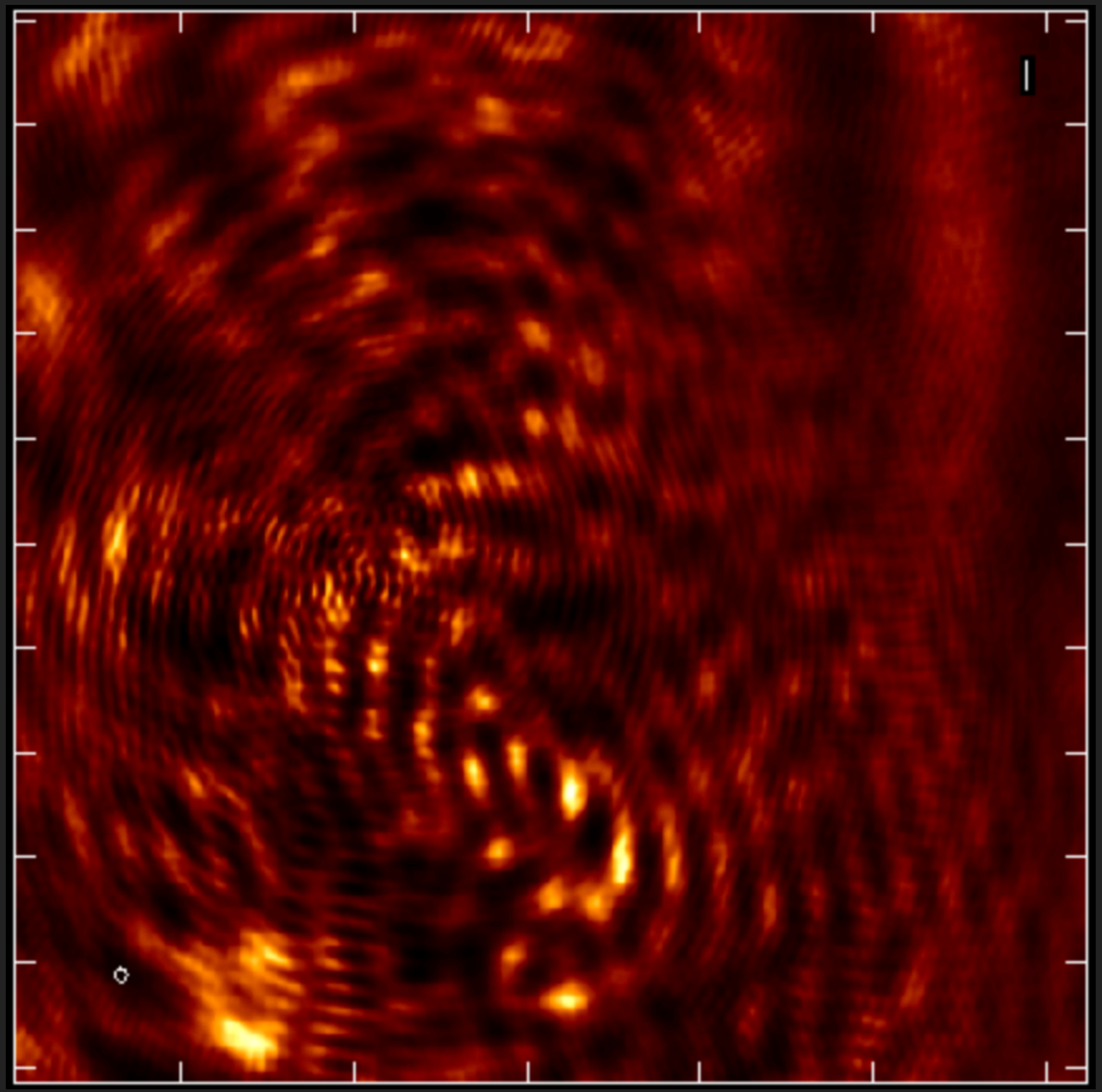


3C277.1 SELF-CALIBRATION WORKSHOP

- The next few slides will outline the self-calibration steps for 3C277.1 that we will do in the workshop.
- The slides will explain the first self-calibration cycle and then you shall continue the process!

TARGET WITHOUT PHASE REFERENCING

First image of the target



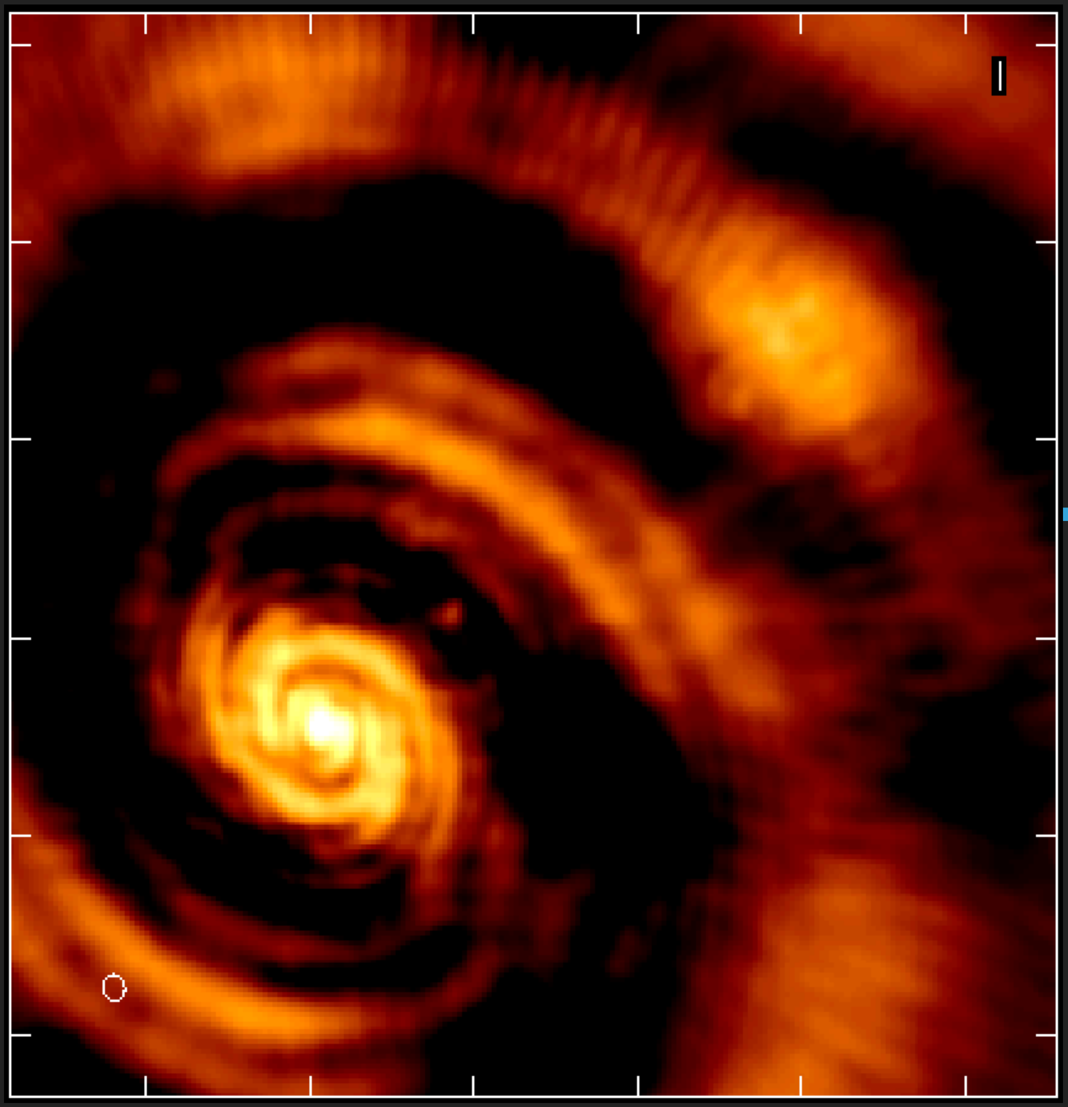
Dirty map

- Raw phase (not corrected)
- Holes and smearing
- No peaks

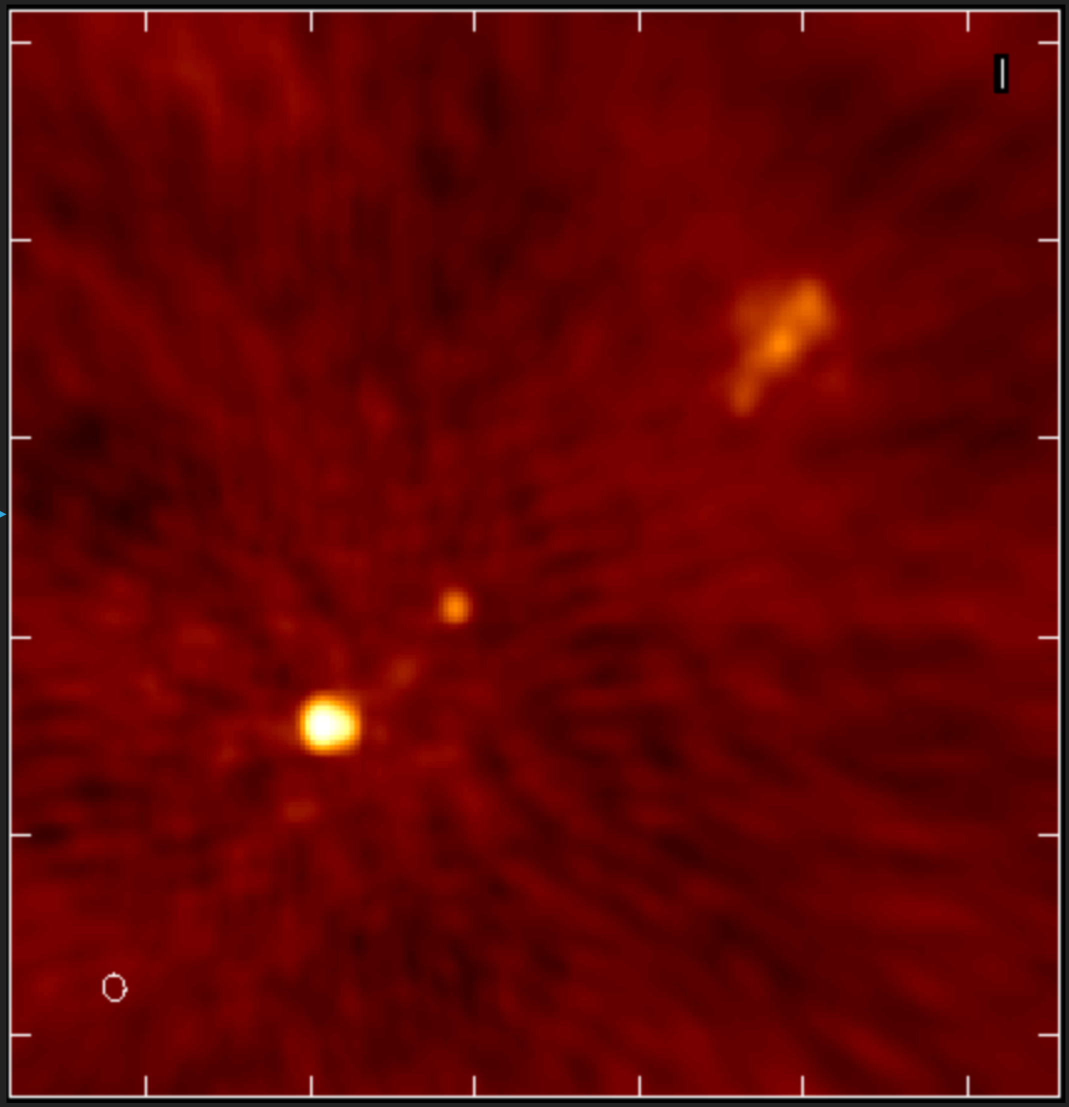
Apply phase referencing corrections to visibilities and Fourier Transform to image plane to get….

PHASE-REFERENCED TARGET IMAGE

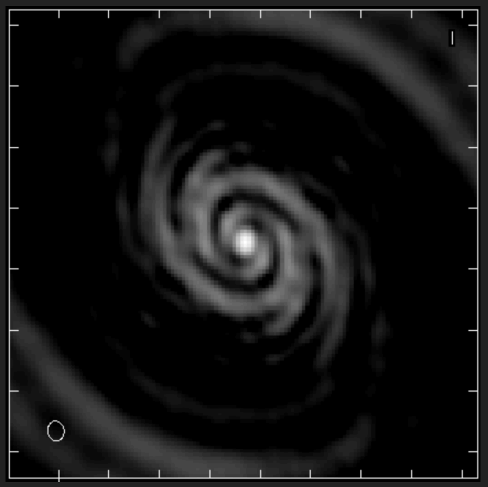
Dirty image



First CLEAN image



Deconvolve iteratively

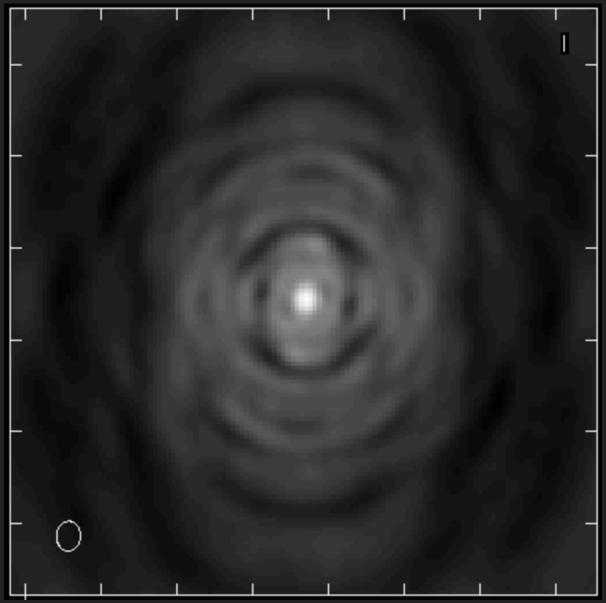


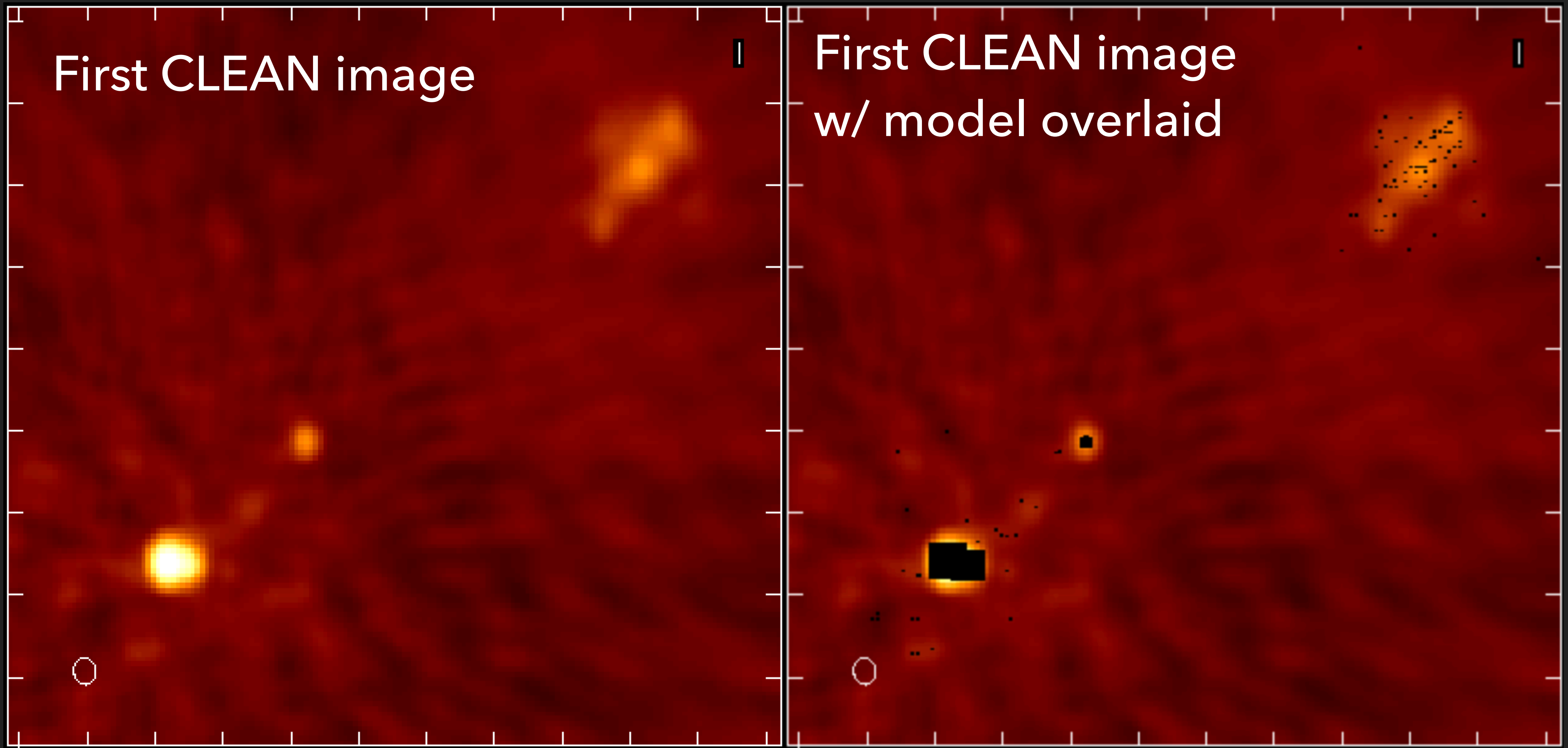
with the dirty beam

MODEL GENERATION

Take care in setting mask (clean boxes)!

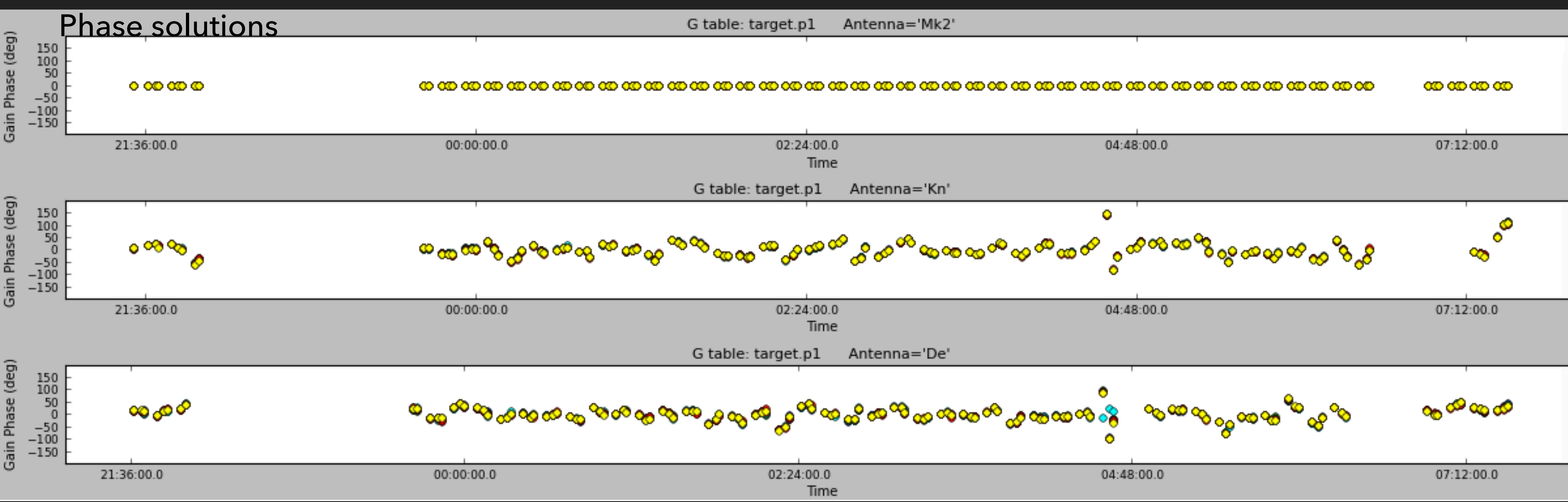
- FT of CLEAN components / model image will be used as model
- Plot dirty beam to help avoid cleaning sidelobes





PHASE SELF-CALIBRATION

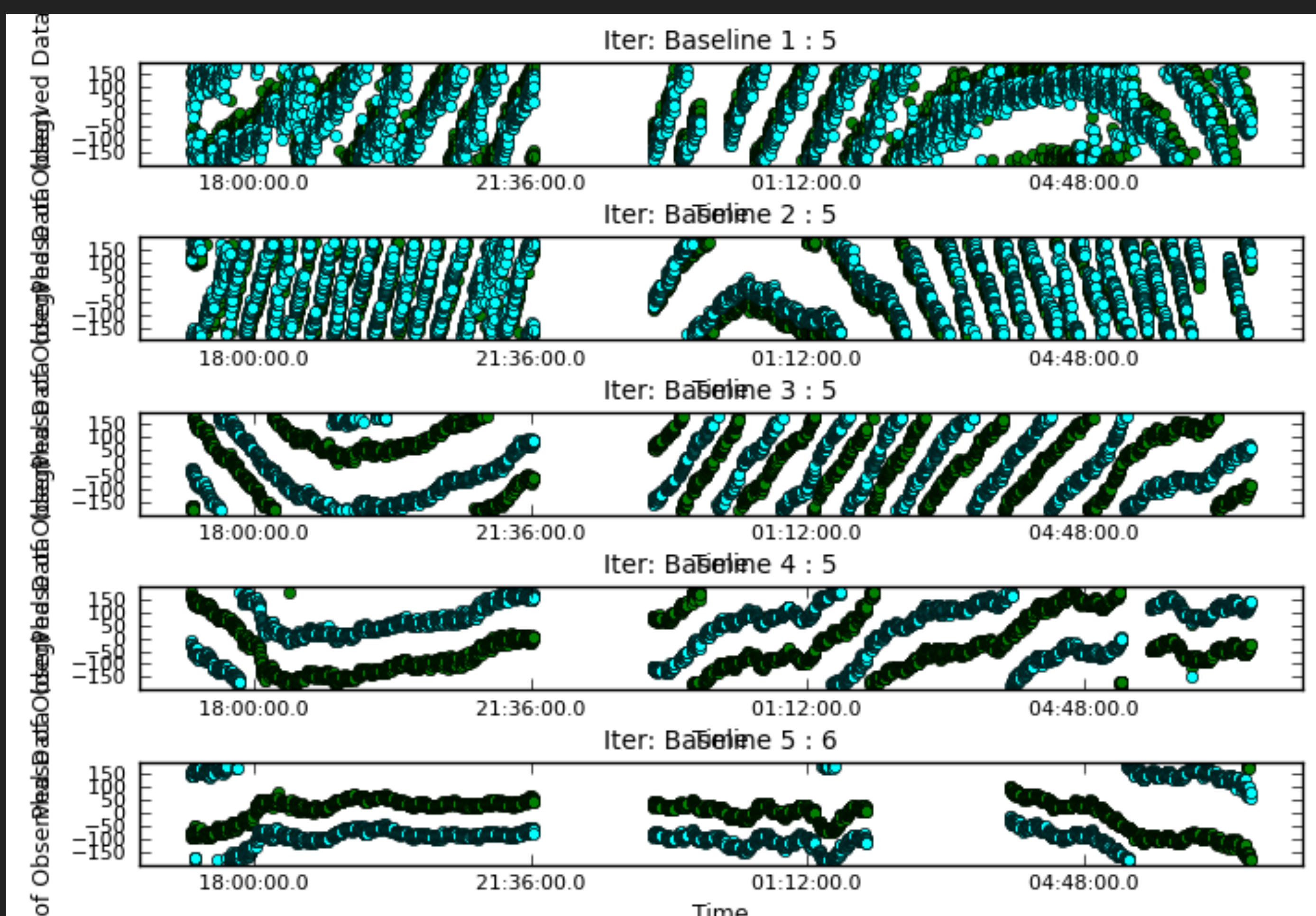
- FT of model image are stored in MS 'model' column if savemodel='modelcolumn' in CASA **tclean** is set.
- Alternatively use task **ft** to convert .model file into model visibilities!
- Use **gaincal** to obtain phase solutions contained in a calibration file - **gaincal** compares model with target visibility data



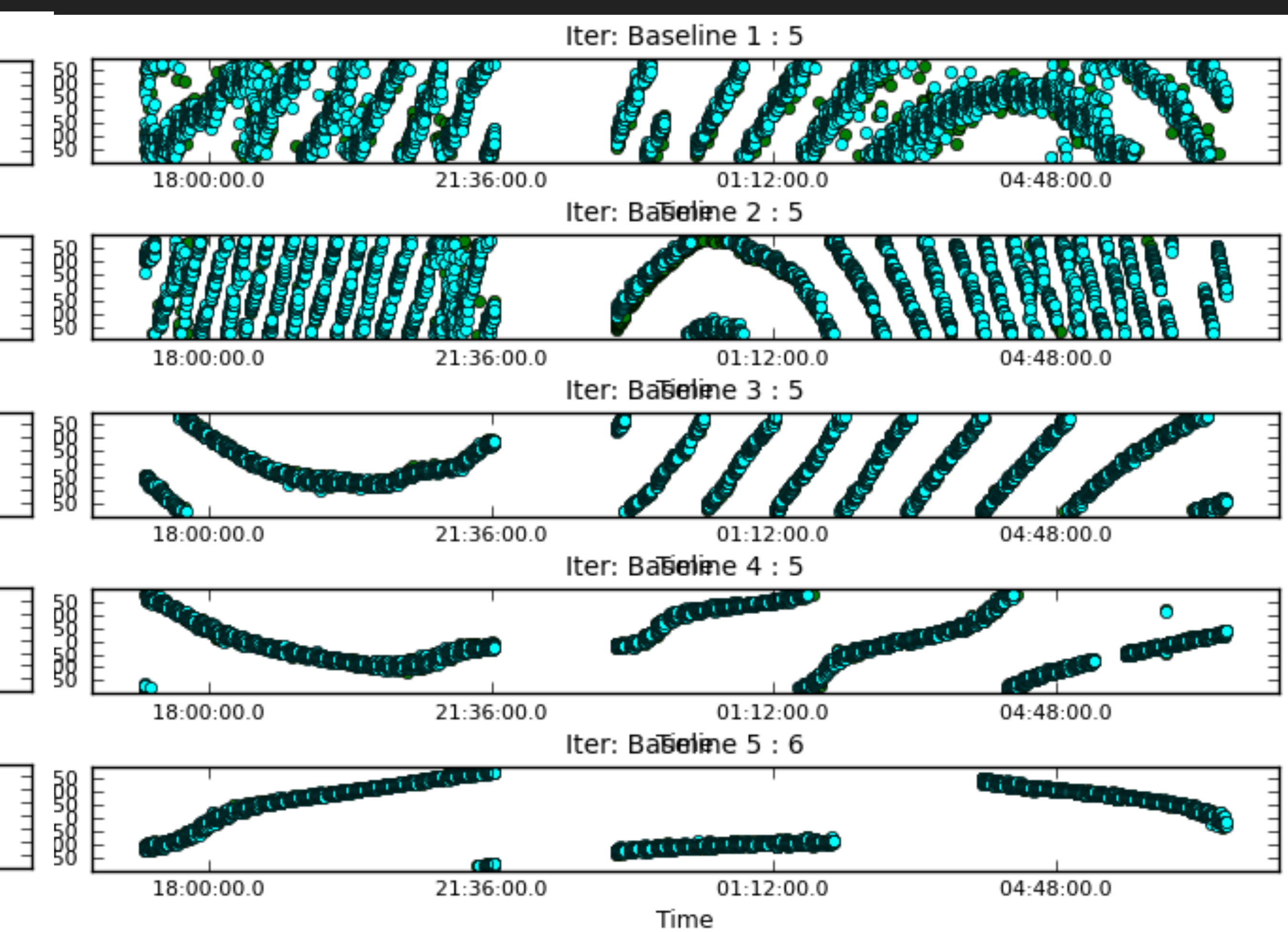
PHASE SELF-CALIBRATION

- Use CASA task **applycal** to apply calibration file (containing self-cal solutions) to the MS
- Much improved and smoother target phases - reduced scatter!
- **This is how far the data you have has got to.**

BEFORE SELF-CAL



AFTER SELF-CAL

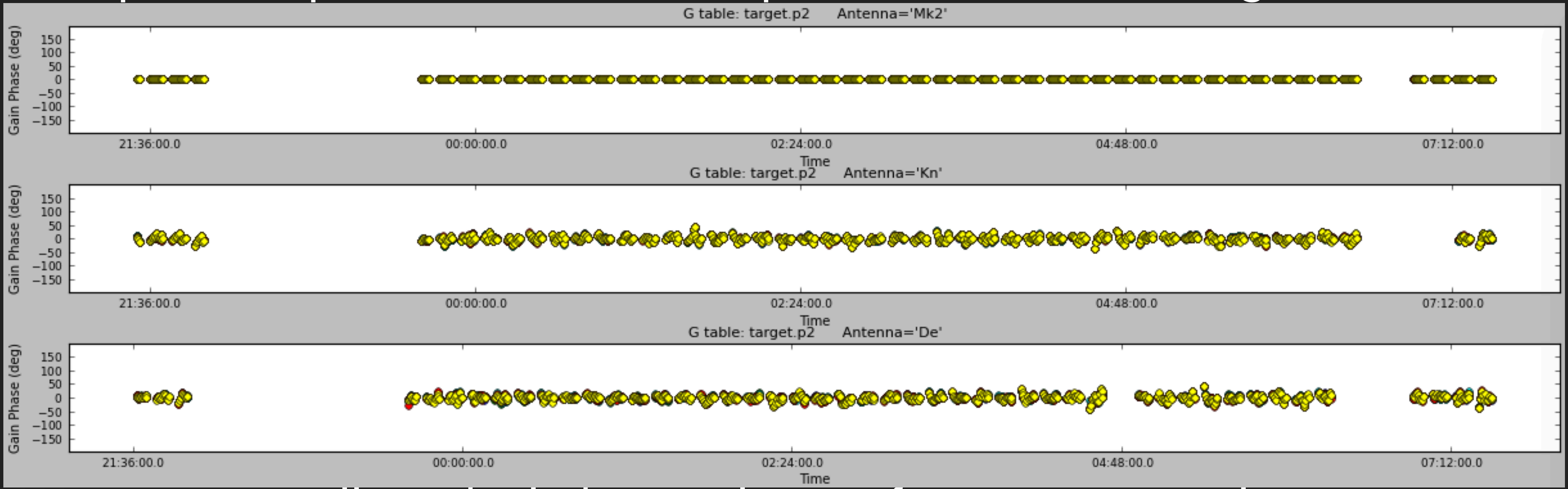


So what do we do next!?

ITERATIVE SELF-CALIBRATION

- We are going to do the following in this workshop…
- Solutions have been applied so we image again!

  Note that:
  - The old model column is overwritten!
  - Use task **ft** (with previous model image) if you want to go back to a previous CLEAN model.
- Repeat this process until residual phase solutions converge on zero.



We expect small residual phase solutions from next gaincal

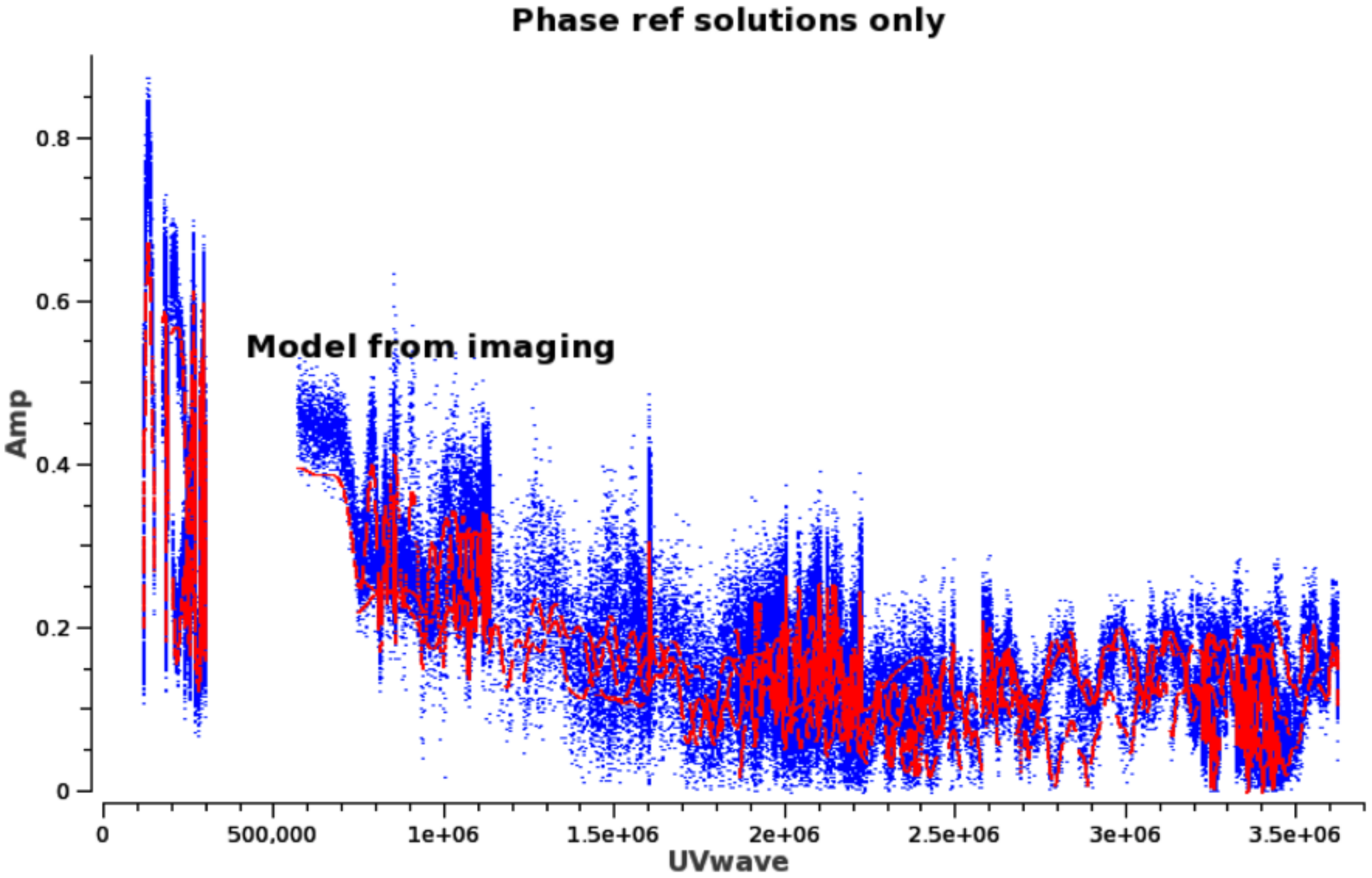
ITERATIVE SELF-CALIBRATION

- A good iterative self-cal cycle to follow is outlined below. The phase calibration increases the model S/N enough for amp cal in future iterations.
- You should aim to change parameters e.g. solution interval at each round to generate the improvements
- In CASA, you should apply your previous self-cal solutions when deriving new calibration tables

First image
↓
Self-cal phases
↓
Shallow CLEAN
(to avoid CLEAN bias)
↓
Self-cal. phase
↓
Deep CLEAN
↓
Self-cal. amplitude + phase
(amp. solutions should be near 1!)
↓
Deep CLEAN

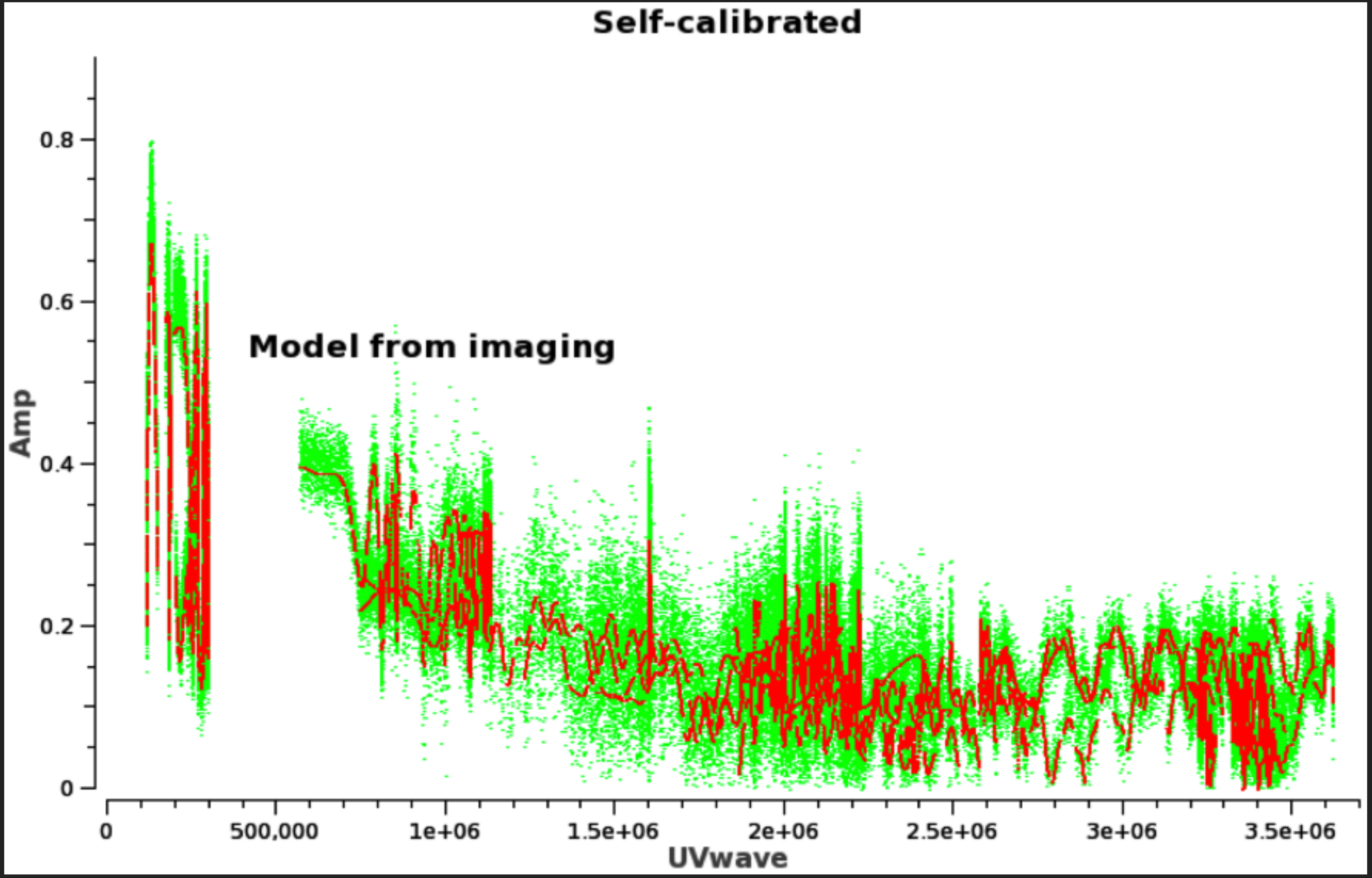
SELF-CALIBRATION IMPROVEMENTS

- And your visibilities change from this… (phase referencing only)



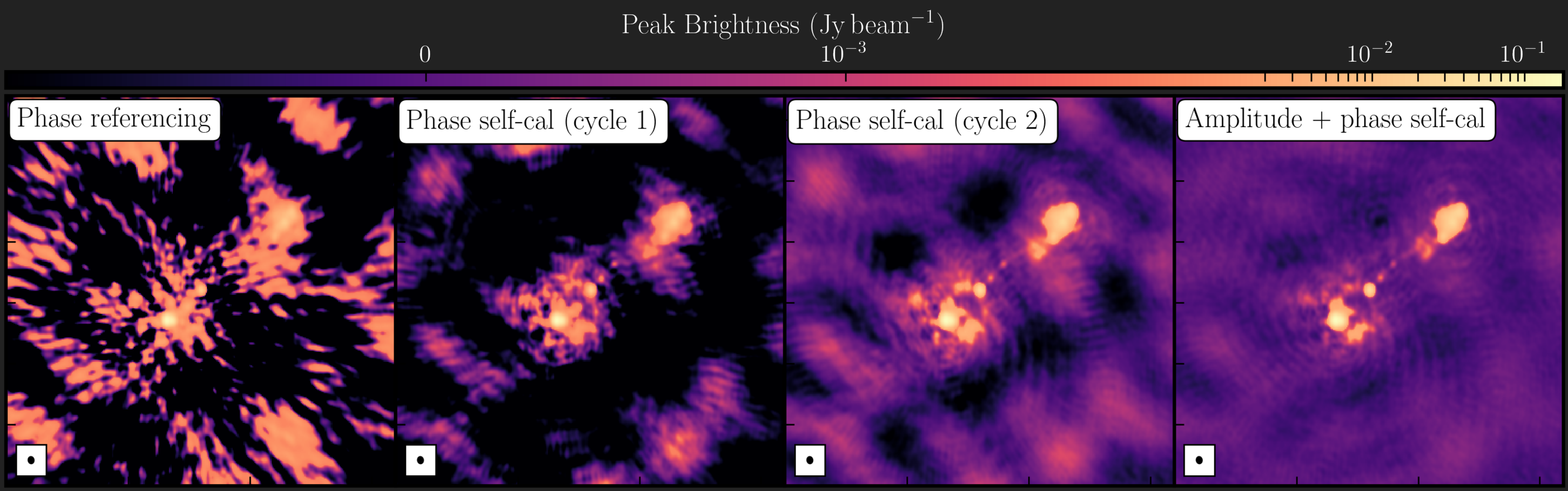
SELF-CALIBRATION IMPROVEMENTS

- To this… (phase referencing + self calibration)



ITERATIVE SELF-CALIBRATION

After rounds of self-calibration, you get a vastly improved image with increased signal to noise:



CHOICES IN SELF-CALIBRATION

1. *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 initially!)
    - Simple model-fitting in uv plane

2. *Self-calibrate phases or amplitudes?*
    - Usually phases first
    - Phase errors cause anti-symmetric structures in images
    - e.g. For VLA and VLBA, amplitude errors tend to be relatively unimportant at dynamic ranges < 1000 or so
    - Safe bet is to follow iterative self calibration cycle shown on previous slide.

CHOICES IN SELF-CALIBRATION

3. *Which baselines?*
   - For a simple source, all baselines can be used
   - For a complex source, with structure on various scales, start with a model that includes the most compact components, and use only the longer baselines

4. *What solution interval should be used?*
   - Generally speaking, use the shortest solution interval that gives 'sufficient' signal/noise ratio (SNR)
   - If solution interval is too long, data will lose coherence - solutions will not track the atmosphere optimally.

GETTING ENOUGH SNR FOR SELF-CALIBRATION

- Can self-calibrate if signal to noise ratio (SNR) on most baselines is greater than one.
- For a point source, the error in the gain solution, σ_g is:

$$\text{Phase only} \quad \sigma_g \approx \frac{1}{\sqrt{N-2}}\frac{\sigma_V}{S}$$

$$\text{Amp+phase} \quad \sigma_g \approx \frac{1}{\sqrt{N-3}}\frac{\sigma_V}{S}$$

S = source flux density, σ_V = noise per visibility sample, N = # antennas

- This is a good measure of whether you can self-cal at all and what solution interval should be used.
- If gain error is much less than 1, then the noise in the final image can be close to theoretical
- If you are desperate, you can average polarisations to get enough SNR or longer solution interval

WE ❤️ RADIONET

RadioNet has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730562

TO 3C277.1 SELF-CALIBRATION

- Good luck!