Radio interferometric imaging in astronomy

Tuomas Savolainen
Aalto University

Image credit: NRAO

Outline

- Aperture synthesis – visibility sampling
- Solving the inverse problem
- Image quality and error recognition

Aalto University
School of Electrical Engineering

Aperture synthesis

Recap: Visibility function $\leftarrow\mathcal{F}\rightarrow$ sky brightness distribution

The relationship between the visibility function $\mathcal{V}(u,v)$ and the sky brightness distribution $I(l,m)$ is a 2-D Fourier transform (assuming a small field of view and neglecting the w-coordinate):

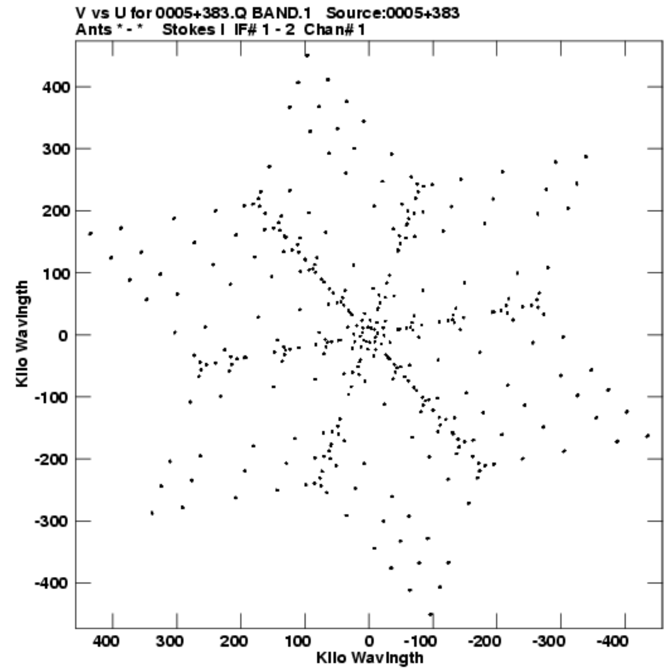
$$\mathcal{V}(u,v) = \iint I(l,m)\, e^{-i2\pi(ul+vm)} dl dm$$

Let's assume that you have a well-calibrated visibility data set. The next step is to turn this data set into an image.

Aperture synthesis

- In principle, inverting $\mathcal{V}(u,v) = \iint I(l,m)\, e^{-i2\pi(ul+vm)} dl dm$ gives the sky brightness distribution. This however requires measuring $\mathcal{V}(u,v)$ everywhere in the (*u*,*v*) plane. Not possible!
- In reality, we aim to sample $\mathcal{V}(u,v)$ sufficiently well in order to constrain $I(l,m)$. What is suffiently well? Well, that is a complicated question… In any case *"(u,v)* coverage" is one of the main decisive factors between a high quality image and rubbish.
- To do well, we want:
  - Many telescopes, since the number of instantaneous (*u*,*v*) samples is N(N-1), where N is the number of telescopes
  - Long synthesis time for changing baseline projections as Earth rotates. However, be careful if the source is variable!

Examples of (*u*,*v*) plane sampling

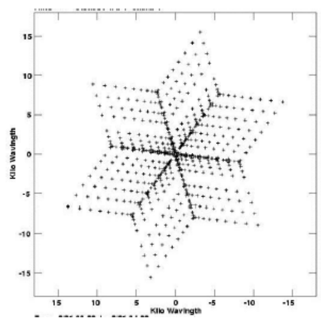


Visibility sampling for a VLA snapshot

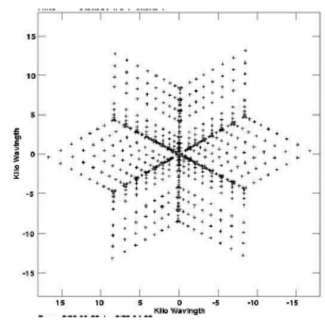
Examples of (*u*,*v*) plane sampling

Sample VLA (U,V) plots for 3C147 (δ = 50)

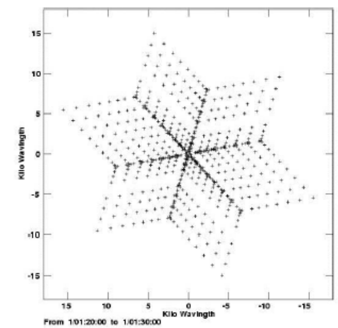
- Snapshot (u,v) coverage for HA = -2, 0, +2 (with 26 antennas).



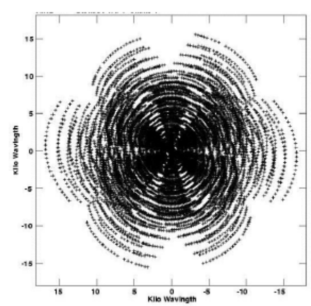
HA = -2h



HA = 0h



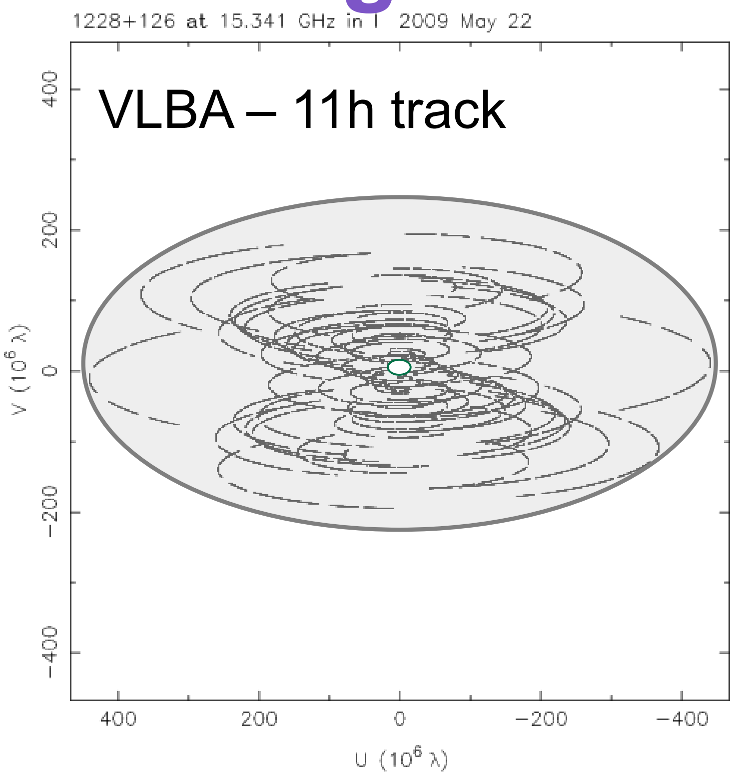
HA = 2h



Coverage over all four hours.

Image credit: Rick Perley

What does (*u,v*) coverage mean to your image?



- Outer boundary limits the angular resolution
- Inner boundary limits the sensitivity to large-scale emission structure
- Imperfect sampling in-between limits the image fidelity – there is information missing!

Formal description of a discrete sampling of the (u,v) plane

Visibility plane is sampled at discrete points given by sampling function:

$$S(u,v) = \sum_k \delta(u - u_k)\delta(v - v_k)$$

If we take an inverse FT of the sampled visibility function, we get a "dirty" image:

$$I^D(l,m) = \mathcal{F}^{-1}(S(u,v)\mathcal{V}(u,v))$$

Convolution theorem says:

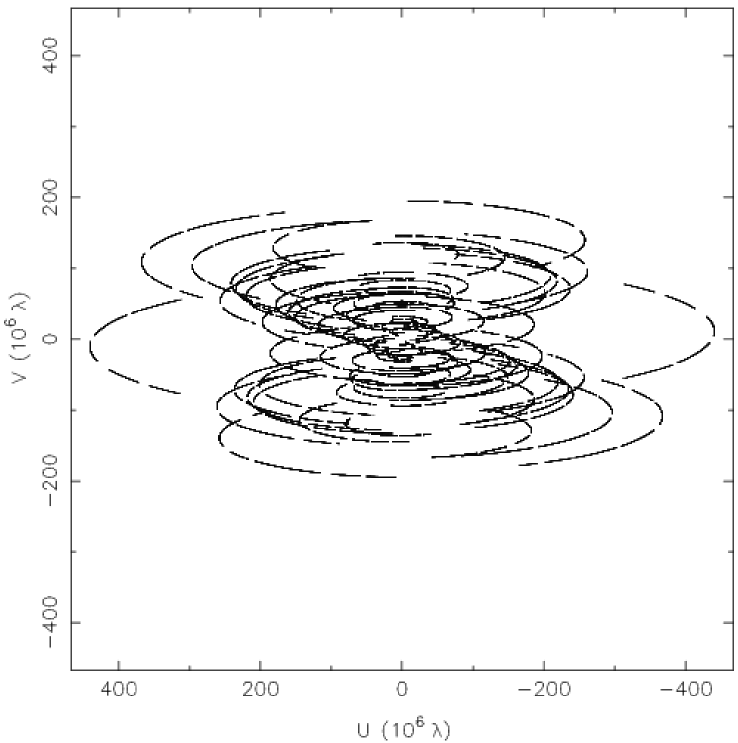
$$I^D(l,m) = b(l,m) * I(l,m)$$

So, $I^D(l,m)$ is a convolution of the true sky brightness distribution and the interferometer beam:

$$b(l,m) = \mathcal{F}^{-1}(S(u,v))$$

Interferometer beam

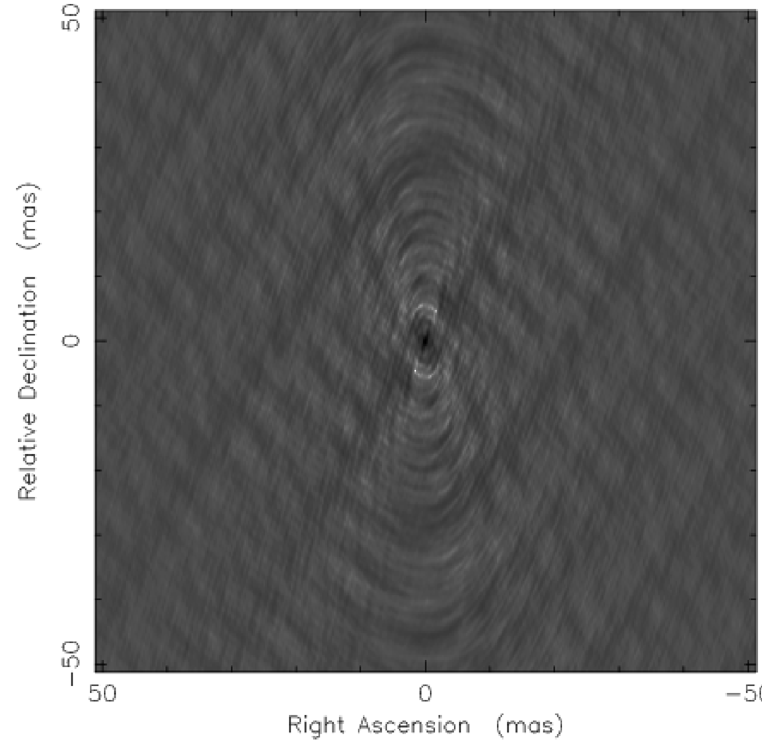
(u,v) plane sampling



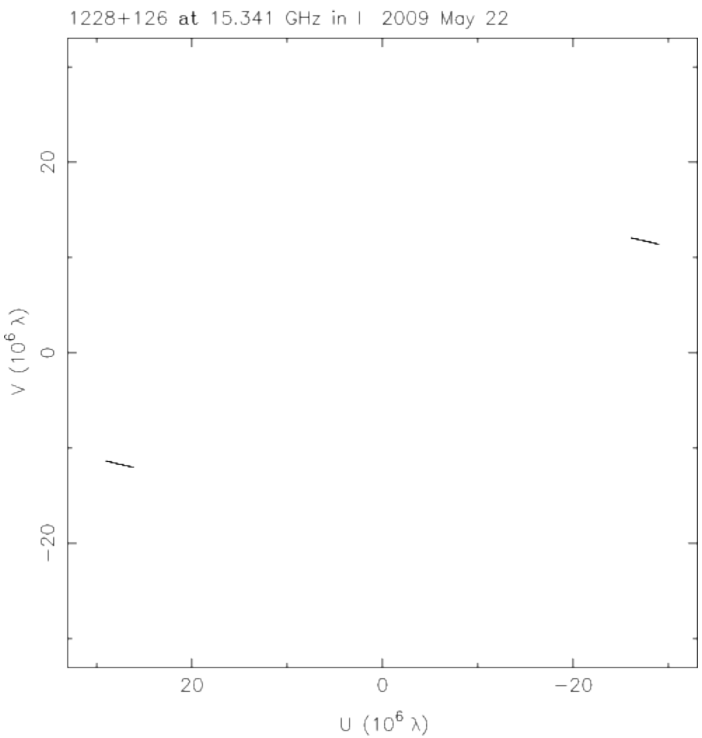
$\mathcal{F}$

$\leftrightharpoons$

Interferometer beam



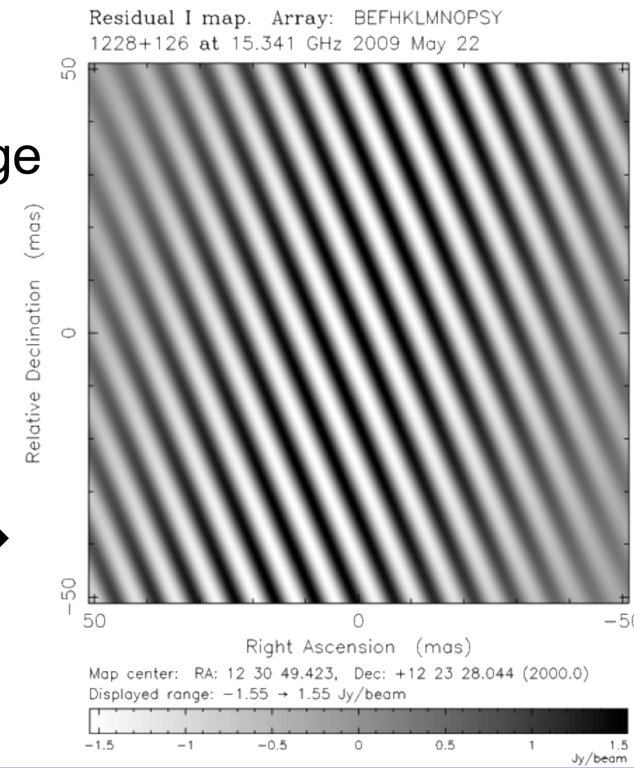
Example: Beam shape with increasing number of (u,v) samples



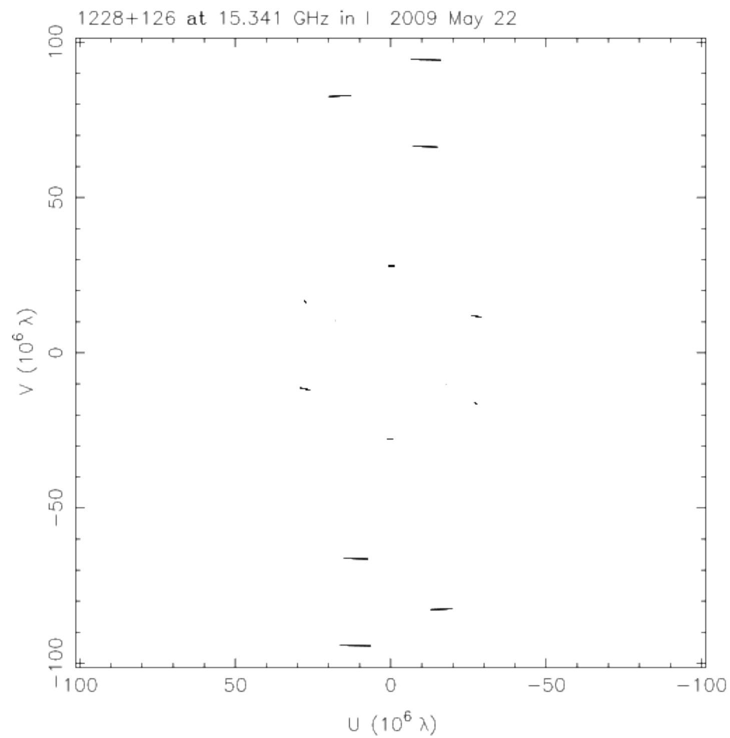
←(u,v) coverage

2 stations
25 min

Dirty image →



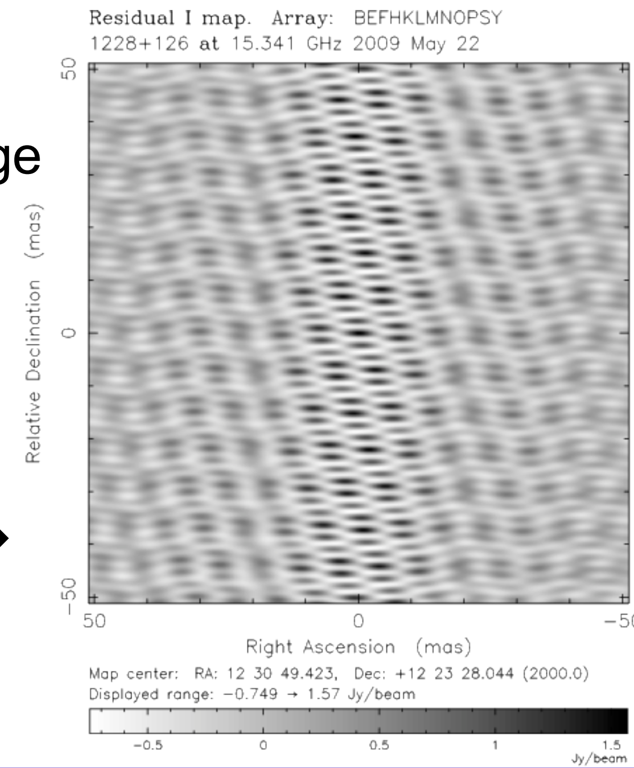
Example: Beam shape with increasing number of (u,v) samples



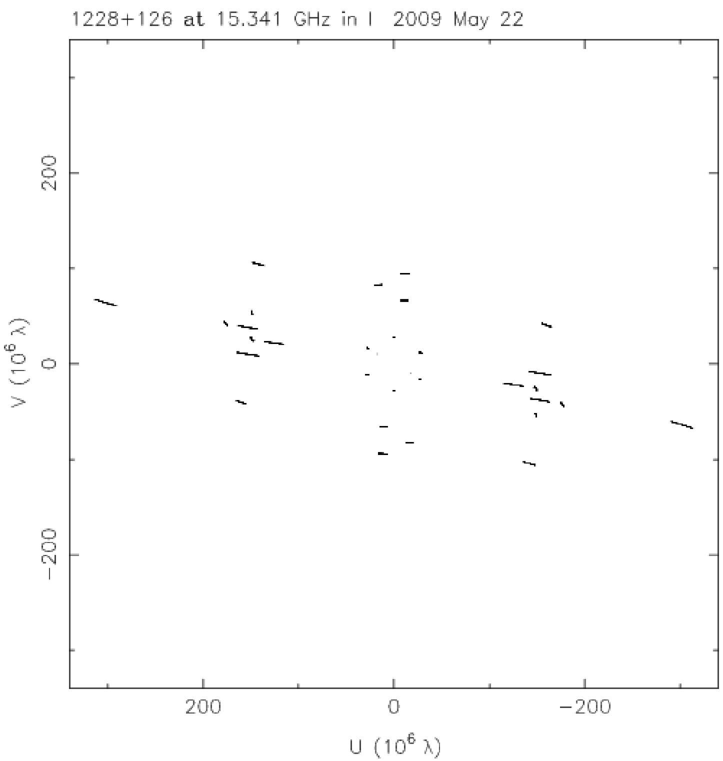
←(u,v) coverage

4 stations
25 min

Dirty image →



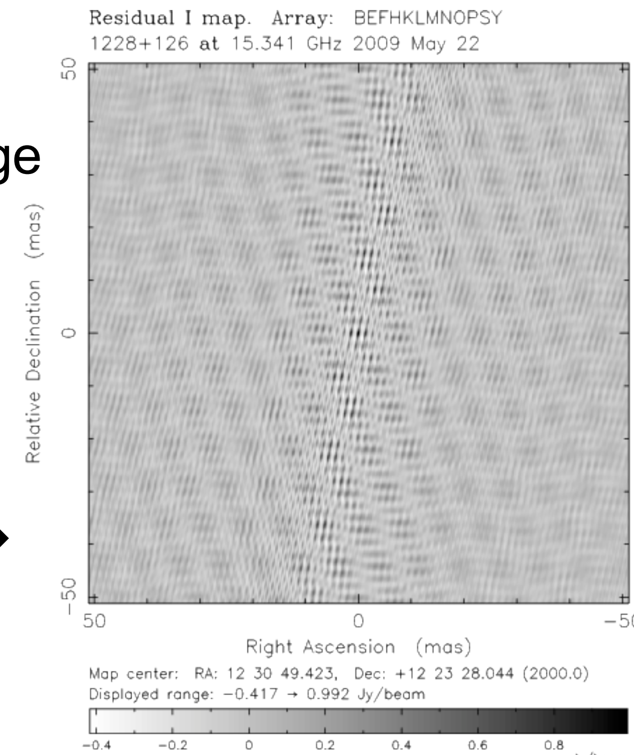
Example: Beam shape with increasing number of (u,v) samples



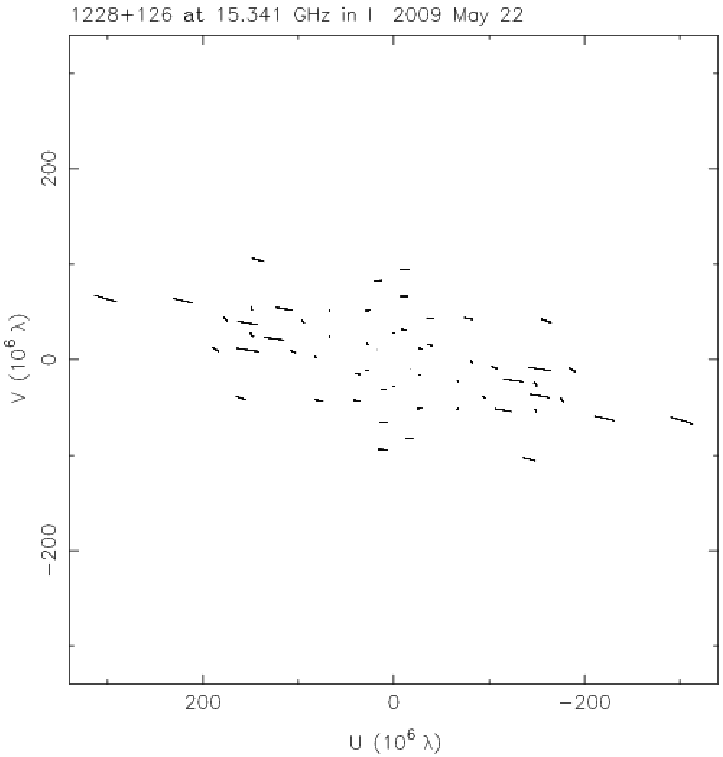
←(u,v) coverage

6 stations
25 min

Dirty image →



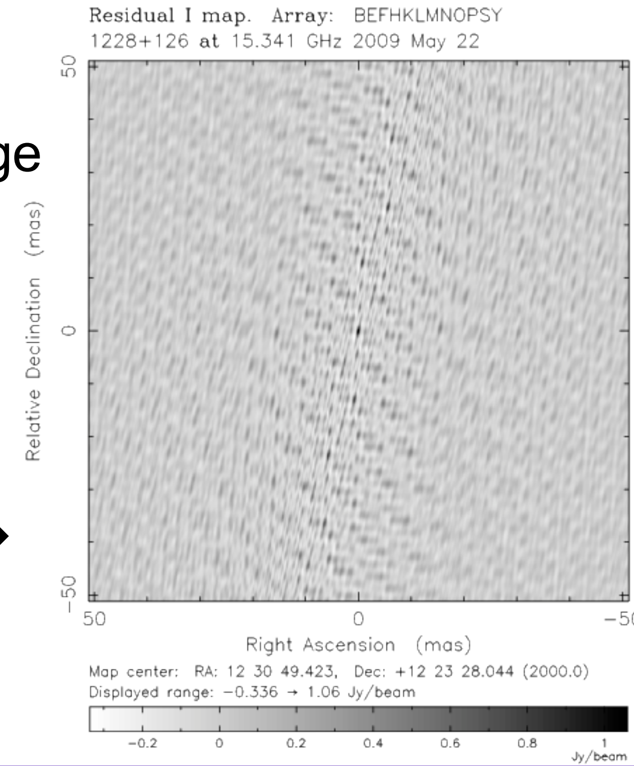
Example: Beam shape with increasing number of (u,v) samples



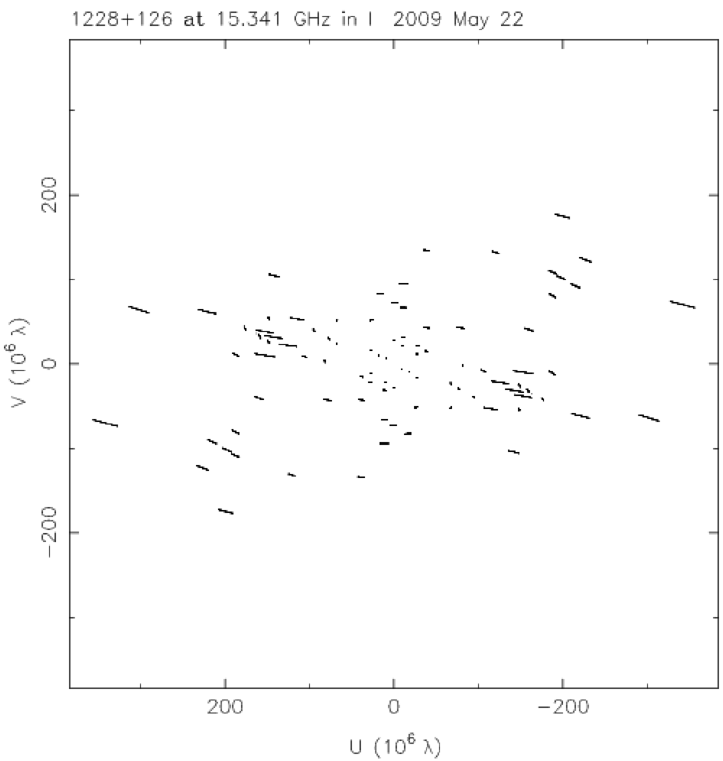
←(u,v) coverage

8 stations
25 min

Dirty image →



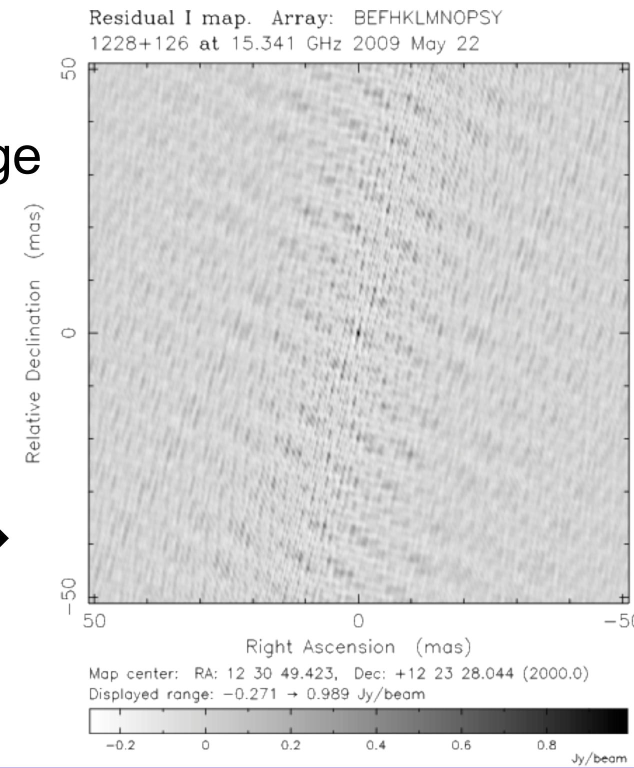
Example: Beam shape with increasing number of (u,v) samples



←(u,v) coverage

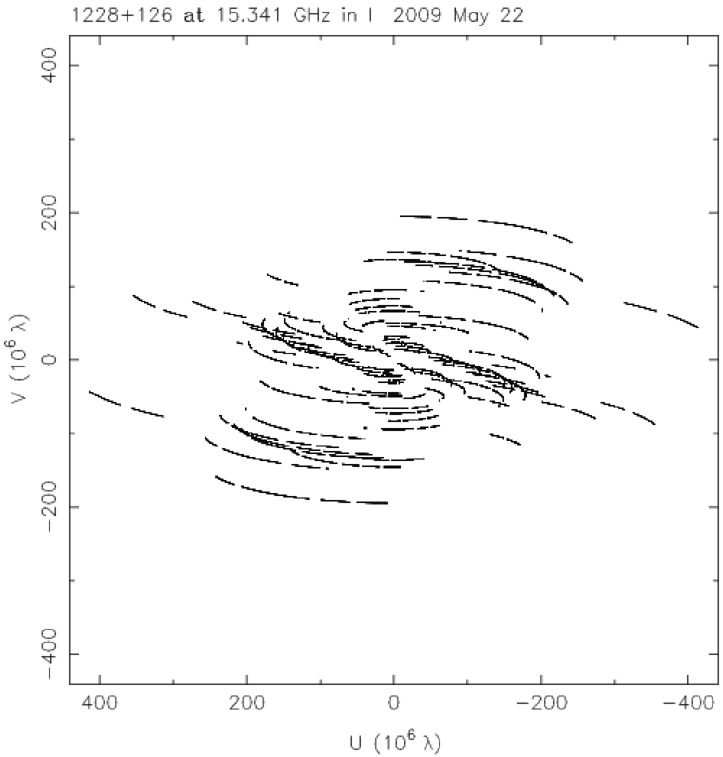
10 stations
25 min

Dirty image →



Aalto University
School of Electrical Engineering

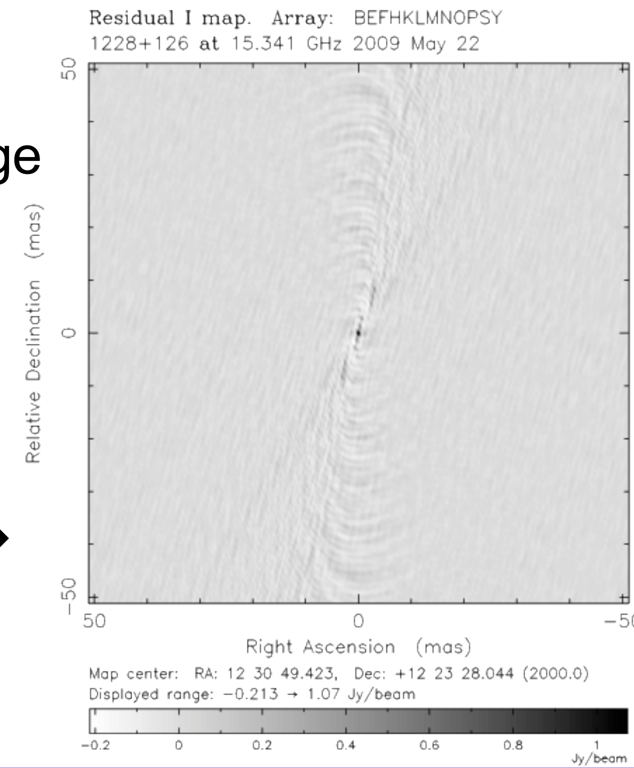
Example: Beam shape with increasing number of (u,v) samples



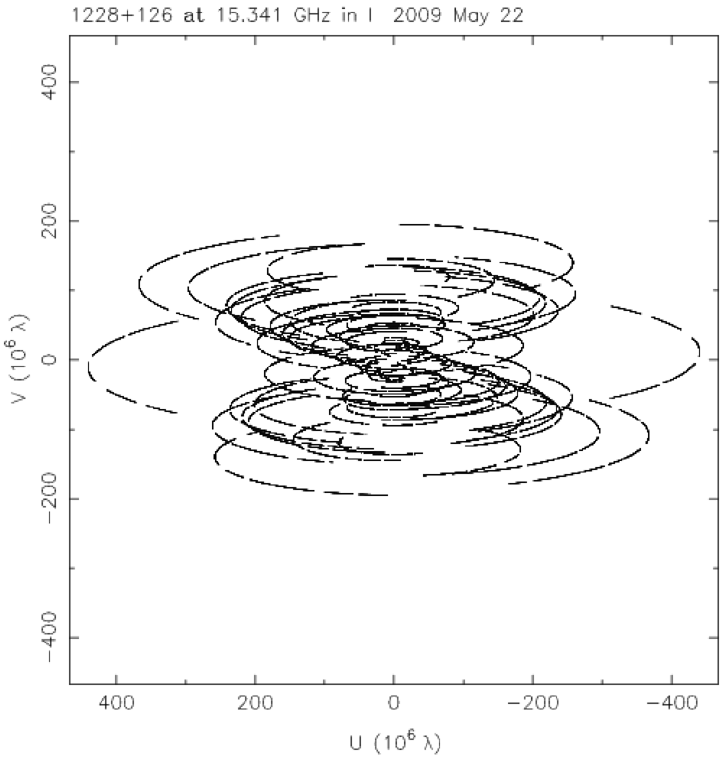
←(u,v) coverage

10 stations
5 hours

Dirty image →



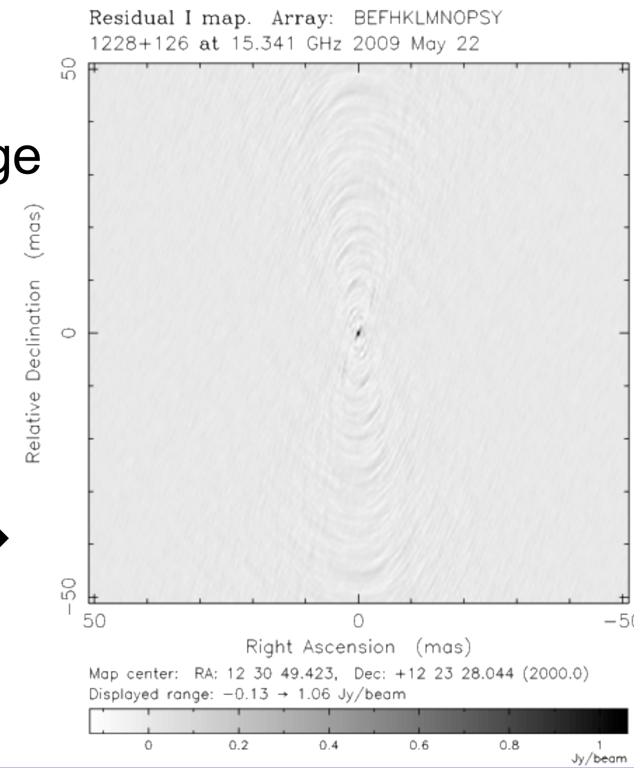
Example: Beam shape with increasing number of (u,v) samples



←(u,v) coverage

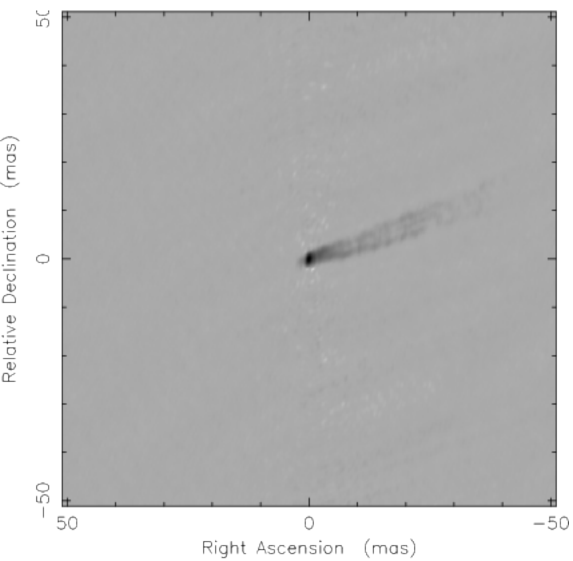
10 stations
11 hours

Dirty image →

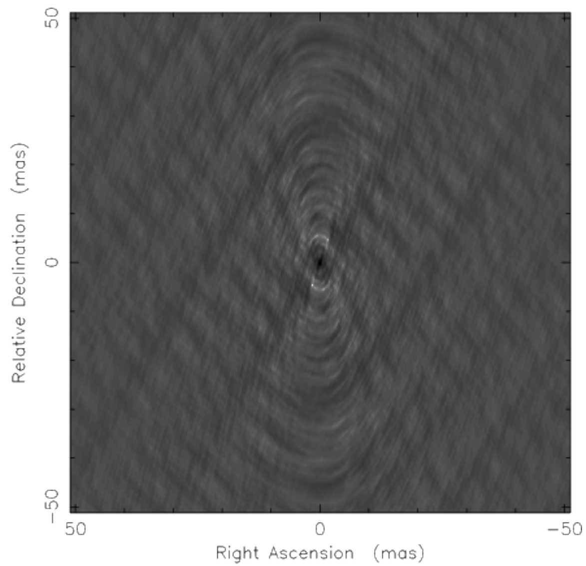


Dirty image

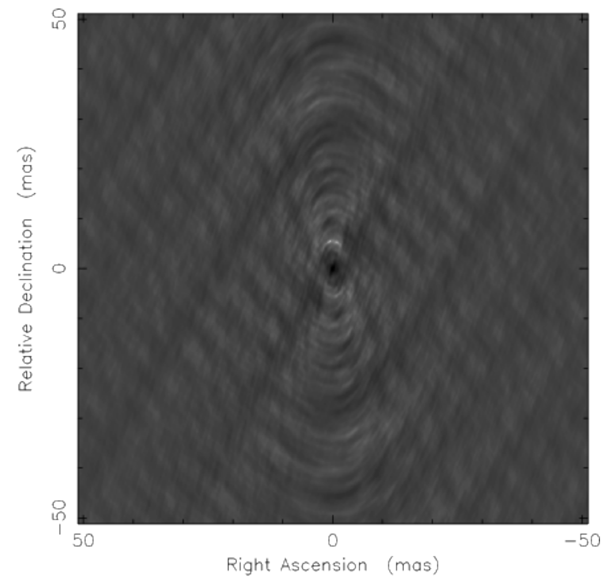
source structure **interferometer beam** **dirty image**



*



=



Aalto University
School of Electrical Engineering

Solving the inverse problem

Going beyond the dirty image

- Because of imperfect *(u,v)* plane sampling, imaging interferometric data is an ill-posed inverse problem.
- Can we deconvolve $I^D(l,m) = b(l,m) * I(l,m)$ to obtain $I(l,m)$? Unfortunately, standard linear deconvolution does not work, since the sampling function has zeroes and thus cannot be divided out.
- Reconstructing images requires information, assumptions or constraints beyond the interferometric measurements. Luckily, usually quite simple assumptions suffice, e.g., 1) finite source size, 2) positivity, 3) smoothness and/or 4) sparseness of the true brightness distribution.
- Two approaches: 1) Inverse modeling by non-linear deconvolution algorithms (CLEAN), 2) forward modeling either by regularized maximum likelihood algorithms or by a full Bayesian approach.

Non-linear deconvolution with a CLEAN algorithm

CLEAN is the most widely used algorithm (implementations in CASA, AIPS, Difmap …)

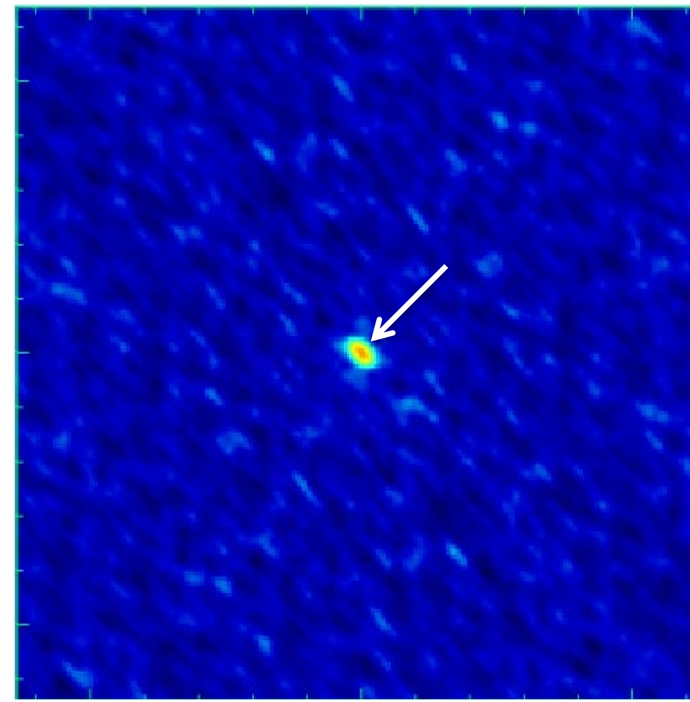
- Fits and subtracts the interferometer beam iteratively
- Original version by Högbom (1974), several improvements later
- Assumes that source structure can be presented as a sum of a finite number of point sources
- User can supply *a priori* information by restricting the area where CLEAN is allowed to work (“CLEAN windows”)
- Has problems with diffuse emission (creates “spotty” structures)
- Instabilities: striping around extended sources is a common artefact

Deconvolution with CLEAN algorithm

Basic algorithm:

Initialize: *residual map* = *dirty map* and list of δ-components = empty

1. Find the peak in the residual map, identify it as a point source

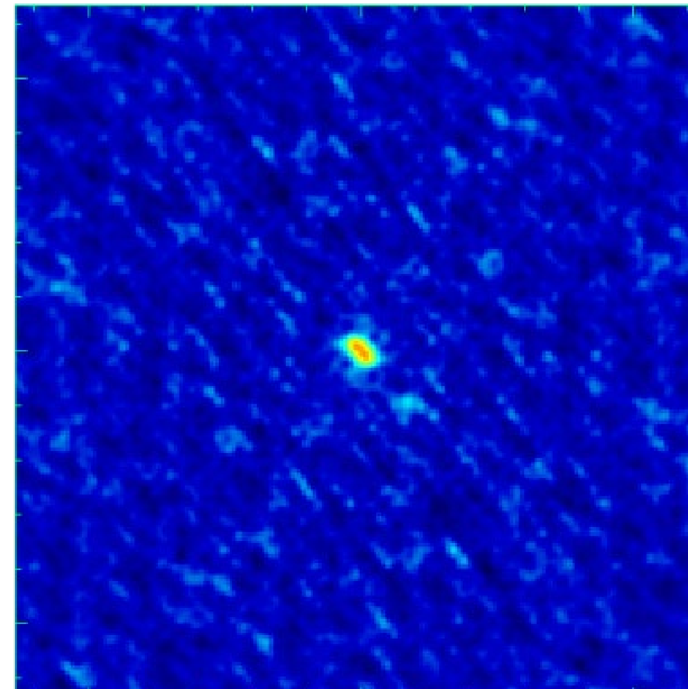


Deconvolution with CLEAN algorithm

Basic algorithm:

Initialize: *residual map* = *dirty map* and list of δ-components = empty

1. Find the peak in the residual map, identify it as a point source
2. Subtract this point source, scaled by *loop_gain* and convolved with the interferometer beam, from the residual image

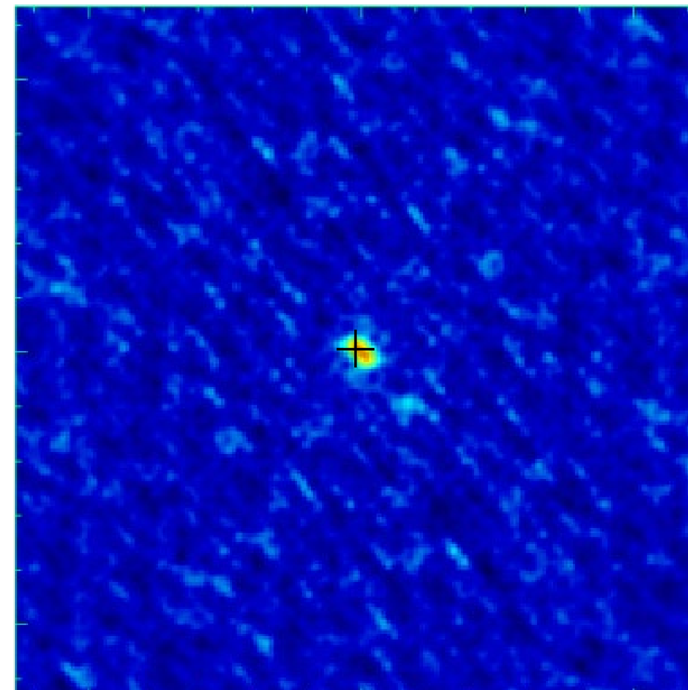


Deconvolution with CLEAN algorithm

Basic algorithm:

Initialize: *residual map* = *dirty map* and list of δ-components = empty

1. Find the peak in the residual map, identify it as a point source
2. Subtract this point source, scaled by *loop_gain* and convolved with the interferometer beam, from the residual image
3. Save the position and subtracted flux to the list of δ-components

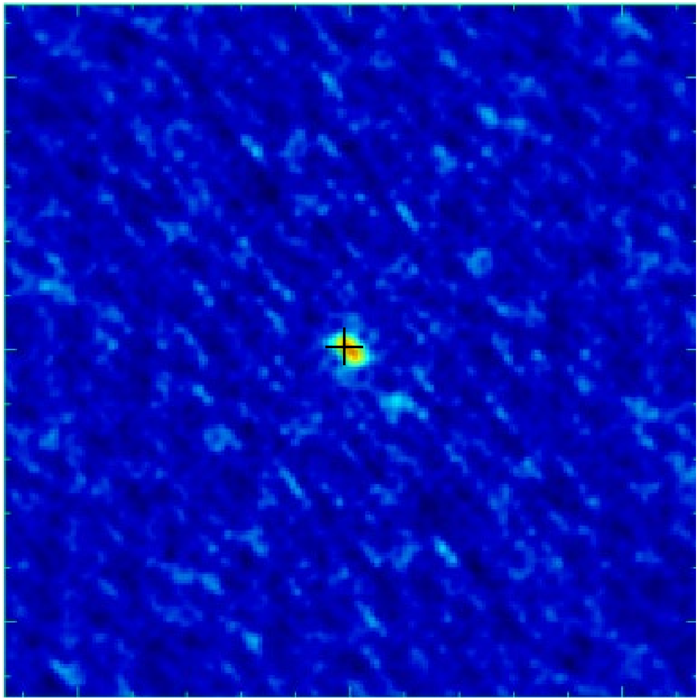


Deconvolution with CLEAN algorithm

Basic algorithm:

Initialize: *residual map* = *dirty map* and list of δ-components = empty

1. Find the peak in the residual map, identify it as a point source
2. Subtract this point source, scaled by *loop_gain* and convolved with the interferometer beam, from the residual image
3. Save the position and subtracted flux to the list of δ-components
4. If stopping criteria are not met, go to step 1



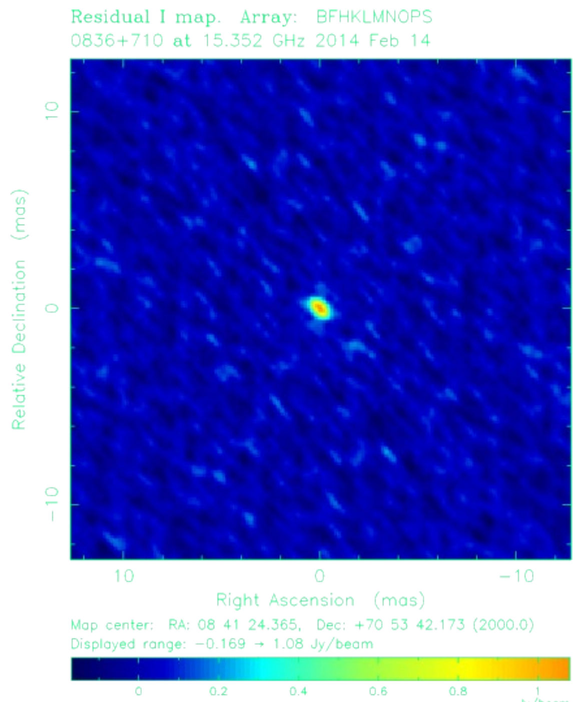
Deconvolution with CLEAN algorithm

- Stopping criteria? Target noise level reached, target SNR reached, or some maximum number of iterations reached.
- Final step – make “restored” image:
  - Make a model image from the final list of δ-components
  - Convolve the model image with a “CLEAN beam”, which is typically a Gaussian fitted to the central peak of the interferometer beam
  - Add the last residual map to show possible imaging artefacts
- The resulting image is an estimate of $I(l, m)$.
- The units are typically Jy / clean_beam_area.

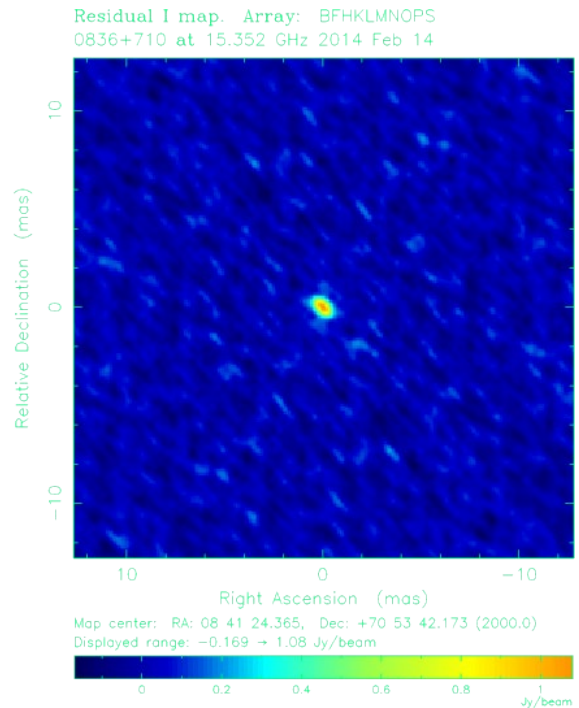
CLEAN example

CLEAN iterations = 0

Dirty image



Residual image

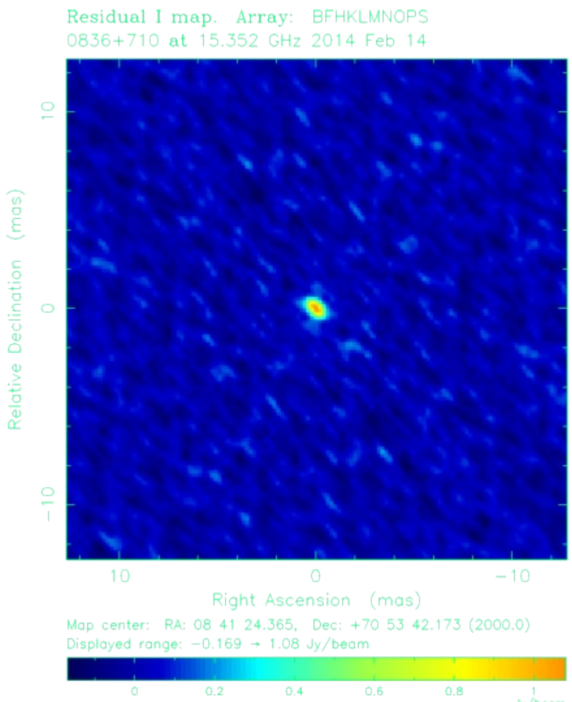


CLEAN image (log)

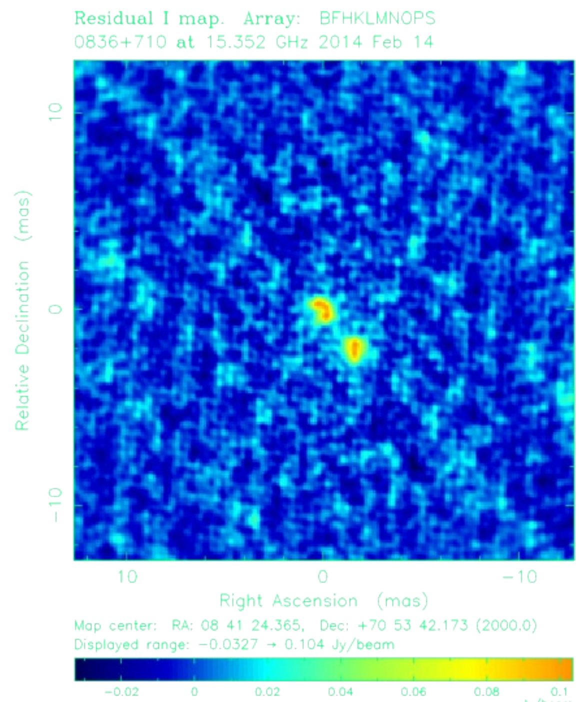
CLEAN example

CLEAN iterations = 100

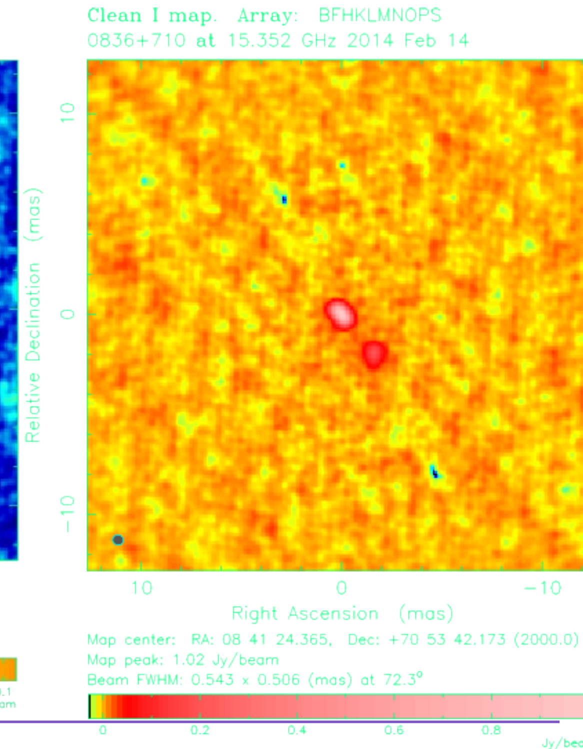
Dirty image



Residual image



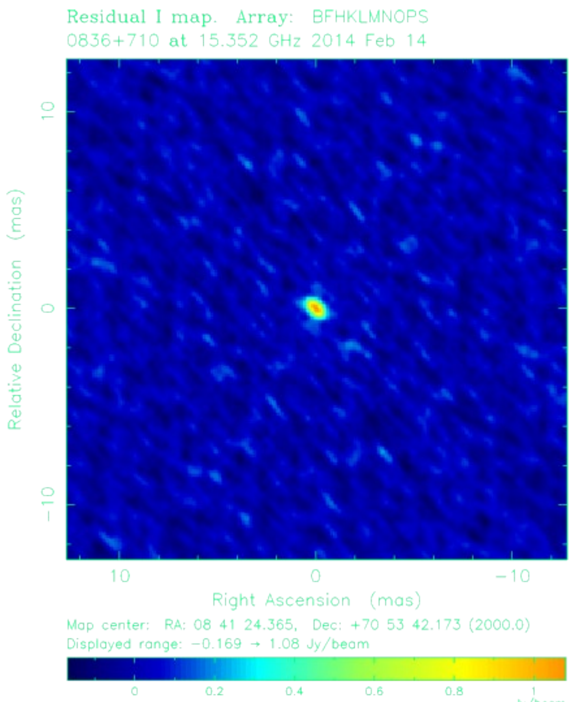
CLEAN image (log)



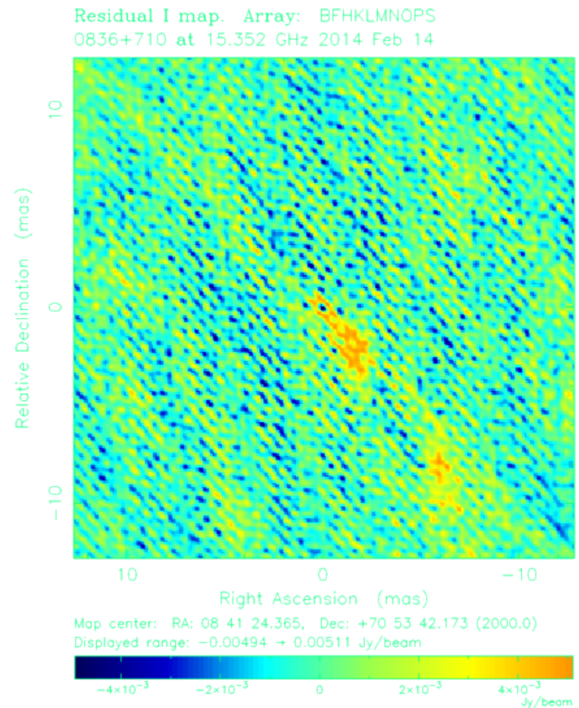
CLEAN example

CLEAN iterations = 500

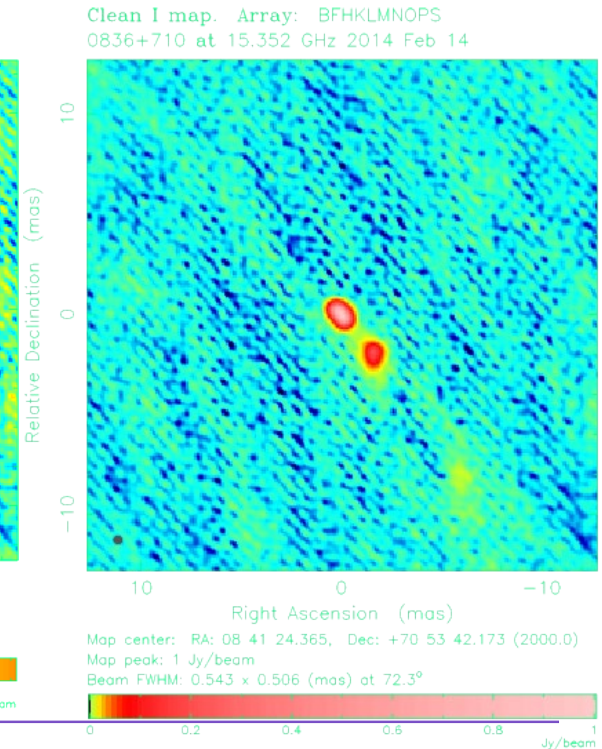
Dirty image



Residual image



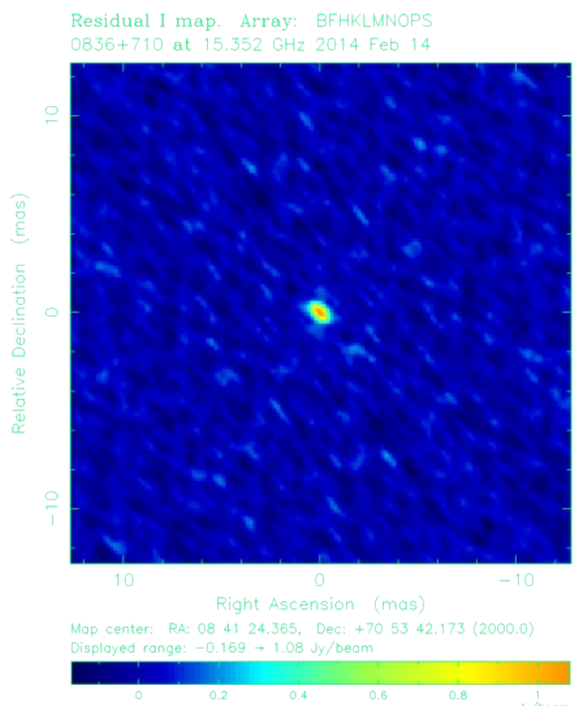
CLEAN image (log)



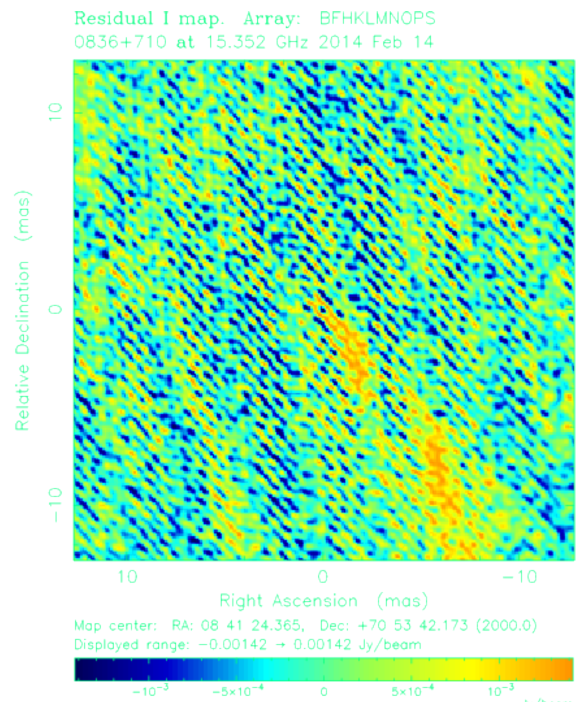
CLEAN example

CLEAN iterations = 1500

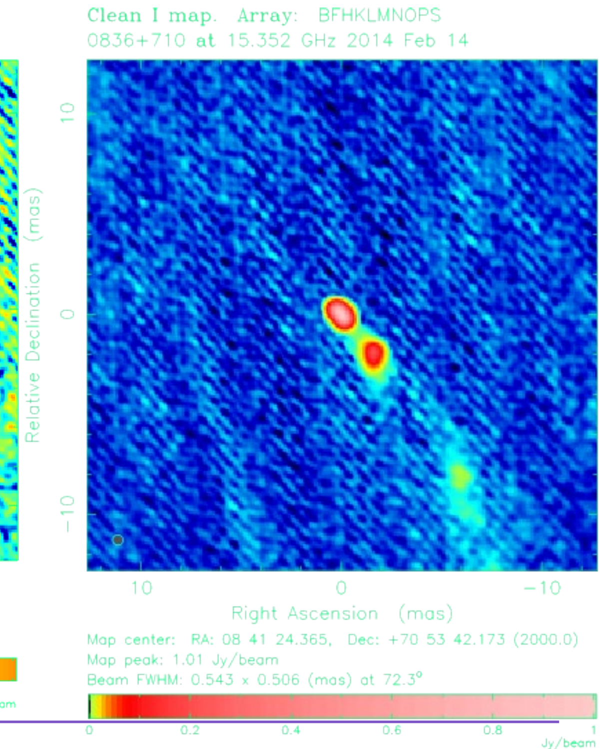
Dirty image



Residual image



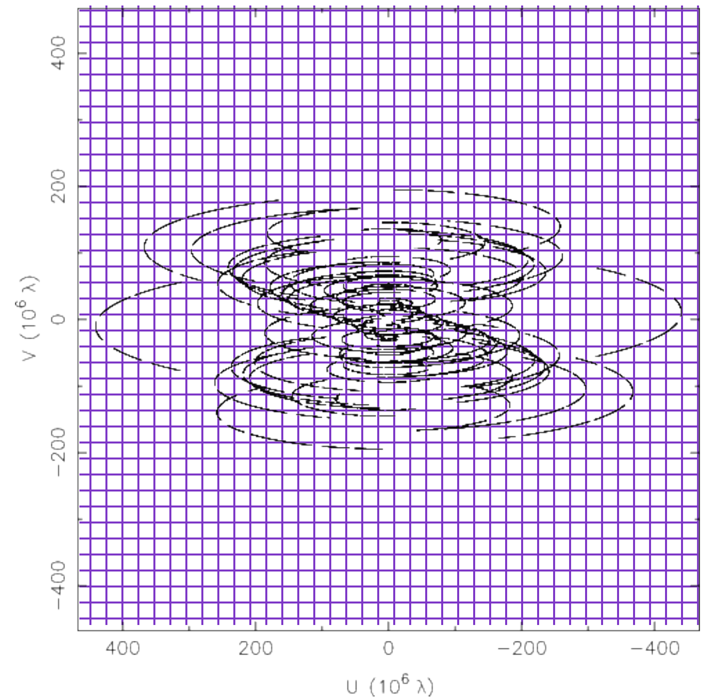
CLEAN image (log)



Aalto University
School of Electrical Engineering

A note about practical Fourier transformation

- Fast Fourier Transform (FFT) is typically used to invert the data, since it is much faster than direct FT ($\mathcal{O}(N^2 \log_2 N)$ vs. $\mathcal{O}(N^4)$) for an image of $N \times N$ pixels and $\sim N^2$data points
- FFT requires data points on a rectangular grid → $\mathcal{V}(u,v)$ needs to be interpolated and resampled for FFT



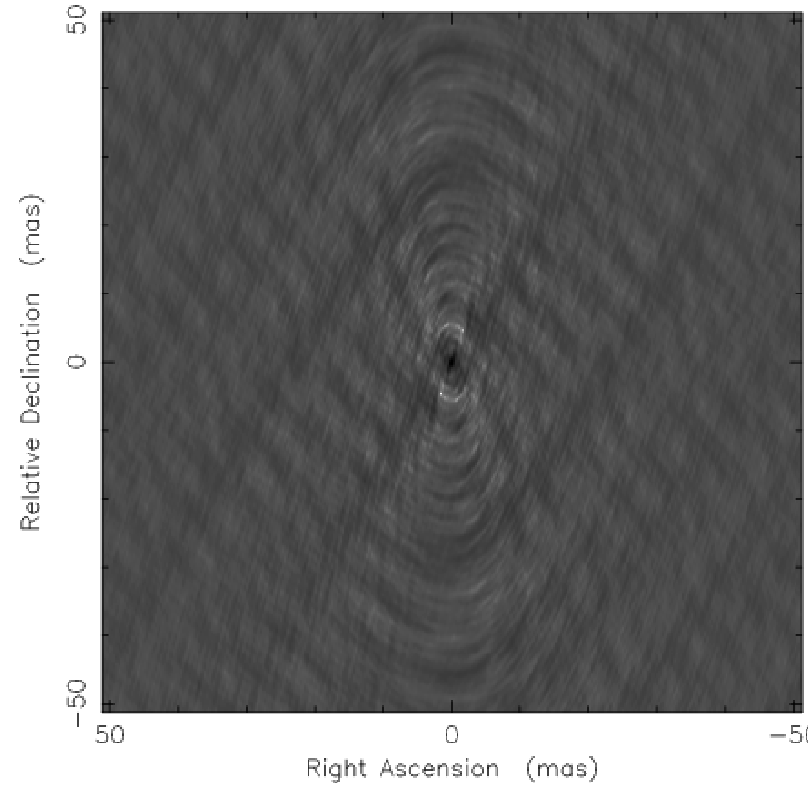
Weighting of the visibility data

Modify the sampling function by a weighting function *W*(*u*,*v*)

- *S*(*u*,*v*) → *W*(*u*,*v*)*S*(*u*,*v*)
- Modifies the interferometer beam

Natural weighting

- $W(u,v) = 1/\sigma_{u,v}^2$ in occupied cells, 0 elsewhere
- Maximizes point source sensitivity
- Typically weights more the short baselines → loses in angular resolution



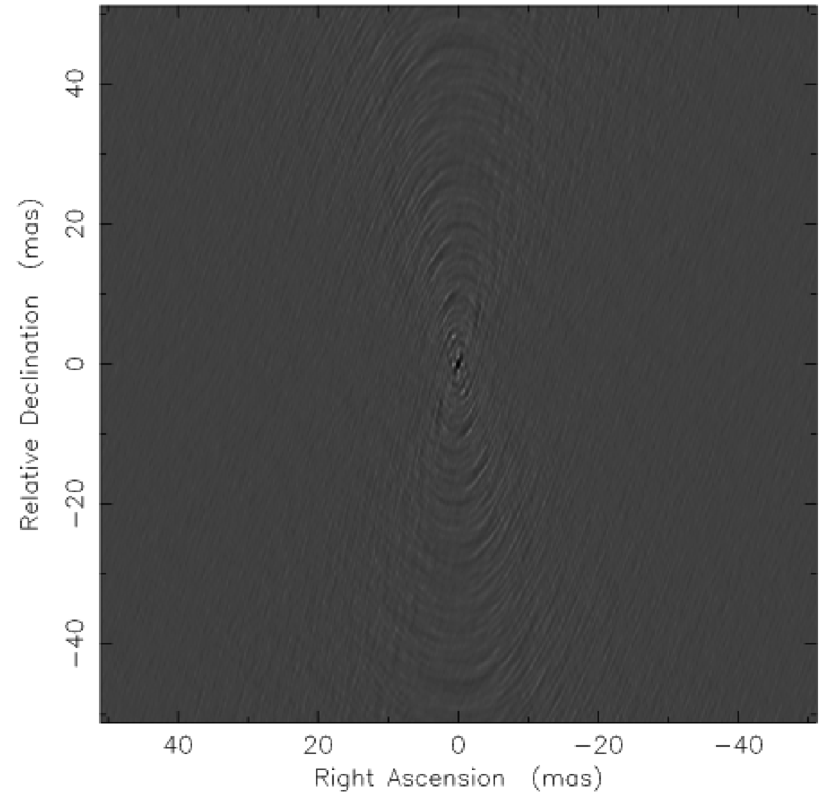
Weighting of the visibility data

Uniform weighting

- $W(u, v) = 1/\rho(u, v)$ where $\rho(u, v)$ is the local density of visibilities in the (u,v) plane. Depends on selected "box size".
- Weights more the long baselines, enhancing angular resolution
- Degrades point source sensitivity
- Be careful, if sampling is sparse

Other weighting schemes

- Briggs' robust weighting



Forward modeling

Regularized maximum likelihood (RML) methods

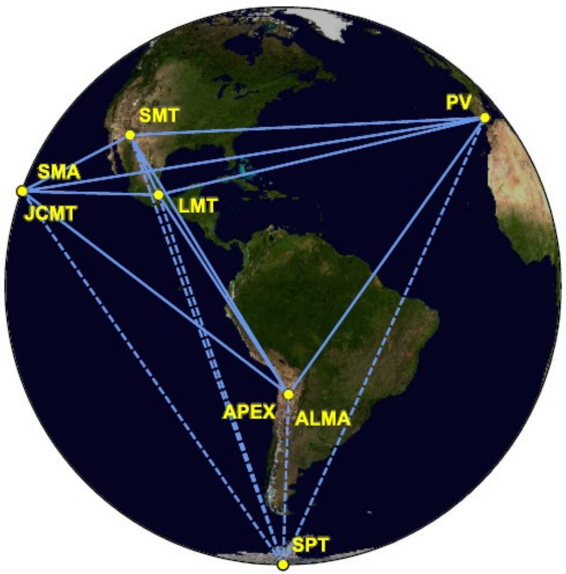
- Represent the image as an array of pixels and Fourier-transform this array to evaluate consistency with the data
- Find the image I that minimizes an objective function

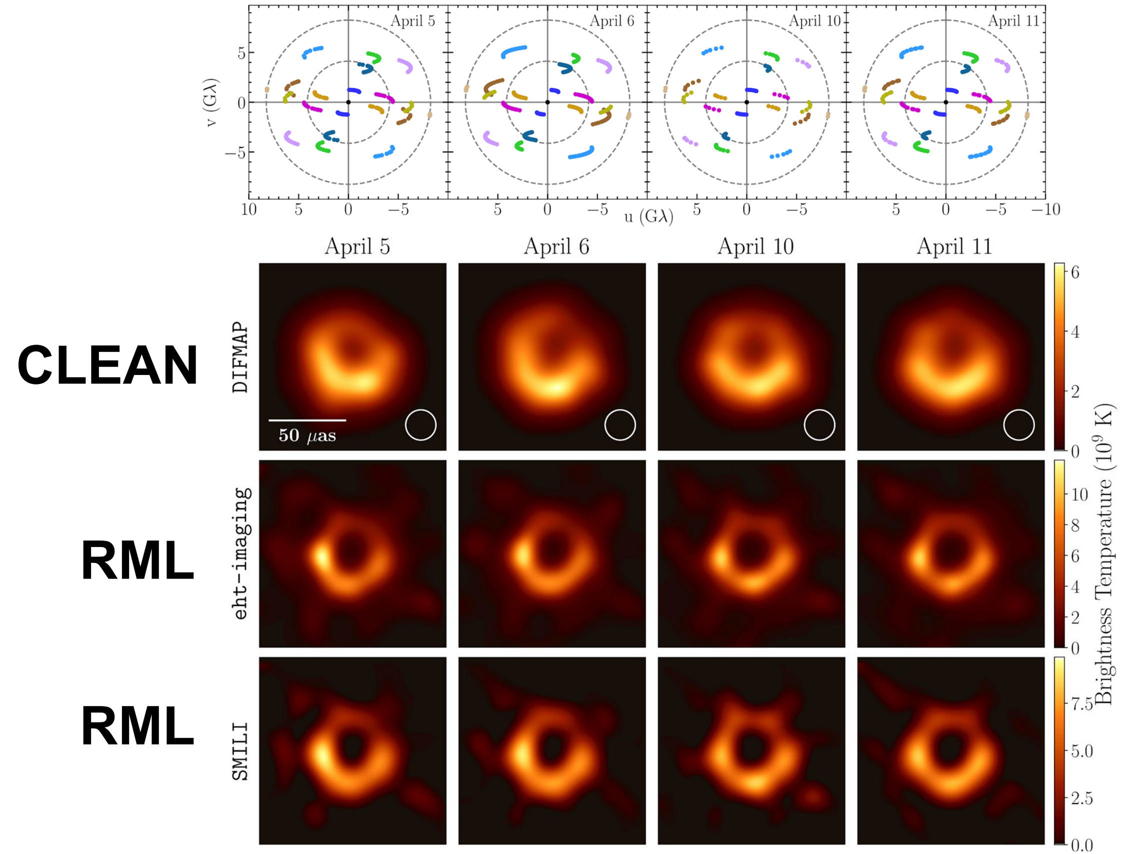
$$J(I) = \sum_{\text{data terms}} \alpha_D \chi^2_D(I) - \sum_{\text{regularizers}} \beta_R S_R(I)$$

where χ^2_D is a goodness-of-fit function, $S_R(I)$ is a regularization term, and α_D and β_R are hyperparameters.

- Typical regularizers: image entropy, smoothness, image sparsity etc.
- No final restoring beam is required

Example: EHT image of the black hole



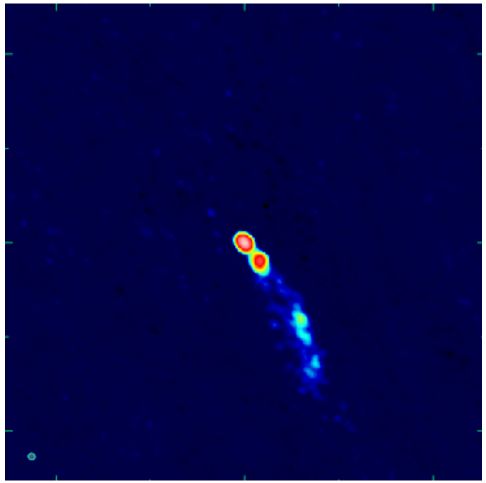


Event Horizon Telescope Collaboration (2019)

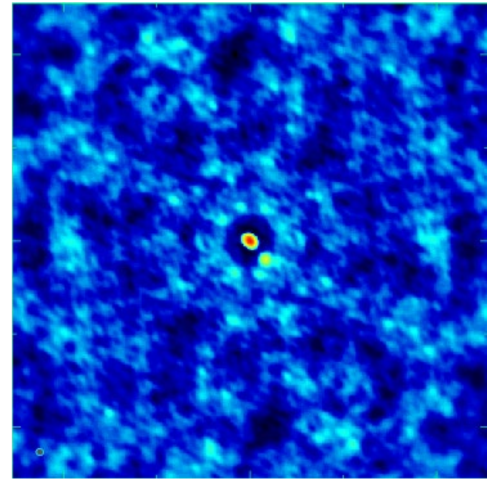
Aalto University
School of Electrical
Engineering

Image quality and error recognition

Is my image ok?



Yes!
Go to do science.



No!
Find the cause and redo some of the calibration / imaging steps.

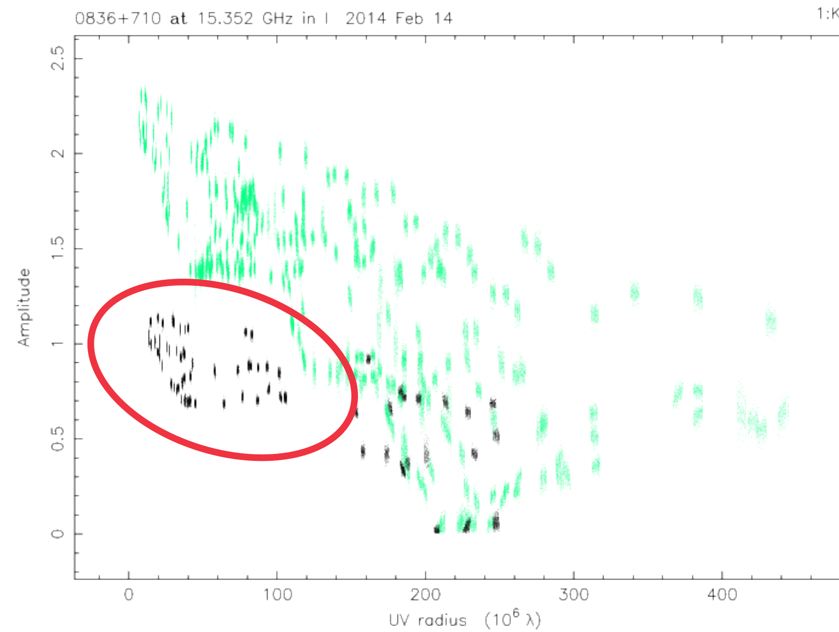
The final image depends on…

- Imaging parameters (image and pixel size, visibility weighting, gridding)
- Imaging method (used algorithm and its parameters)
- Any errors in calibration and/or editing of the visibilities, i.e. existence of bad data
- Noise

How can I tell?
I. Identifying bad data in the (u,v) plane

Look at data in the (*u*,*v*) plane first:

- Easier to identify outliers in (*u*,*v*) plane – their effect is spread throughout image plane
- Plot visibilities vs. baseline length or time – variations should be smooth
- Fraction *f* of slightly bad data gives errors at the level of *f* in the image – look for gross outliers
- Plot weights – look for large discrepant values
- Beware of RFI – check spectral plots

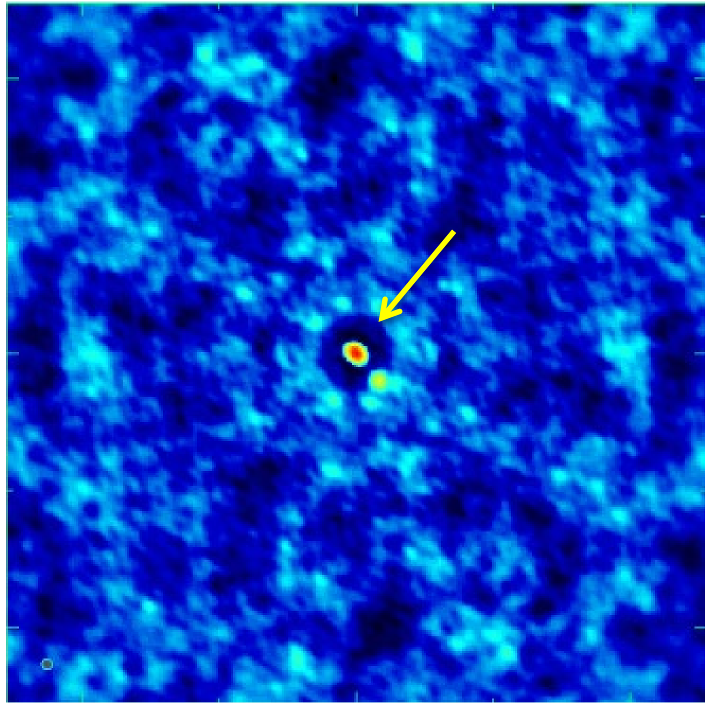


One antenna has amplitudes down by 50%. Increases image noise by a factor of 100!

How can I tell?
II. Imaging artefacts

Persistent errors sometimes easier to find in the image plane:

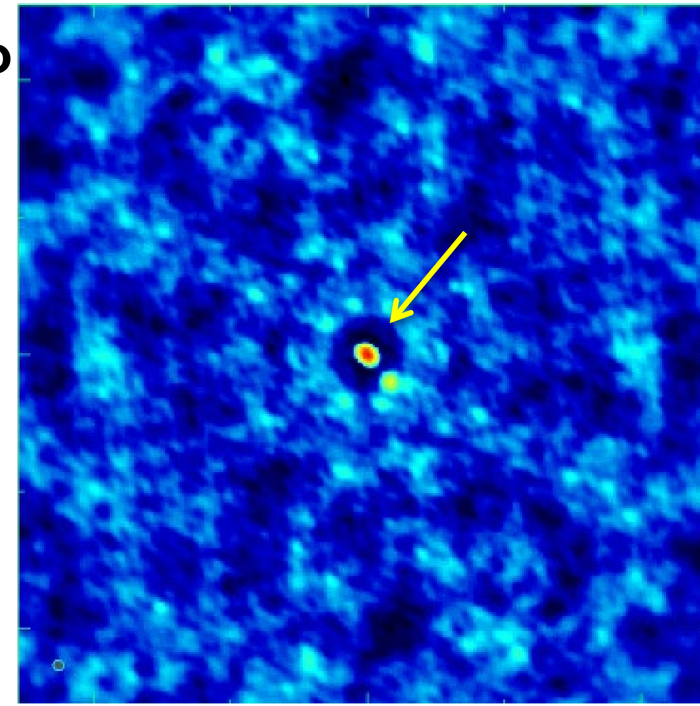
- For example, a 5% antenna gain calibration error is difficult to see in the data, but causes artefacts in the image at 1% level
- Look for unnatural structures in the image:
  - Stripes or rings around bright features
  - Negative bowls around extended structure
  - Spotty on-source structure or short-wavelength ripples
  - Features resembling the interferometer beam



How can I tell?
II. Imaging artefacts

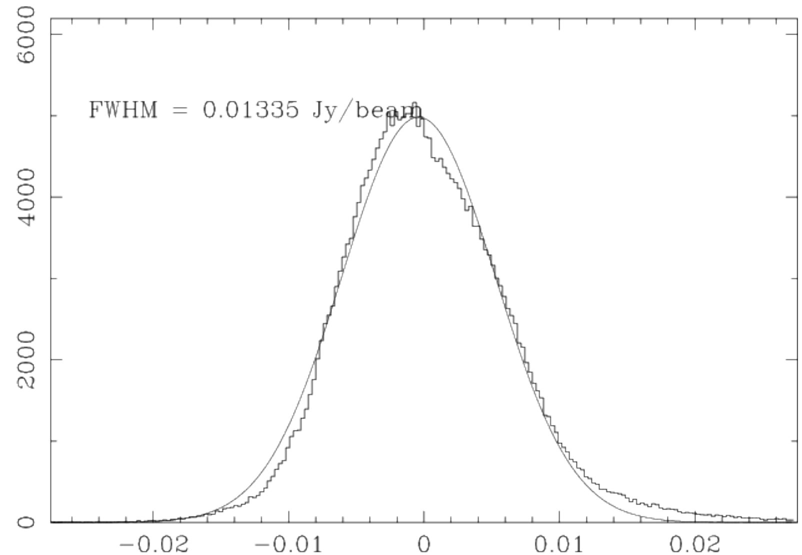
Persistent errors sometimes easier to find in the image plane:

- Are these artefacts additive (constant over the field) or multiplicative (brighter around bright sources)?
- Are they symmetric or antisymmetric around bright sources?



How can I tell?
III. Noise in the image

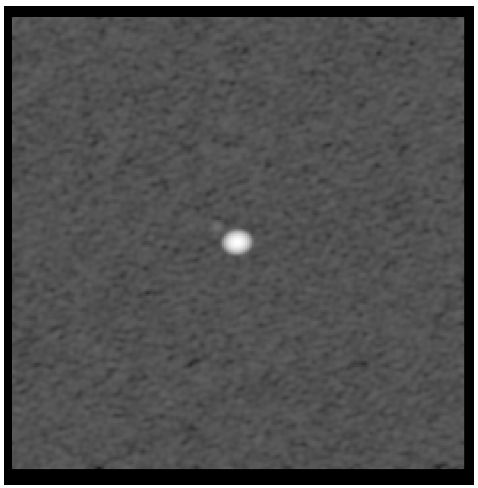
- Calculate the expected thermal noise level in the final image from the sensitivity of your interferometer and integration time
- Measure off-source rms noise by e.g., making a histogram of pixel fluxes and fitting a Gaussian. Is the distribution Gaussian?
- Compare expected and measured rms noise. If you do not reach the thermal noise level, find out why.



Flux distribution of the image from the previous slide. The expected rms noise was 0.0001 Jy/beam!

Identifying bad data in the image plane - short burst of bad data at all antennas

Results for a point source using VLA. 13 x 5min observation over 10 hr. Images shown after editing, calibration and deconvolution.



No errors (rms 0.11 mJy/beam)

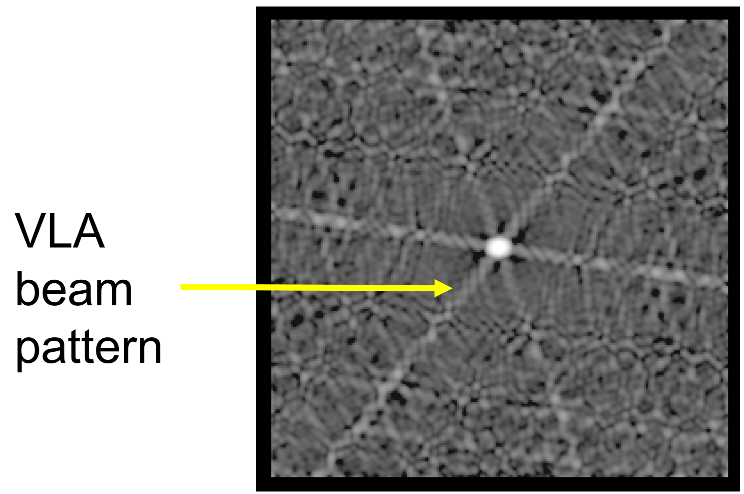
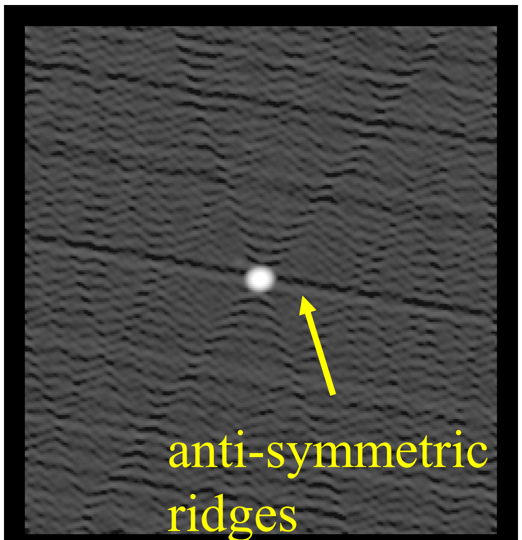


Image credit: Greg Taylor

10% amplitude error for all antennas at one time (rms 2.0 mJy/beam)

Identifying bad data in the image plane - short burst of bad data at one antenna

10 deg phase error for one antenna at one time (rms 0.49 mJy/beam)



20% amplitude error for one antenna at one time (rms 0.56 mJy/beam)

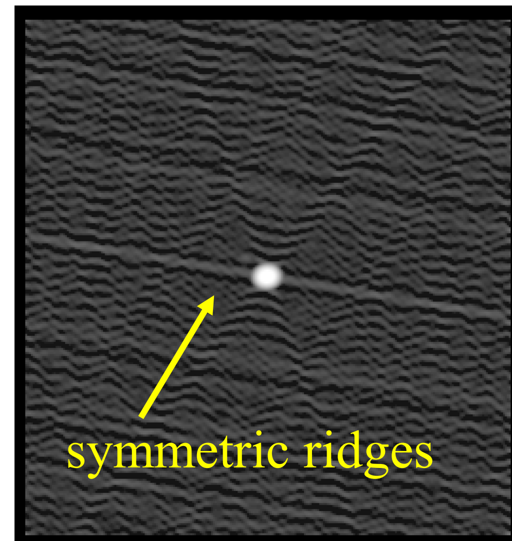
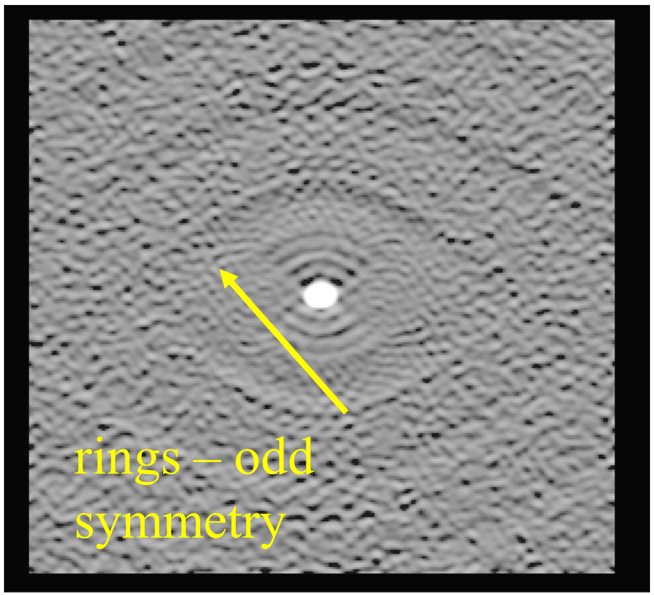


Image credit: Greg Taylor

Identifying bad data in the image plane - persistent bad data

10 deg phase error for one antenna all times (rms 2.0 mJy/beam)



20% amp error for one antenna all times (rms 2.3 mJy/beam)

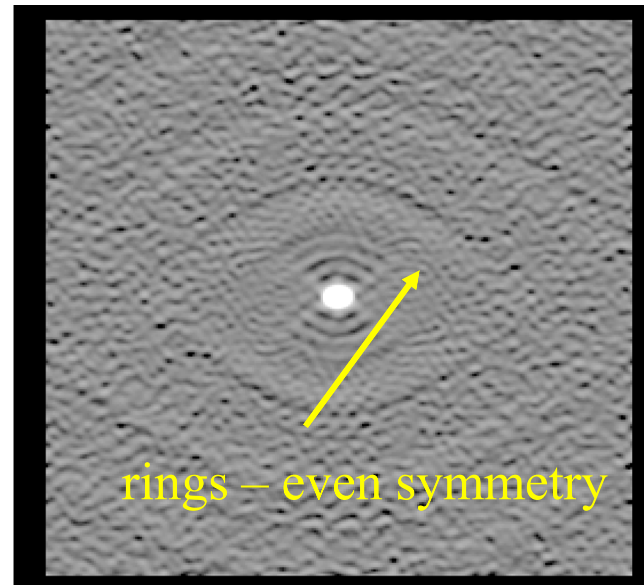


Image credit: Greg Taylor

Other causes of problems: I. Missing short spacings

- If short (*u*,*v*) spacings are missing from the data, there is no information about structures larger than $\sim {}^{\lambda}/_{2B_{min}}$
- Negative bowl around an extended source is often a sign of unmeasured power at short (*u*,*v*) spacings

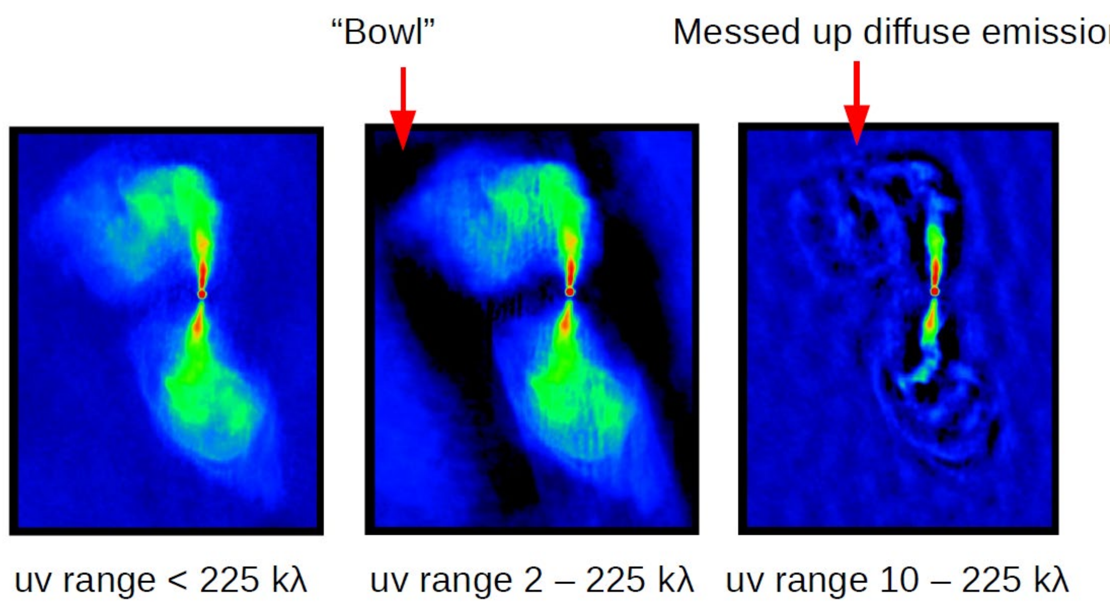


Image credit: Robert Laing

Other causes of problems: II. Deconvolution errors

- Wrongly selected CLEAN windows
- Too shallow or too deep CLEAN
- Poor choice of weighting

Effect of CLEAN windows

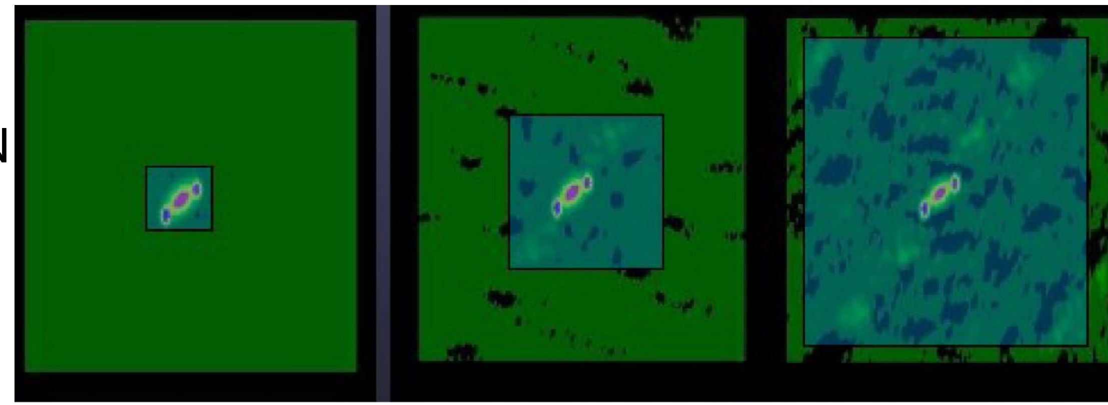
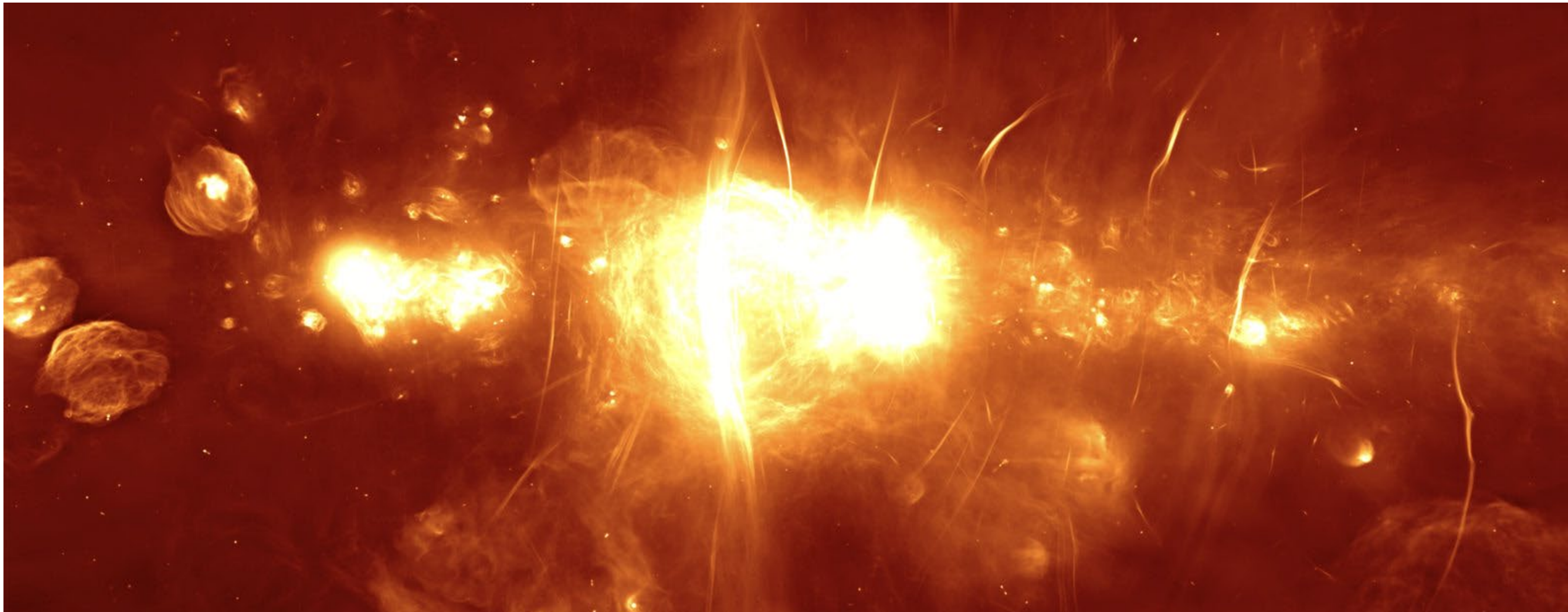


Image credit: Robert Laing

Correct Too big Far too big

Select tight enough CLEAN boxes to avoid CLEANing noise interacting with sidelobes.

A proof that all this actually works…



Aperture synthesis image of the Galactic Center made by MeerKAT array

Summary

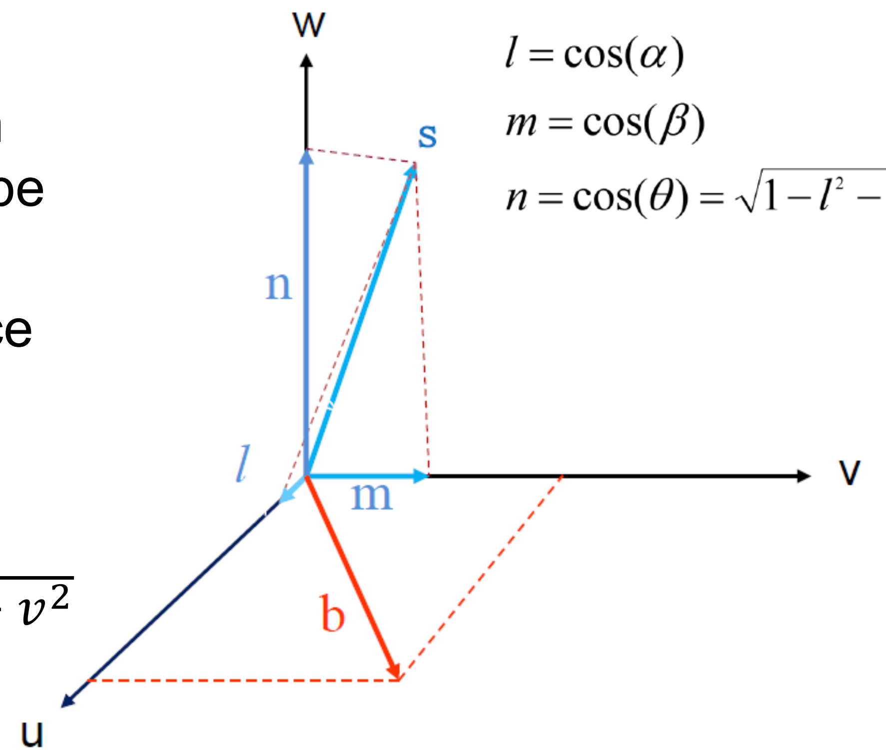
- Interferometer samples Fourier components of the sky brightness distribution
- Inverse Fourier transform of the measured visibilities gives an image
- Due to incomplete sampling of the visibility function, imaging is an ill-posed inverse problem, which can be solved either by inverse modeling (non-linear deconvolution) or by forward modeling (e.g., RML or Bayesian methods).
- There are an infinite number of brightness distributions that can fit the observed visibilities. Astronomers must be cautious and exercise judgement while imaging interferometric data!
- Still, most of the time things do converge!

Aalto University
School of Electrical
Engineering

Extra slides

(*u,v,w*) coordinate system

- (*u,v,w*) coordinates (measured in wavelengths) are used to describe antenna positions and baselines
- *w* points to and follows the source (or phase tracking center), *u* is towards East, *v* towards North celestial pole
- Projected baseline length: $\sqrt{u^2 + v^2}$
- (*l,m,n*) are direction cosines describing direction vector ***s***



$$l = \cos(\alpha)$$
$$m = \cos(\beta)$$
$$n = \cos(\theta) = \sqrt{1 - l^2 - m^2}$$

Image credit: Rick Perley

Some properties of Fourier transforms

Fourier transform:

$$\boldsymbol{F(u)} = \mathcal{F}(\boldsymbol{f(x)}) \equiv \int_{-\infty}^{\infty} \boldsymbol{f(x) e^{-i2\pi ux} dx}$$

Inverse Fourier transform:

$$\boldsymbol{f(x)} = \mathcal{F}^{-1}(\boldsymbol{F(u)}) \equiv \int_{-\infty}^{\infty} \boldsymbol{F(u) e^{i2\pi ux} du}$$

Linearity:

$$\mathcal{F}(\boldsymbol{f} + \boldsymbol{g}) = \mathcal{F}(\boldsymbol{f}) + \mathcal{F}(\boldsymbol{g})$$

Shifting:

$$\mathcal{F}(\boldsymbol{f(x - x_0)}) = \boldsymbol{F(u) e^{i2\pi u x_0}}$$

Scaling:

$$\mathcal{F}(\boldsymbol{f(ax)}) = \frac{\boldsymbol{1}}{\boldsymbol{|a|}} \boldsymbol{F}\left(\frac{\boldsymbol{u}}{\boldsymbol{a}}\right)$$

Convolution:

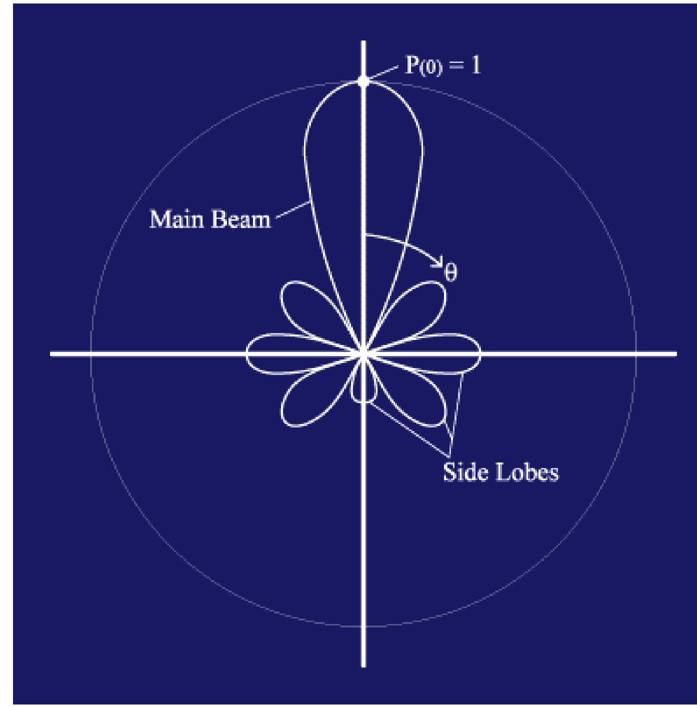
$$\boldsymbol{f(x) * g(x)} \equiv \int_{-\infty}^{\infty} \boldsymbol{f(x') g(x' - x) dx'}$$

$$\mathcal{F}(\boldsymbol{f} * \boldsymbol{g}) = \mathcal{F}(\boldsymbol{f})\mathcal{F}(\boldsymbol{g})$$

Effect of the antenna reception pattern

The antenna reception pattern *A*(*l*,*m*) is not uniform

- One needs to correct for the direction-dependent sensitivity
- Luckily, it is usually simple: dividing *I*(*l*,*m*) by *A*(*l*,*m*) in the image plane is enough



Reading (and watching) material

- **Condon, J. & Ransom, S.: "*Essential Radio Astronomy*" (see Chapter 3)**
  - *https://science.nrao.edu/opportunities/courses/era*
- **Thompson, A.R., Moran, J.M. & Swenson, G.W.: *"Interferometry and Synthesis in Radio Astronomy" (3rd edition 2017)***
  - *Open access: http://www.springer.com/in/book/9783319444291*
- **Taylor, G. B., Carilli, C. L. & Perley, R. A.: "*Synthesis Imaging in Radio Astronomy II" ASP Conference Series Vol. 180 (1999)***
  - *Contents available online*
- **NRAO Synthesis imaging school 2014 lecture videos are online**
  - *https://science.nrao.edu/science/meetings/2014/14th-synthesis-imaging-workshop*