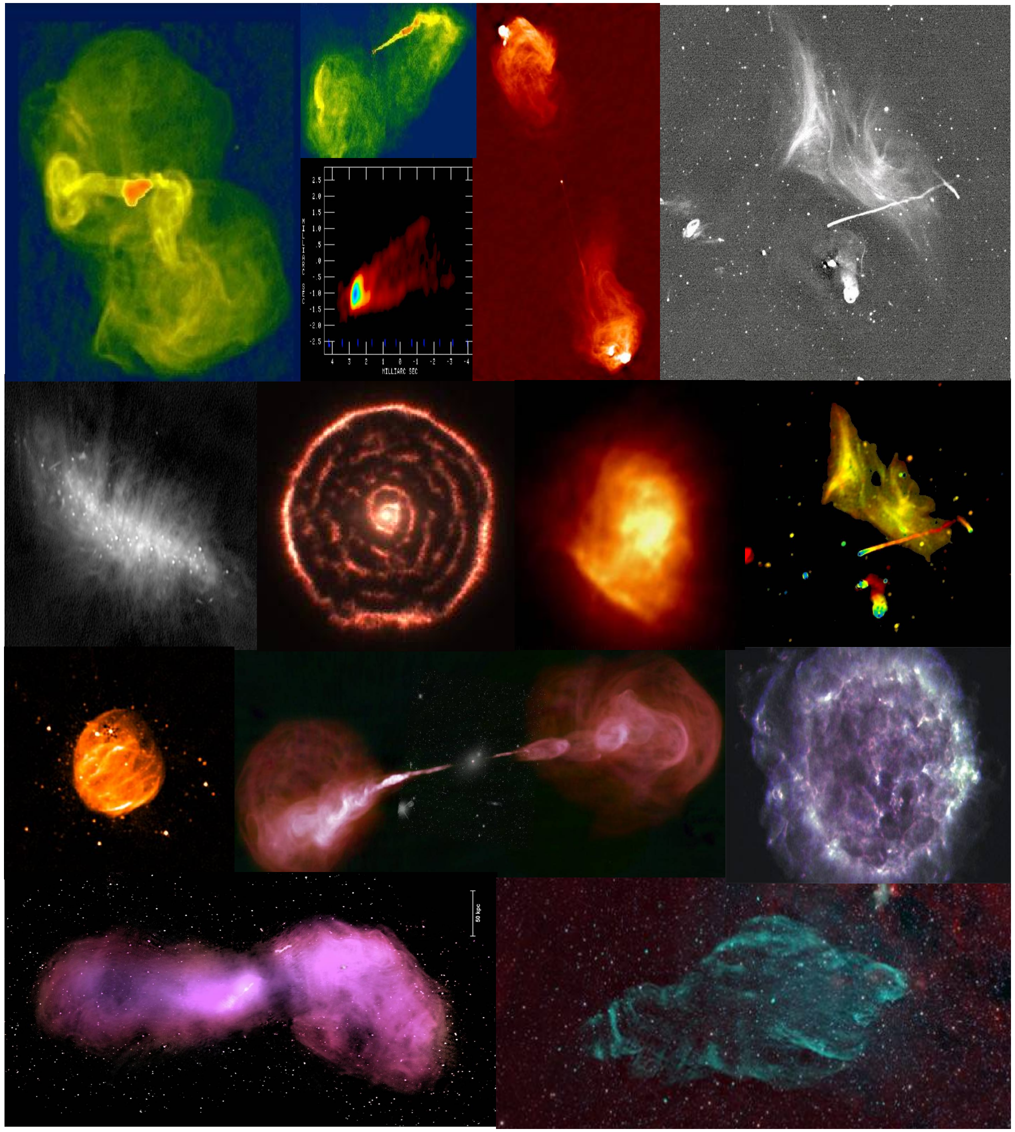
Imaging and Deconvolution

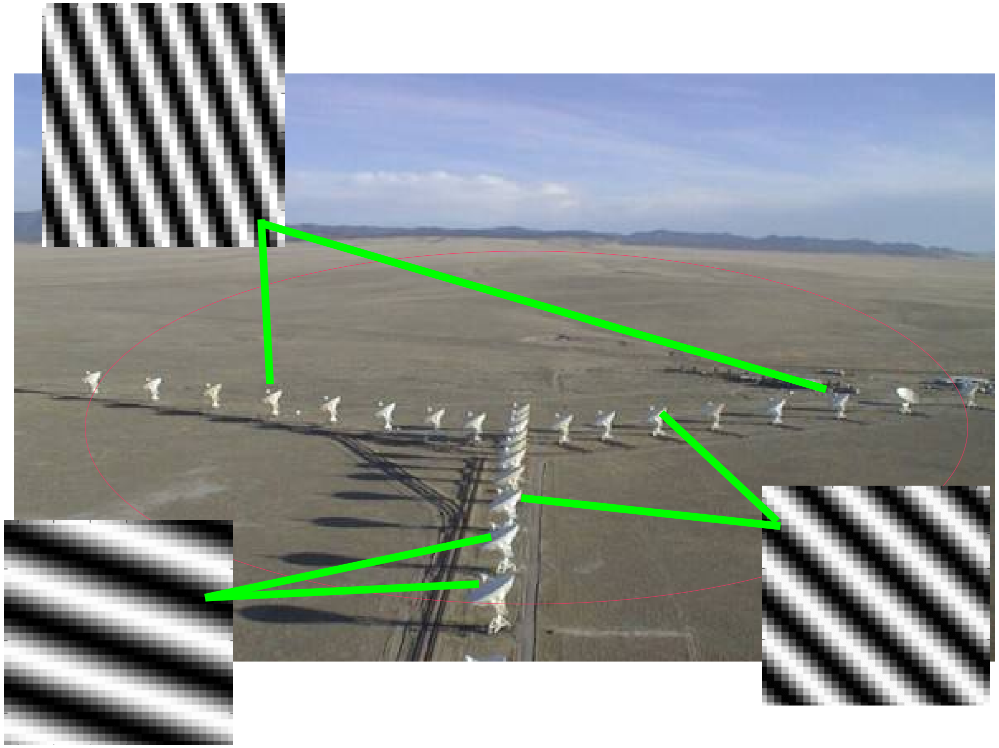


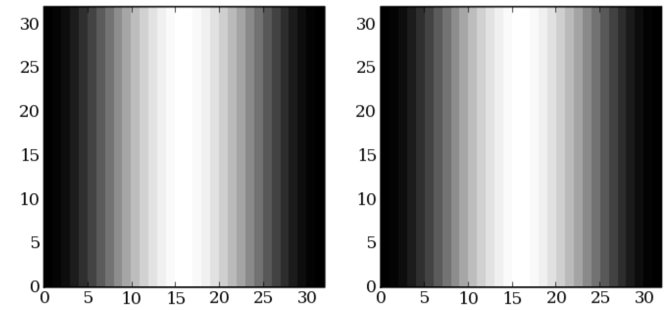
Urvashi Rau

National Radio Astronomy Observatory, Socorro, NM, USA

The van-Cittert Zernike theorem

$$\langle E_i E_j^* \rangle \propto V_{ij}(u,v) = \iint I^{sky}(l,m)\, e^{2\pi i (ul+vm)}\, dl\, dm$$





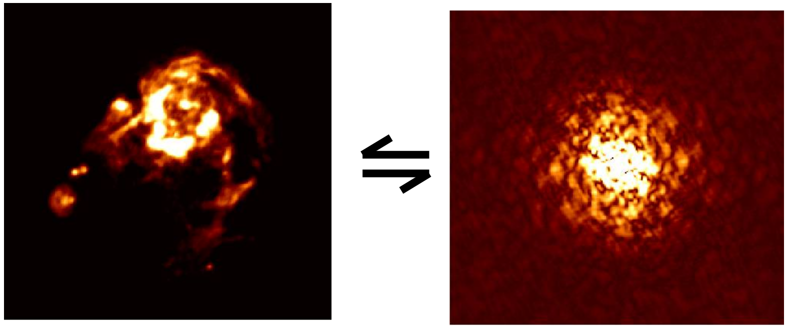
2D Fourier transform :

Image = sum of cosine 'fringes'.

Each antenna-pair measures the parameters of one 'fringe'.

Amplitude, Phase : $\langle E_i E_j^* \rangle$ is complex.

Orientation, Wavelength : Geometry

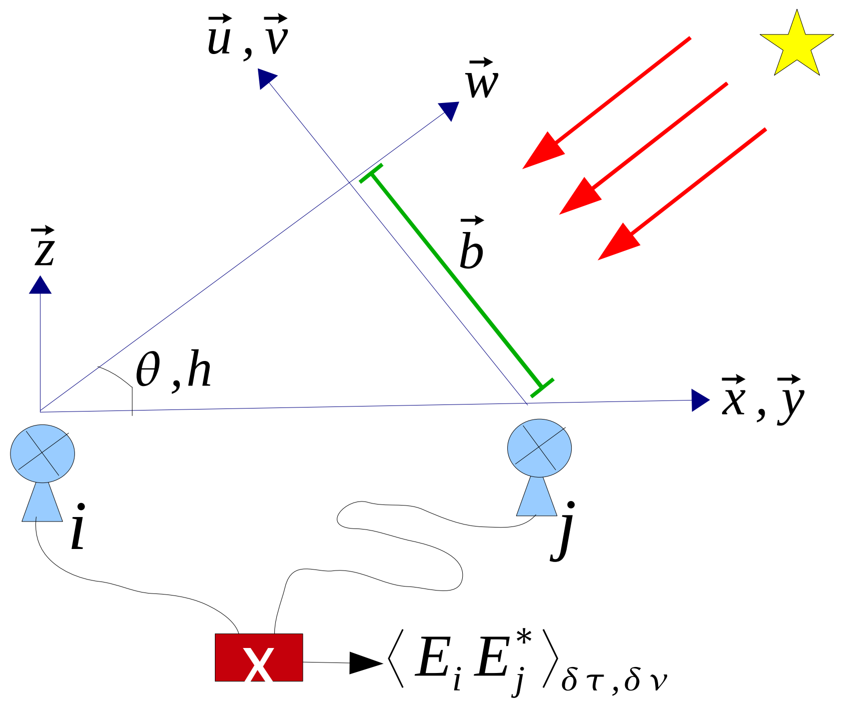


$I^{sky}(l,m)$ $\rightleftharpoons$ $V(u,v)$

Sky Coordinates — Spatial Frequencies

Measuring the visibility function

Spatial Frequency : Length and orientation of the vector between two antennas, projected onto the plane perpendicular to the line of sight.



$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

For each antenna pair, u, v change with time (hour-angle, declination) and observing frequency.

Time and Frequency-resolution of the data samples $\delta\tau, \delta\nu$ decides $\delta u, \delta v$

N antennas
N(N-1)/2 antenna-pairs (baselines)

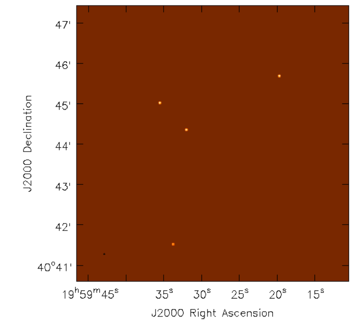
Image is real => Visibility function is Hermitian : $V(u,v) = V^*(-u,-v)$

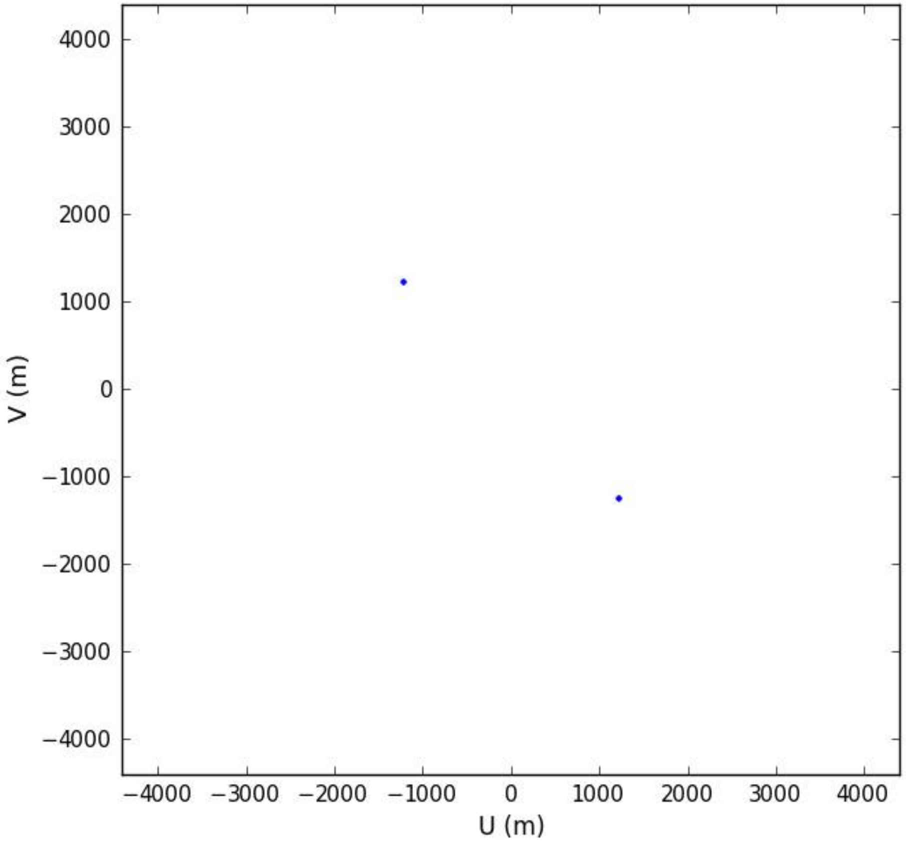
=> One baseline : 2 visibility points

Spatial Frequency (uv) coverage + Observed Image

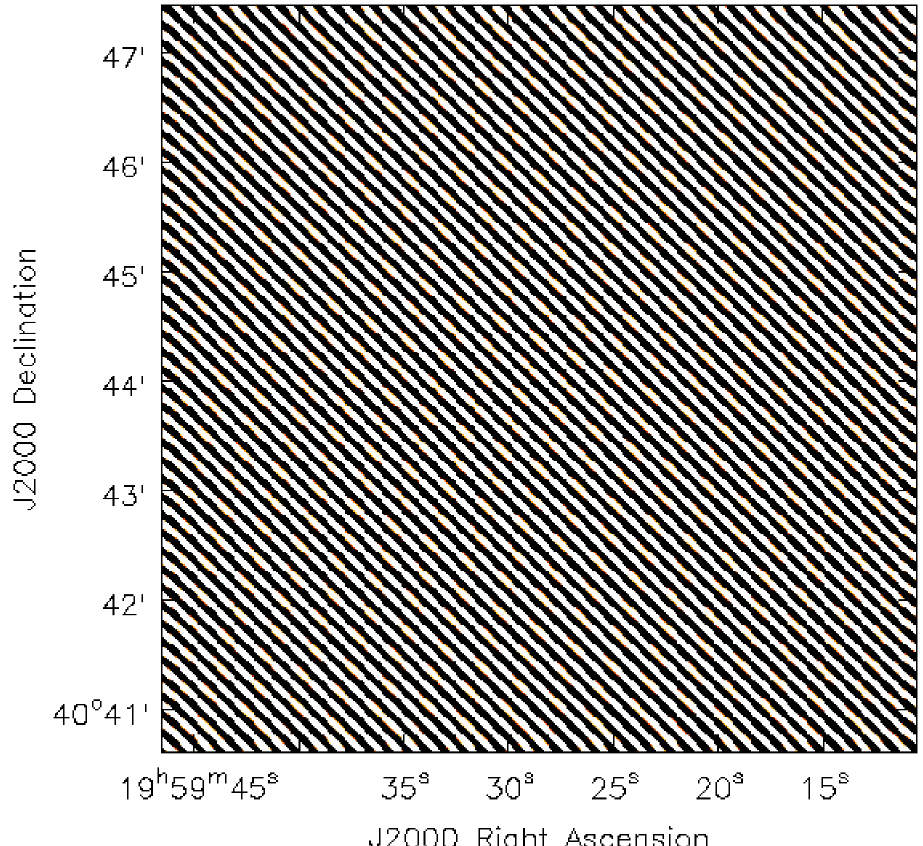
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 2 antennas





$$S(u,v)$$

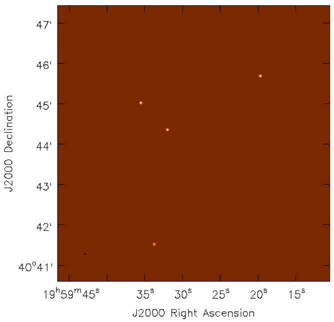


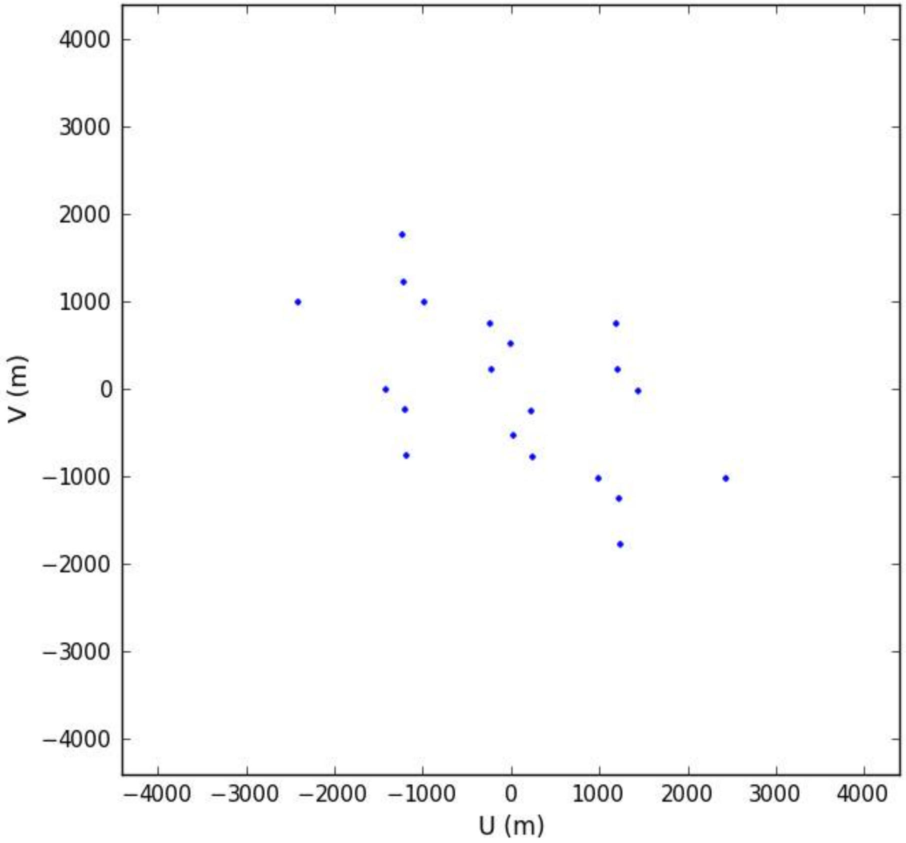
$$I^{obs}(l,m)$$

Spatial Frequency (uv) coverage + Observed Image

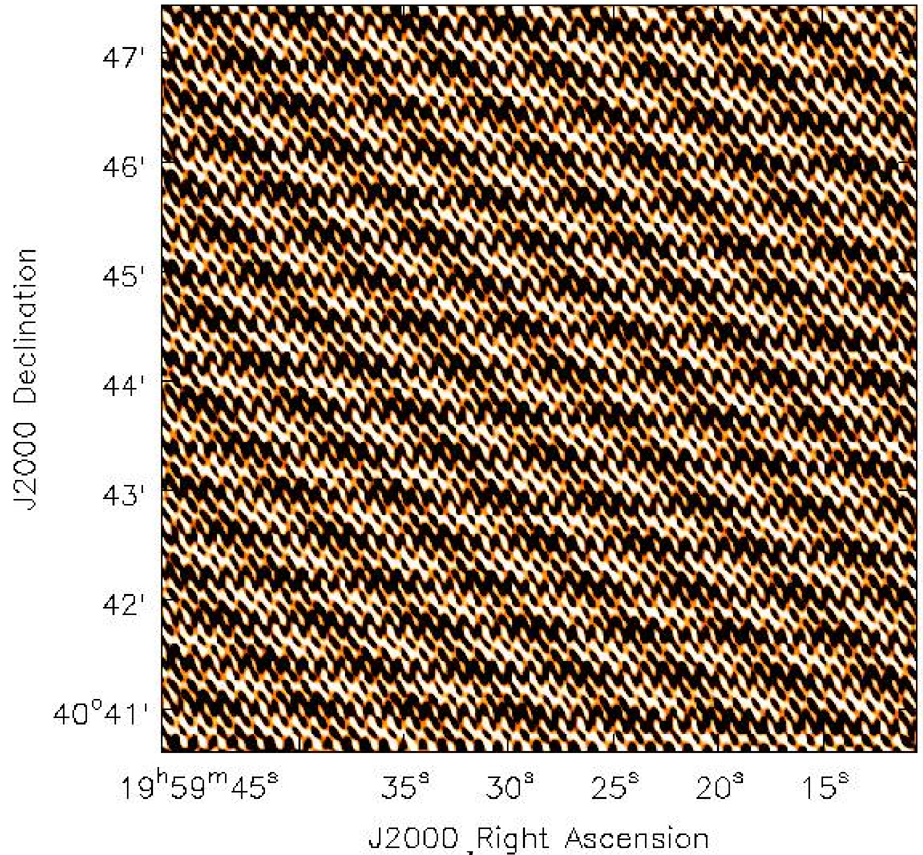
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \left[R(h,\theta) \right] \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 5 antennas





$$S(u,v)$$

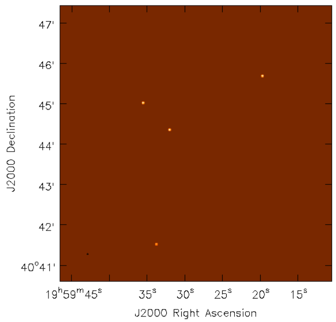


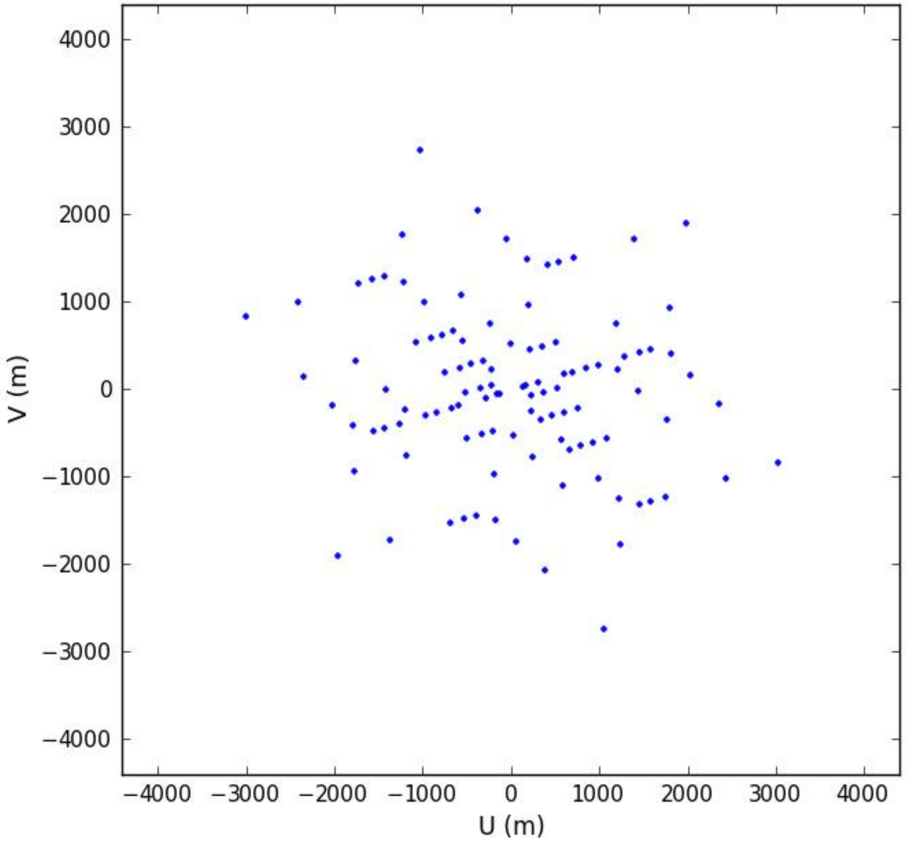
$$I^{obs}(l,m)$$

Spatial Frequency (uv) coverage + Observed Image

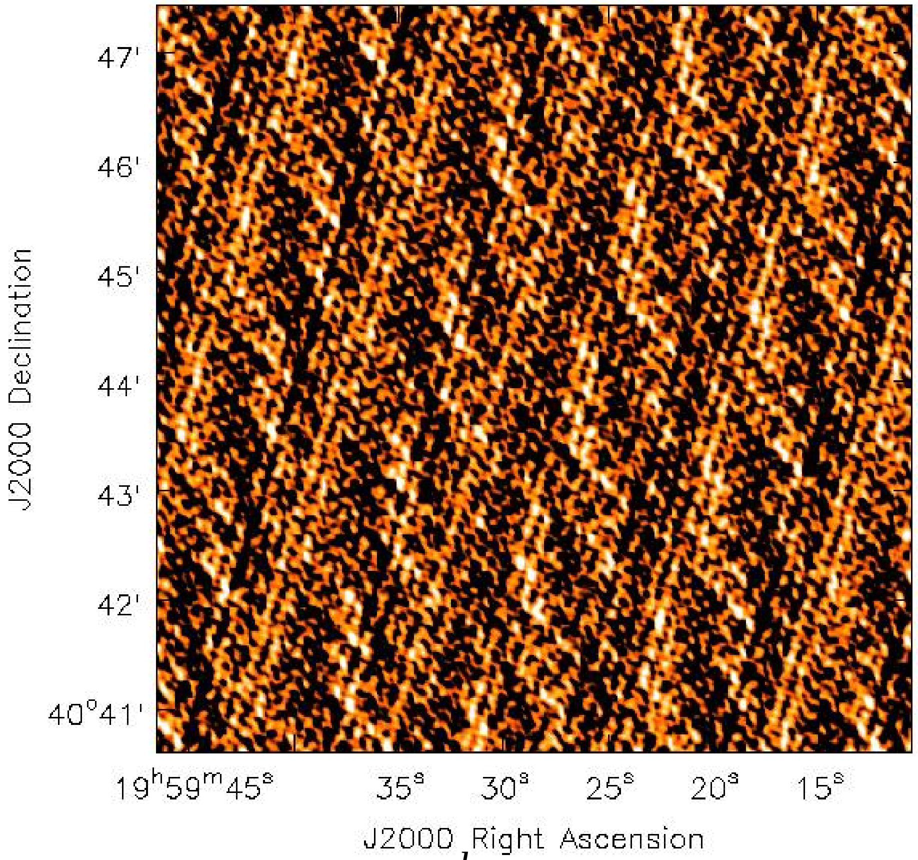
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h, \theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 11 antennas





$S(u, v)$

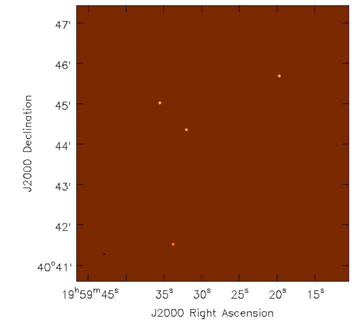


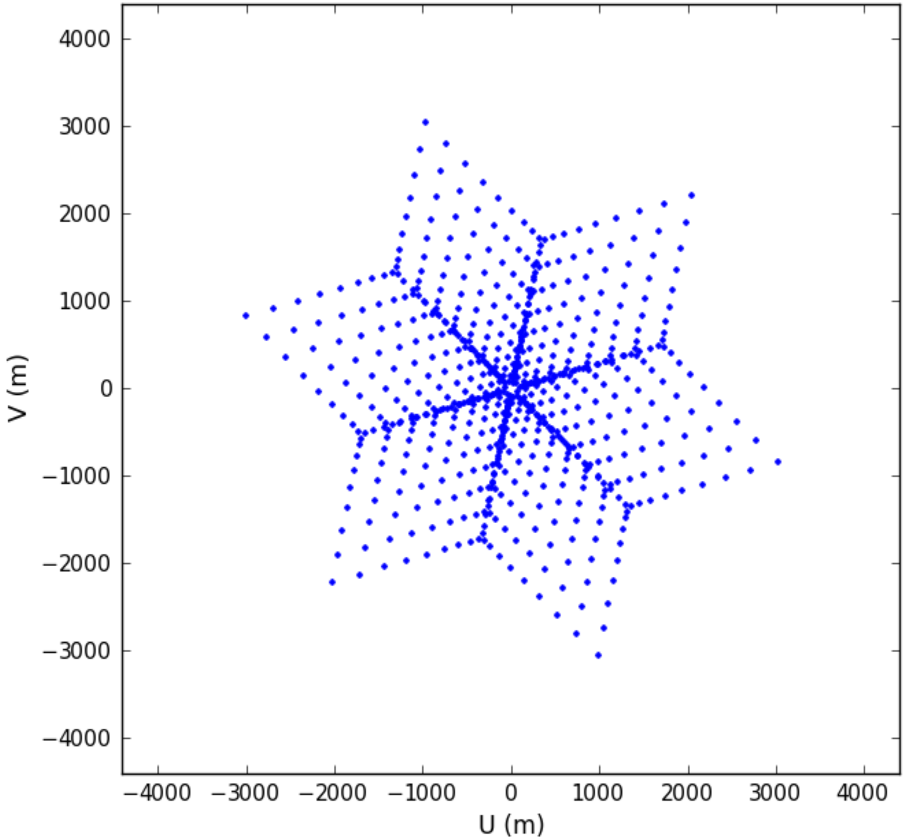
$I^{obs}(l, m)$

Spatial Frequency (uv) coverage + Observed Image

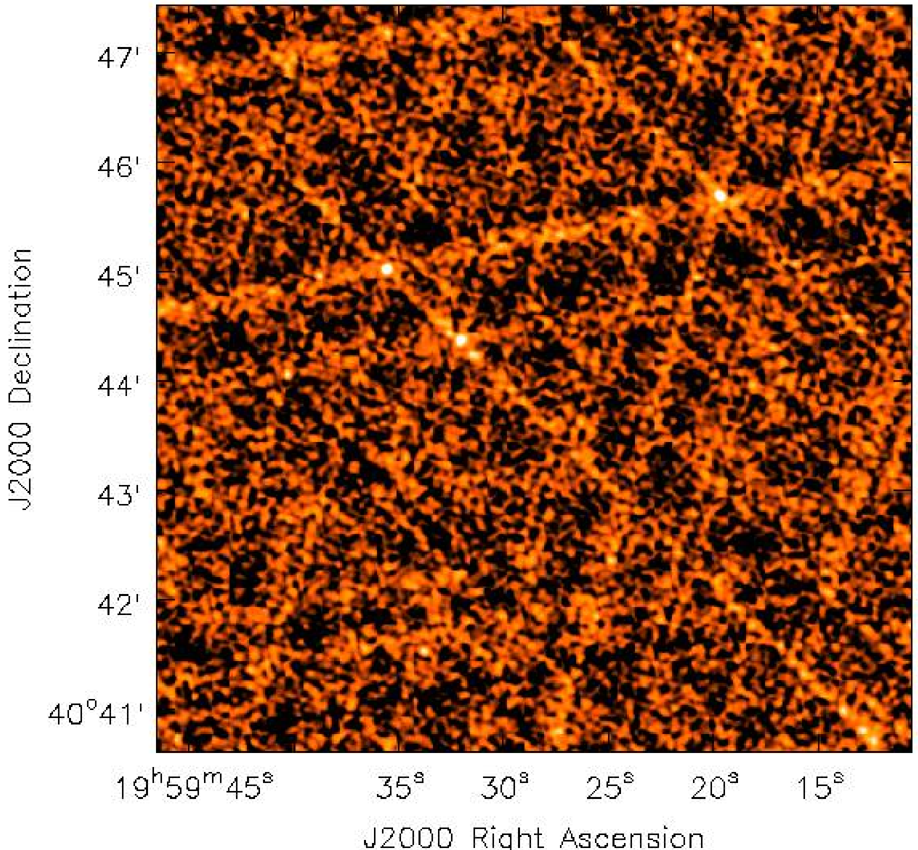
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 27 antennas





$$S(u,v)$$

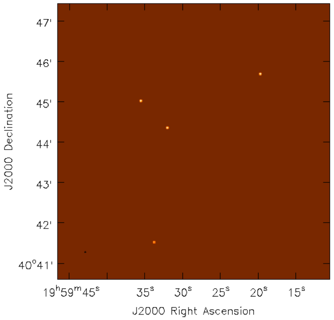


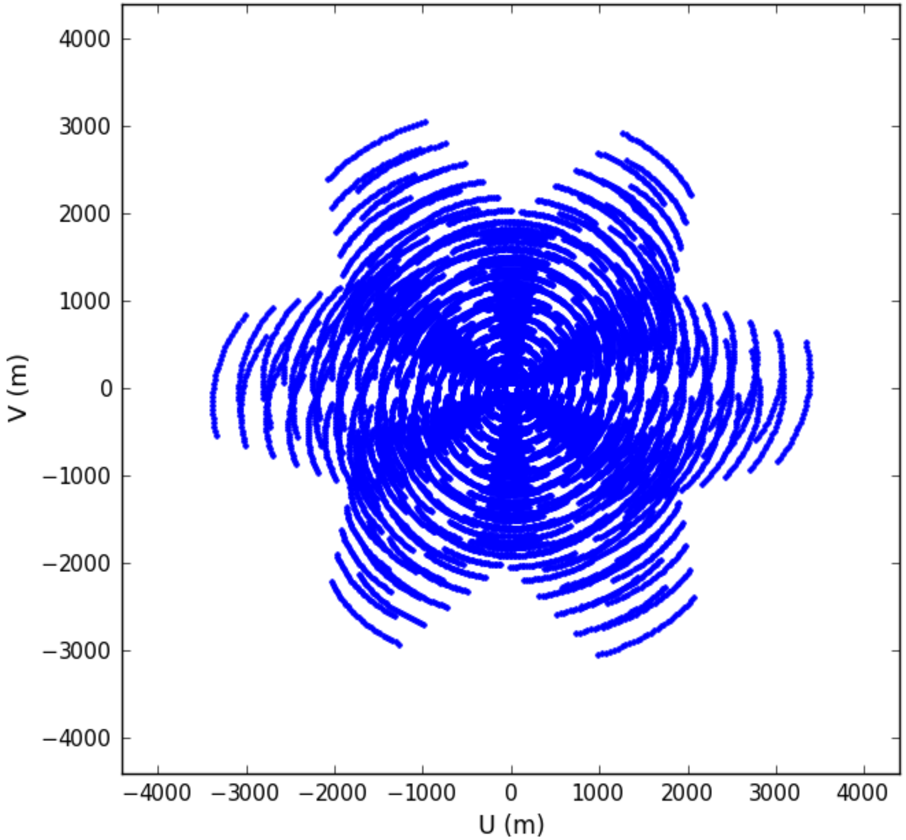
$$I^{obs}(l,m)$$

Spatial Frequency (uv) coverage + Observed Image

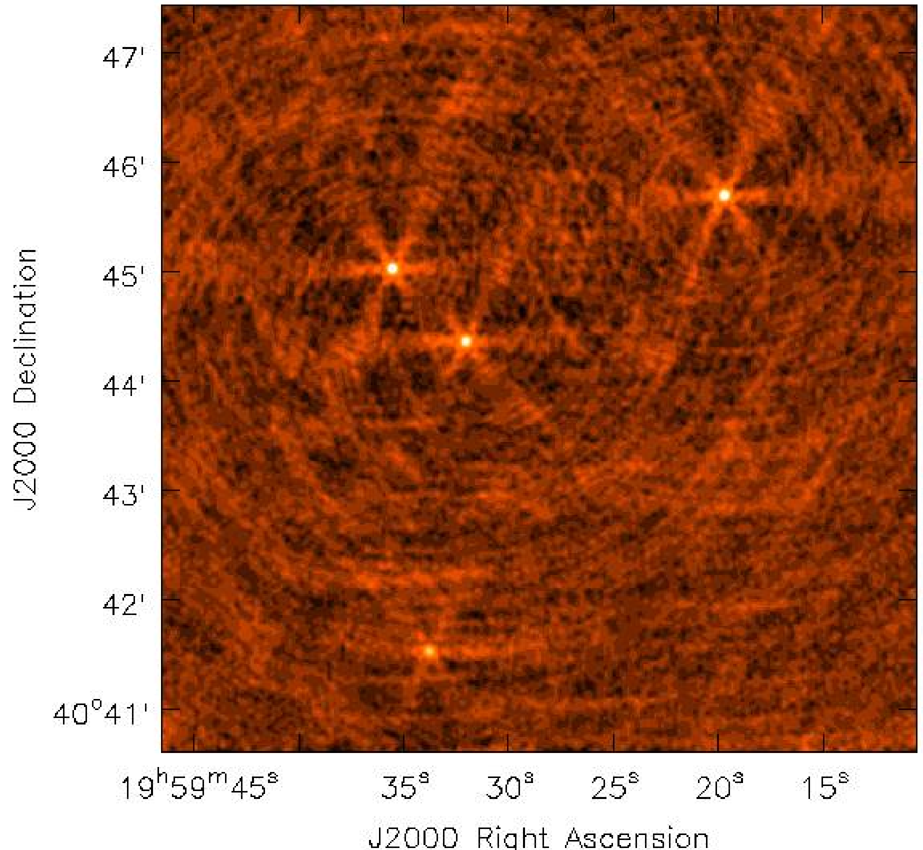
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \left[R(h,\theta) \right] \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky
using 27 antennas
over 2 hours
'Earth Rotation Synthesis'





$S(u,v)$

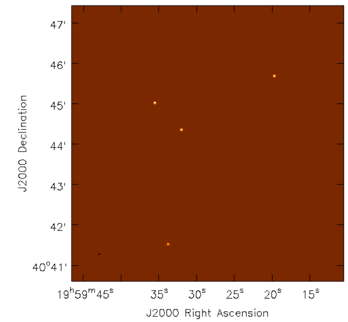


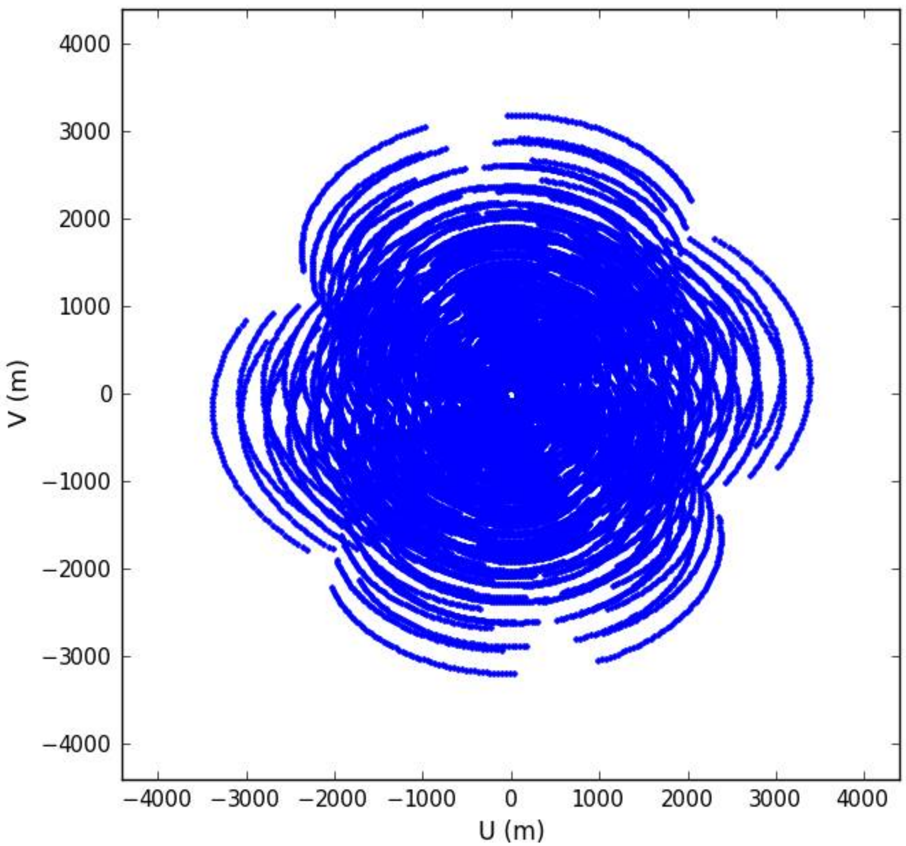
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

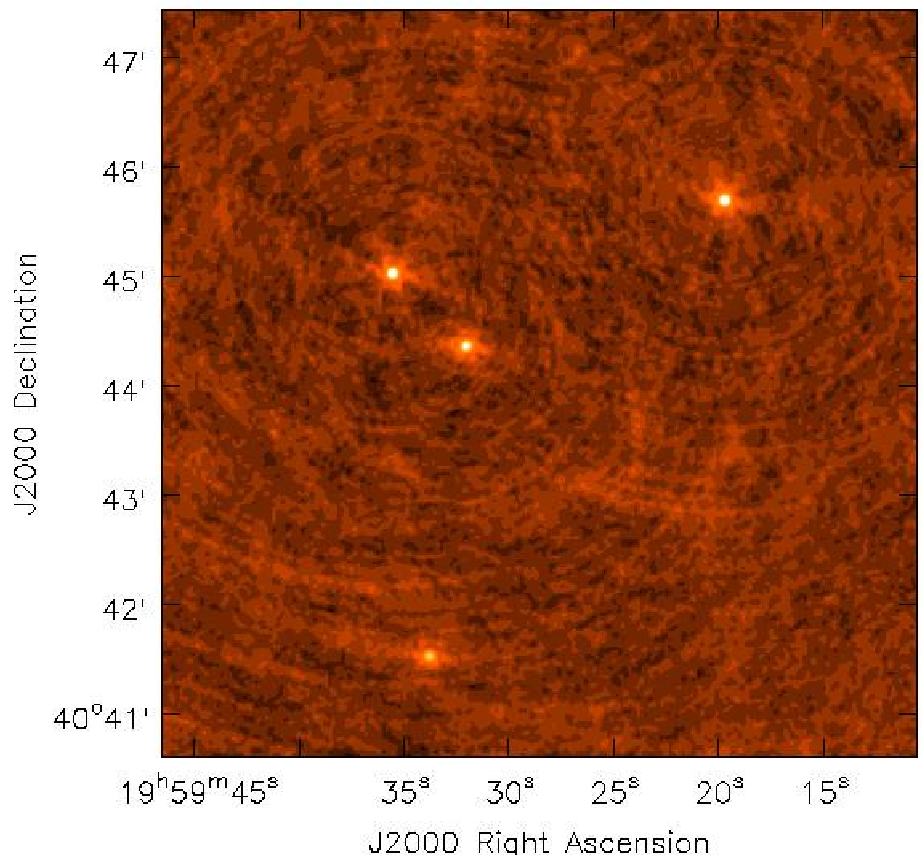
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky
using 27 antennas
over 4 hours
'Earth Rotation Synthesis'





$S(u,v)$

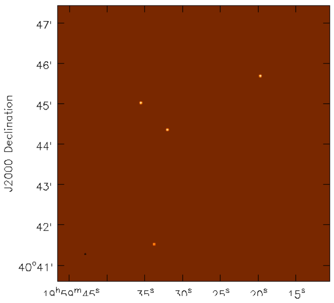


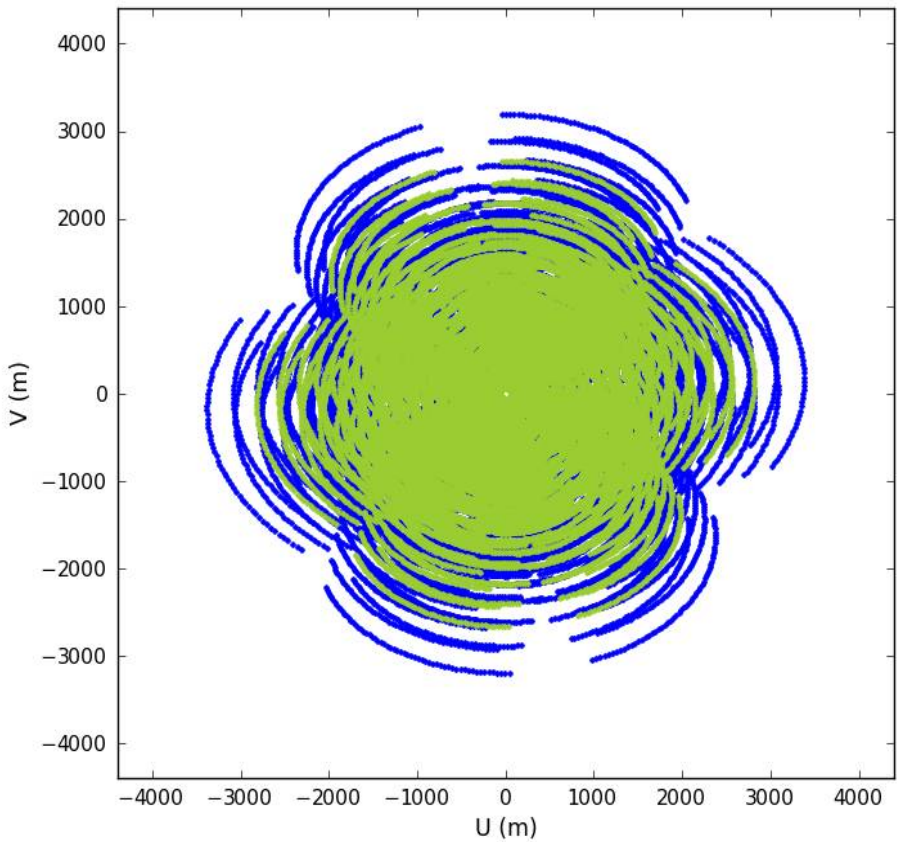
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

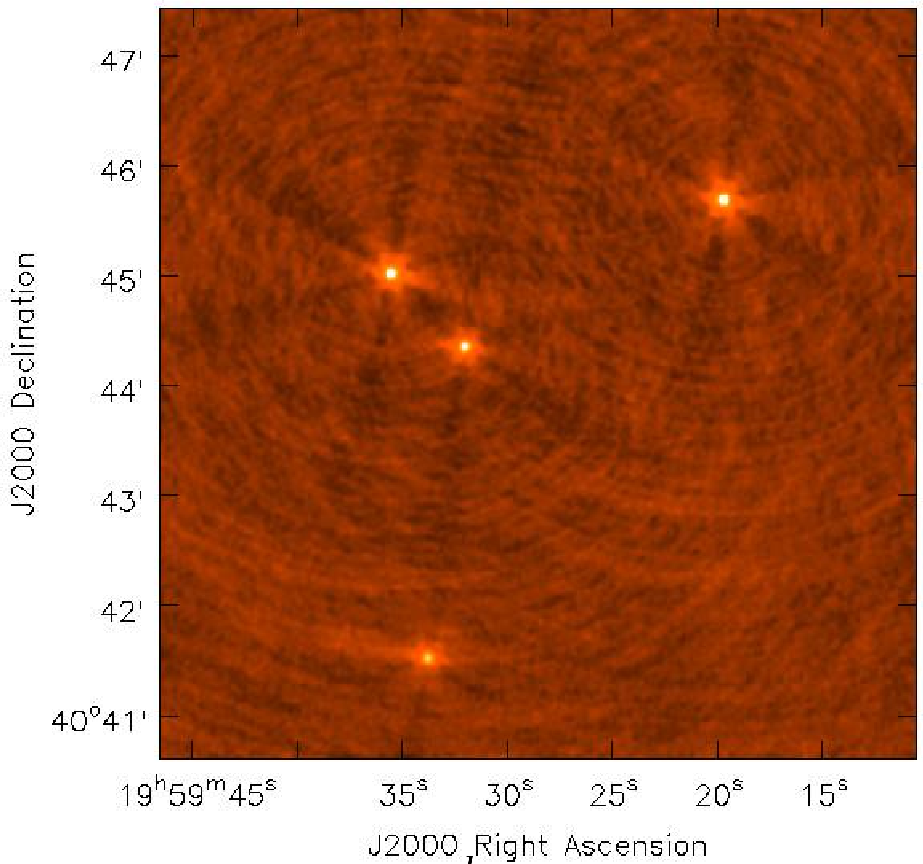
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky
using 27 antennas
over 4 hours, 2 frequencies
'Multi-Frequency Synthesis'





$S(u,v)$

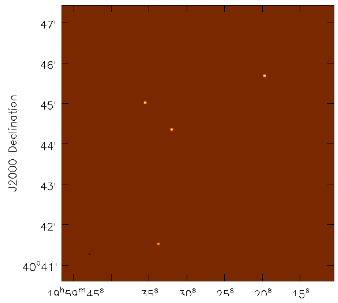


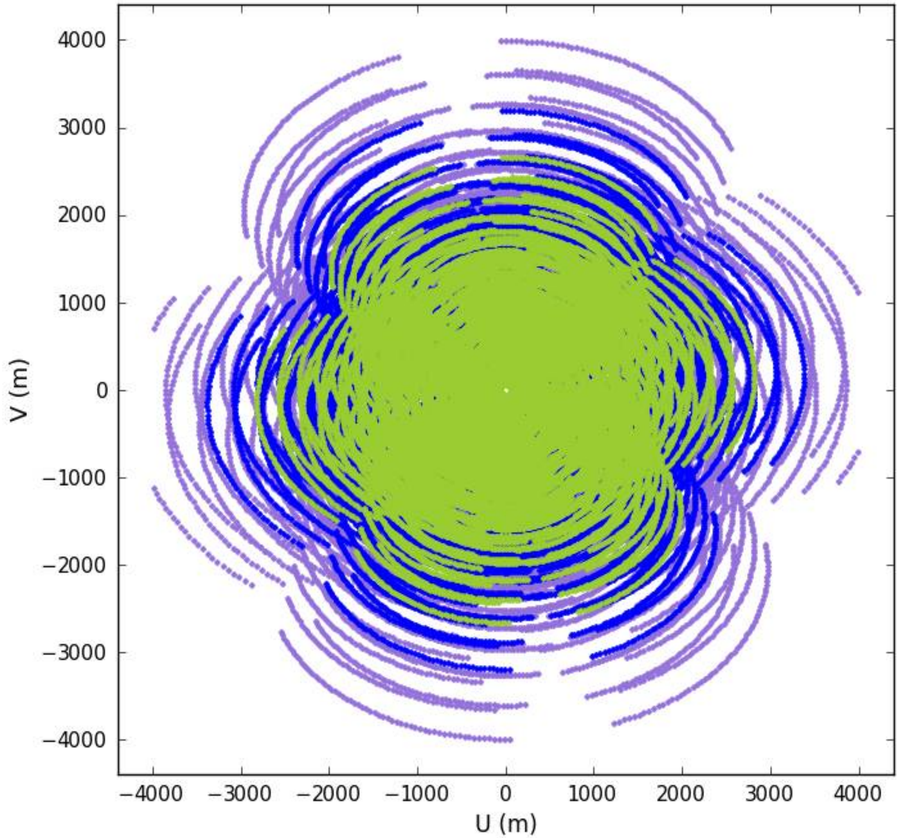
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

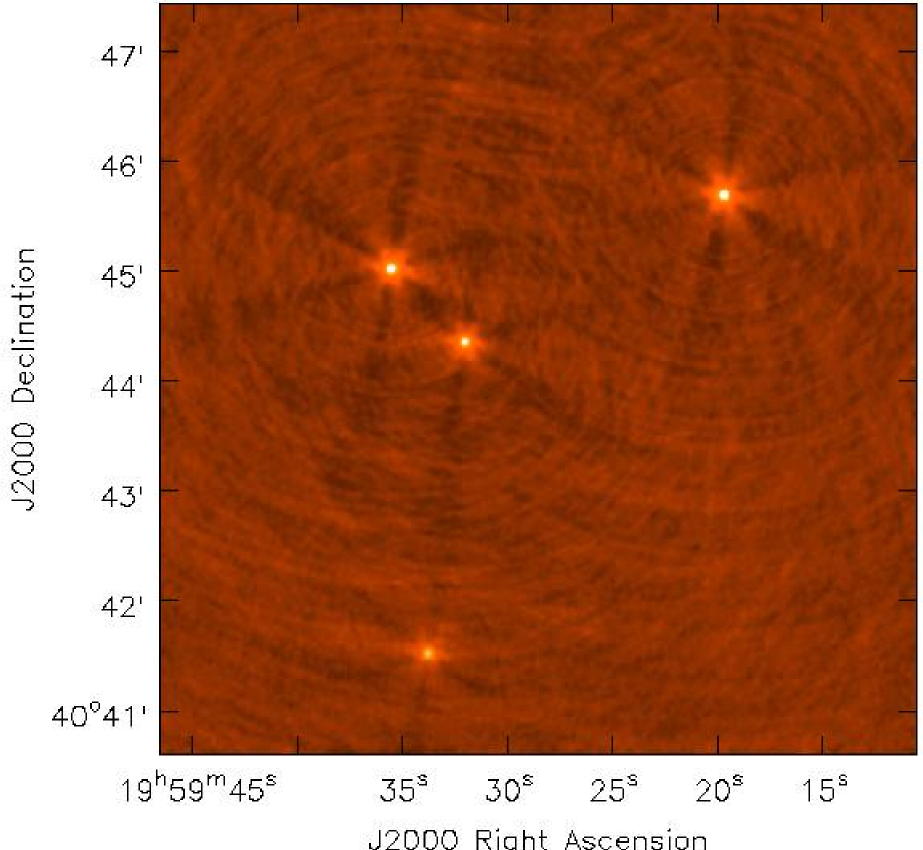
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky
using 27 antennas
over 4 hours, 3 frequencies
'Multi-Frequency Synthesis'





$S(u,v)$

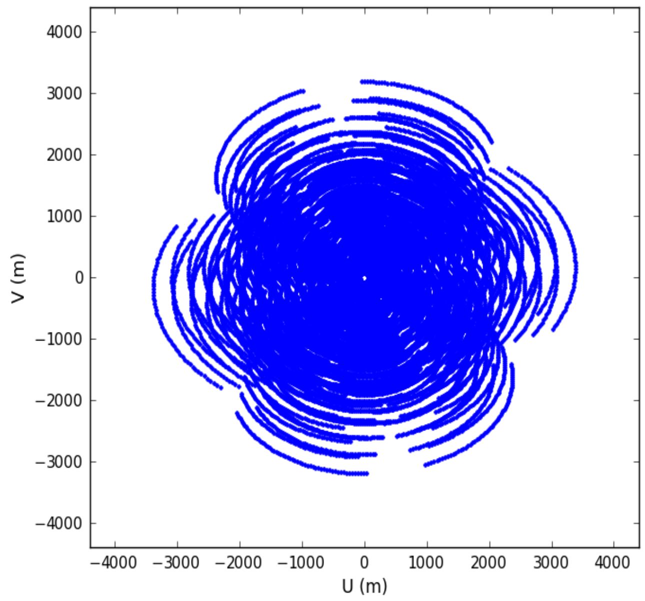


$I^{obs}(l,m)$

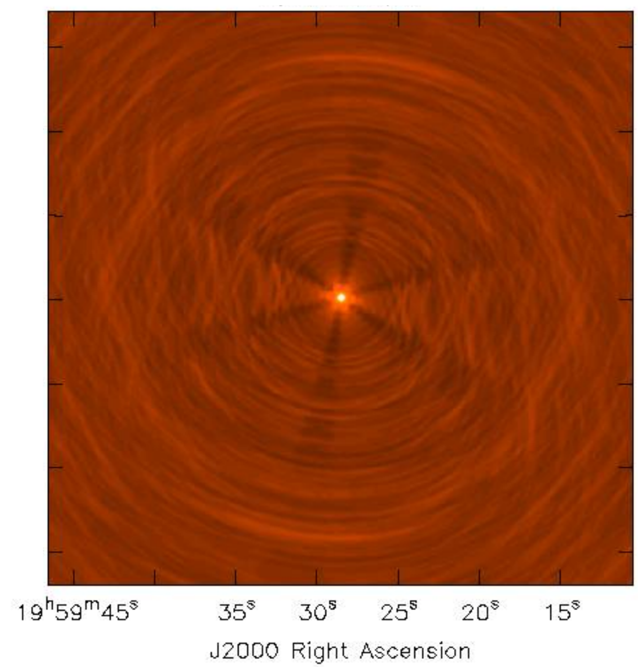
Point Spread Function (PSF)

$$S(u,v) \qquad I^{psf}(l,m) = F^{-1}[S(u,v)]$$

The PSF is the inverse Fourier transform of the UV-coverage



$\rightleftharpoons$



The PSF is

--- the impulse-response of the instrument (image of a point-source)

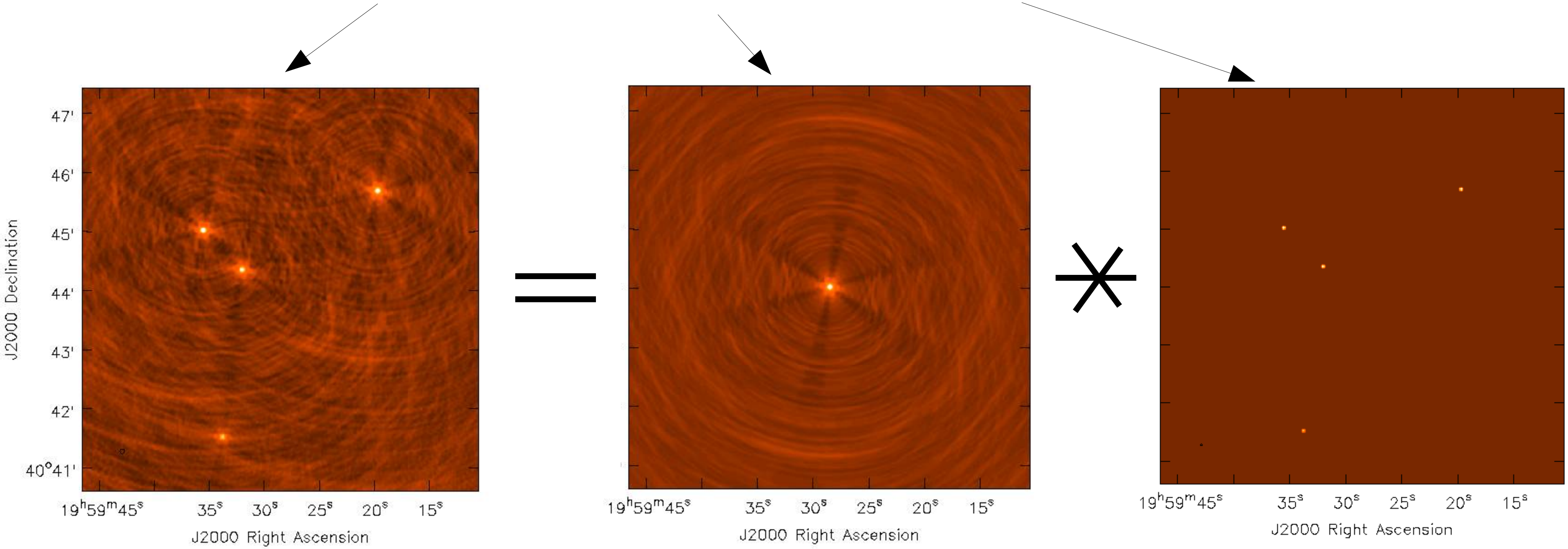
--- the intensity of the diffraction pattern through an array of 'slits' (dishes)

--- a measure of the imaging-properties of the instrument

angular resolution, (max uv-spacing)	peak sensitivity, (# measurements)	sidelobe levels, (missing spacings)	no total power (central uv-hole)

Image formed by an interferometer : Convolution Equation

$$I^{obs}(l,m) = I^{PSF}(l,m) * I^{sky}(l,m)$$



=> You have measured the Convolution of the True Sky with the instrumental PSF.

=> Recovering the True Sky ---------> DE-convolution

First Step : Construct the PSF image, and the OBSERVED image....

Imaging in practice : Choosing image size, cell-size

– Choosing image 'cell' size : Nyquist-sample the main lobe of the PSF

PSF beam width : $\frac{\lambda}{b_{max}} = \frac{1}{u_{max}}$ $radians$(x $\frac{180}{\pi}$ to convert to degrees)

This is the diffraction-limited angular-resolution of the telescope
Ex : Max baseline : 10 km. Freq = 1 GHz. Angular resolution : 6 arcsec

– Choosing image field-of-view (npixels) : As much as desired/practical.

$$\frac{1}{fov_{rad}} = \delta u$$

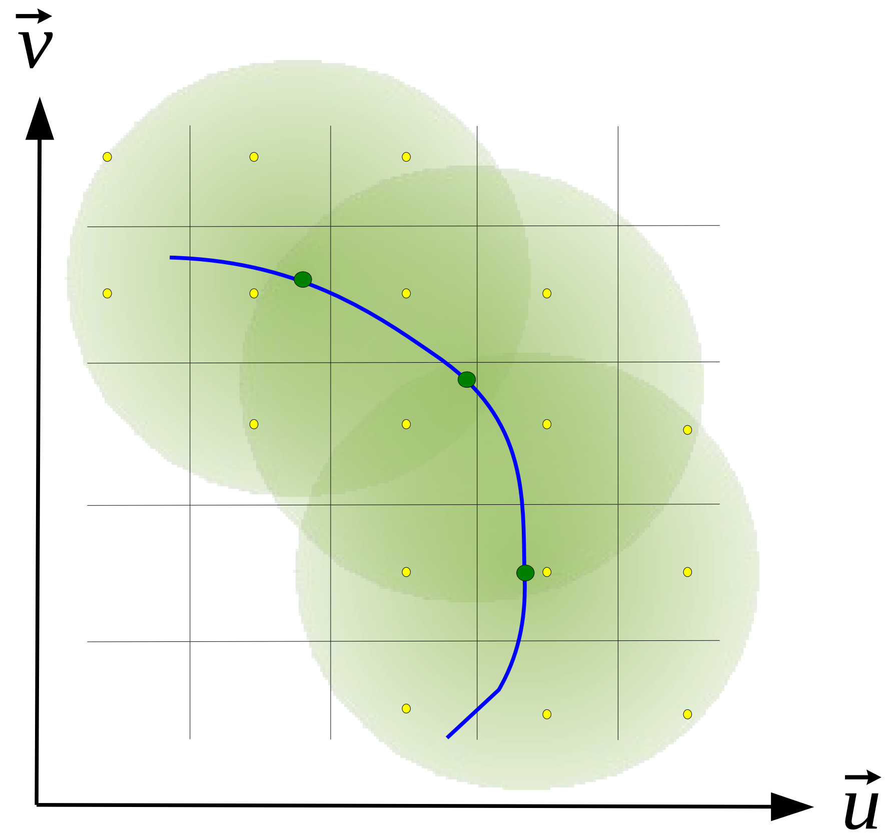
Field of View (fov) controls the uv-grid-cell size $(\delta u, \delta v)$

- Antenna primary-beam limits the field-of-view ('slits' of finite width)

– Gridding + FFT :

- An interferometer measures irregularly spaced points on the UV-plane.
- Need to place the visibilities onto a regular grid of UV-pixels, and then take an FFT

Imaging in practice : Gridding and Weighting

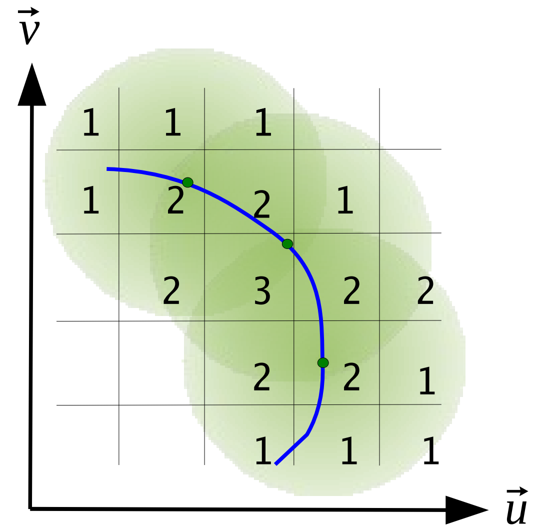


-- Visibility data are recorded onto a regular grid before taking an i-FFT

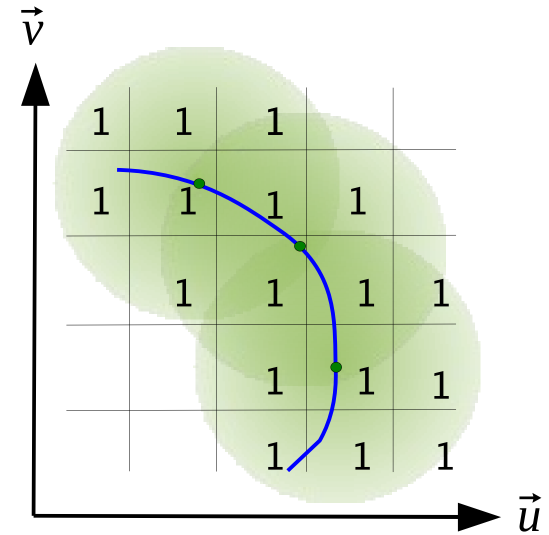
– Convolutional Resampling

=> Use a gridding convolution function
=> Use weights per visibility
(weighted average of all data points per cell)

An Image is a weighted-average of the data.



Natural Weights

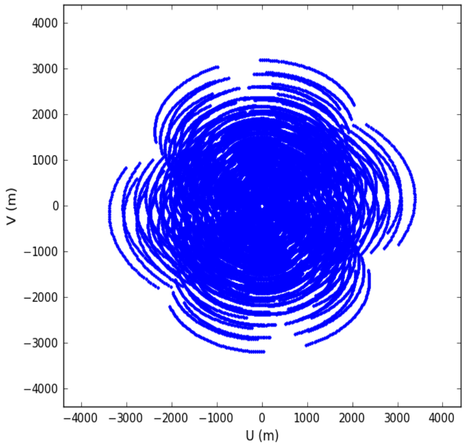


Uniform Weights

Imaging in practice : Weighting schemes

An Image is a weighted average of the data.

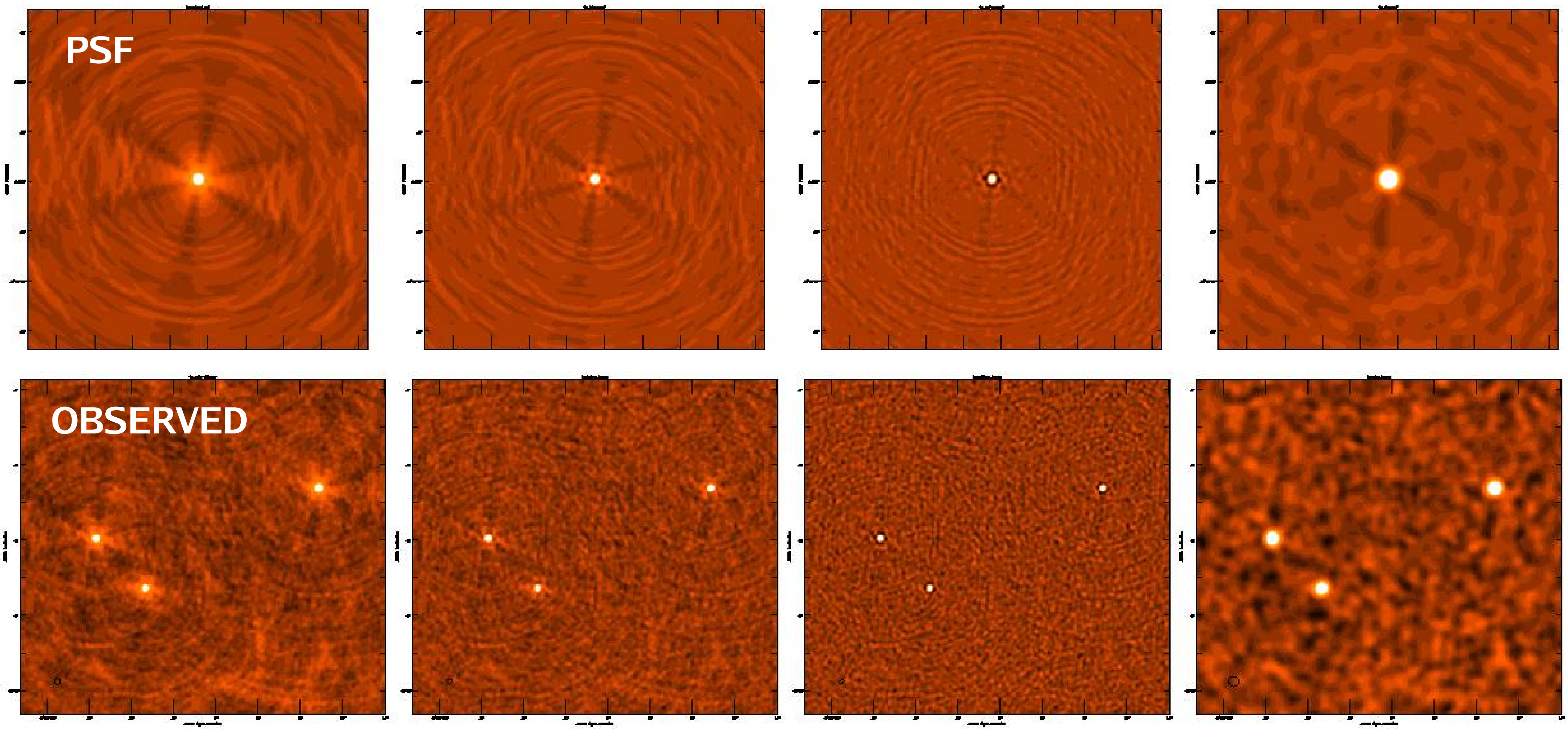
Choose a weighting-scheme => modify the imaging properties of the instrument
=> emphasize features/scales of interest
=> control imaging sensitivity



	Uniform/Robust All spatial-frequencies get equal weight	Natural/Robust All data points get equal weight	UV-Taper Low spatial freqs get higher weight than others
Resolution	higher	medium	lower
PSF Sidelobes (VLA)	lower	higher	depends
Point Source Sensitivity	lower	maximum	lower
Extended Source Sensitivity	lower	medium	higher

Imaging in practice : PSFs and Observed (dirty) Images

Natural	Robust 0.7	Uniform	Tapered Uniform
Bm : 5.6 arcsec 0.1 sidelobe	Bm : 4.0 arcsec 0.05 sidelobe	Bm : 3.2 arcsec +0.03,-0.08 sidelobe	Bm : 8.0arcsec 0.01 sidelobe

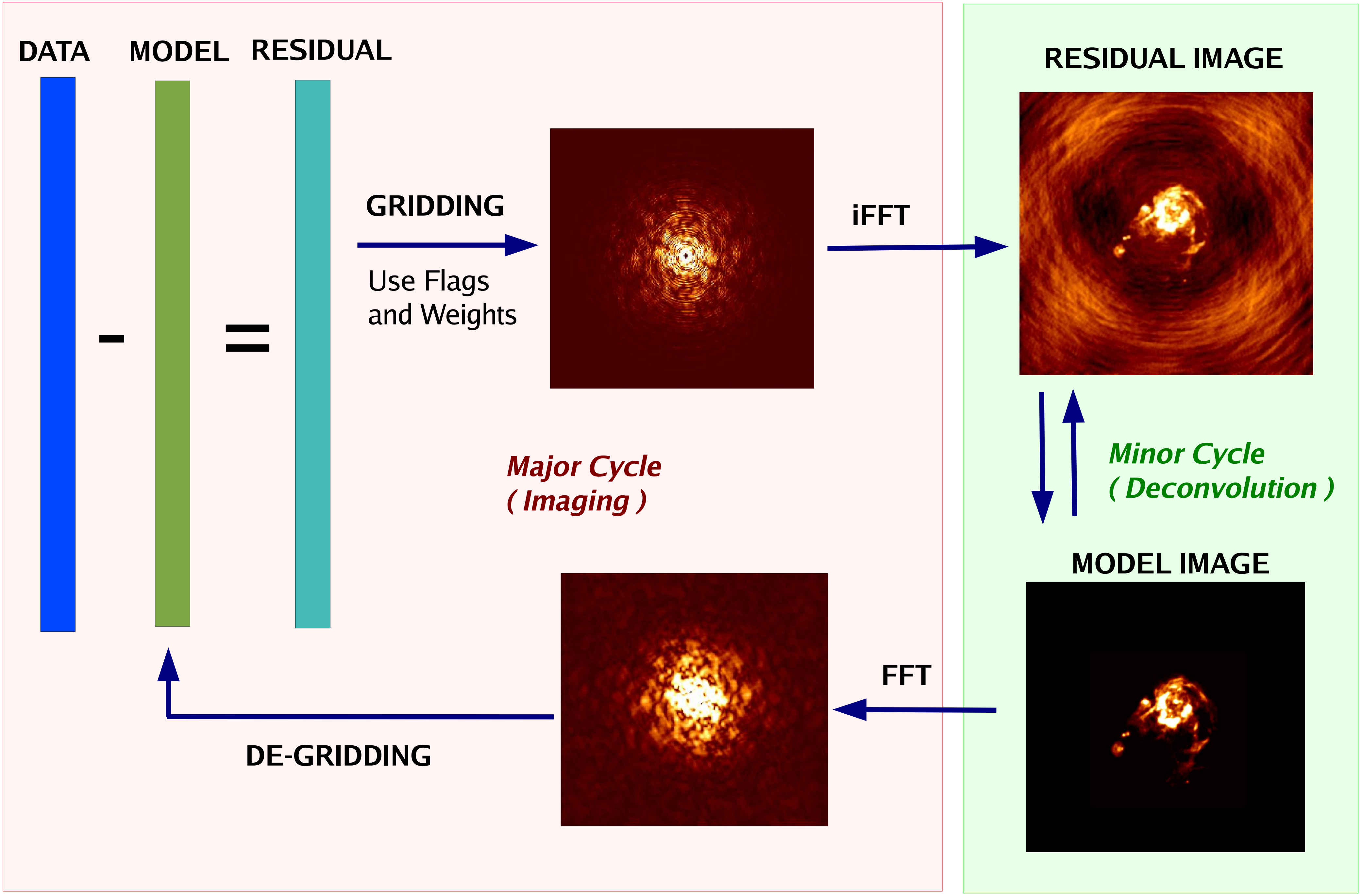


Note the noise-structure. Noise is correlated between pixels by the PSF.

Image Units (Jy/beam)

------ All pairs of images satisfy the convolution relation => Need to deconvolve them

Iterative Image Reconstruction – Major and Minor Cycles



Deconvolution – Hogbom CLEAN

Sky Model : List of delta-functions

(1) Construct the observed (dirty) image and PSF

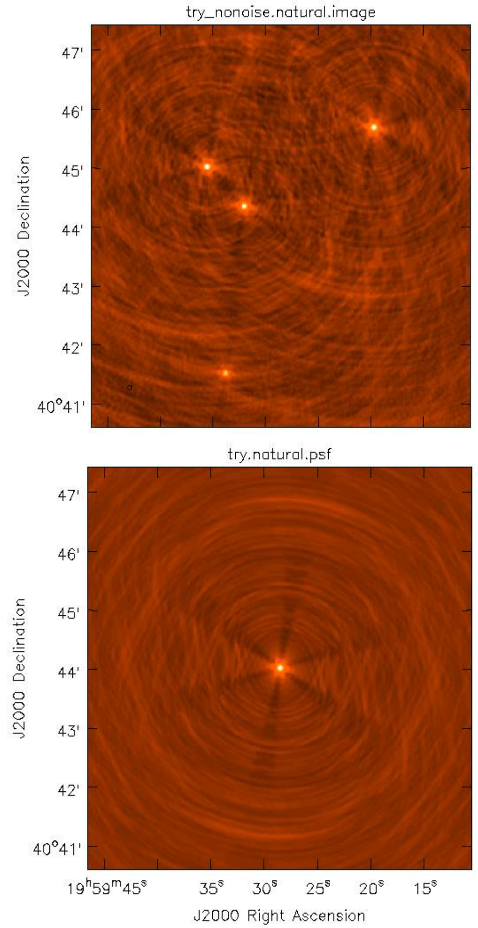
(2) Search for the location of peak amplitude.

(3) Add a delta-function of this peak/location to the model

(4) Subtract the contribution of this component
from the dirty image - a scaled/shifted copy of the PSF

Repeat steps (2), (3), (4) until a stopping criterion is reached.

(5) Restore : Smooth the model with a 'clean beam' and add residuals



The CLEAN algorithm can be formally derived as a model-fitting problem

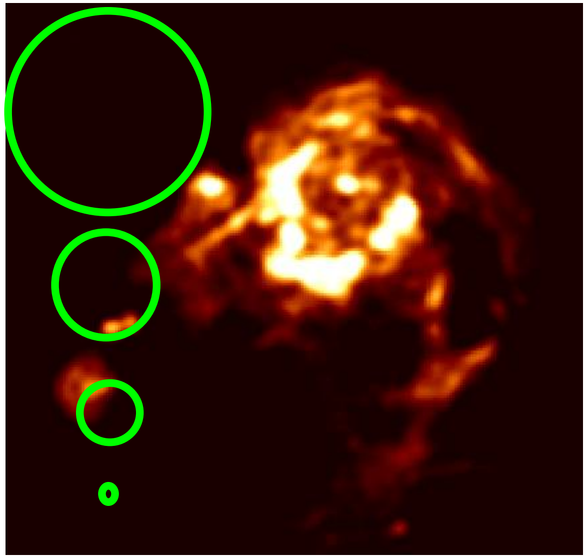
– model parameters : locations and amplitudes of delta functions

– solution process : χ^2 minimization via an iterative steepest-descent algorithm
(method of successive approximation)

Deconvolution – MS-CLEAN

Multi-Scale Sky Model : Linear combination of 'blobs' of different scale sizes

- Efficient representation of both compact and extended structure (sparse basis)



A scale-sensitive algorithm

(1) Choose a set of scale sizes

(2) Calculate dirty/residual images
smoothed to several scales (basis functions)

– Normalize by the relative sum-of-weights
(instrument's sensitivity to each scale)

(3) Find the peak across all scales, update a single multi-scale model as well as all residual images (using information about coupling between scales)

Iterate, similar to Classic CLEAN, and restore at the end.

The MS-CLEAN algorithm can also be formally derived as a model-fitting problem using χ^2 minimization and a basis set consisting of several 'blob' sizes.

Deconvolution – Comparison of Algorithms

CLEAN

Minimize L2
(assume sparsity
in the image)

MEM

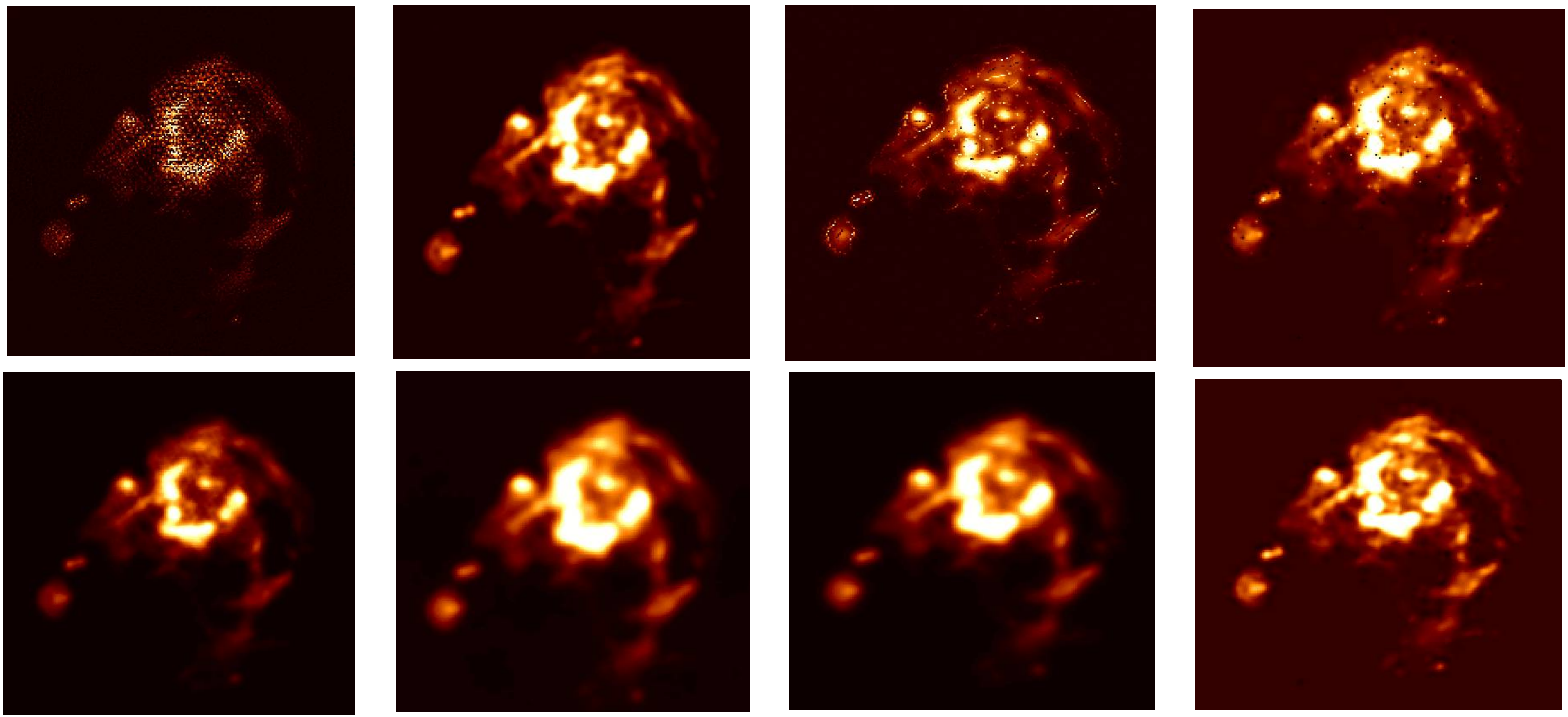
Minimize L2 subject to
an entropy-based prior
(e.g. smoothness)

MS-CLEAN

Minimize L2
(assume a set of
spatial scales)

ASP

Minimize L2 with
TV-based subspace
searches



I^m

I^{out}

Deconvolution – Comparison of Algorithms

CLEAN

Minimize L2
(assume sparsity
in the image)

MEM

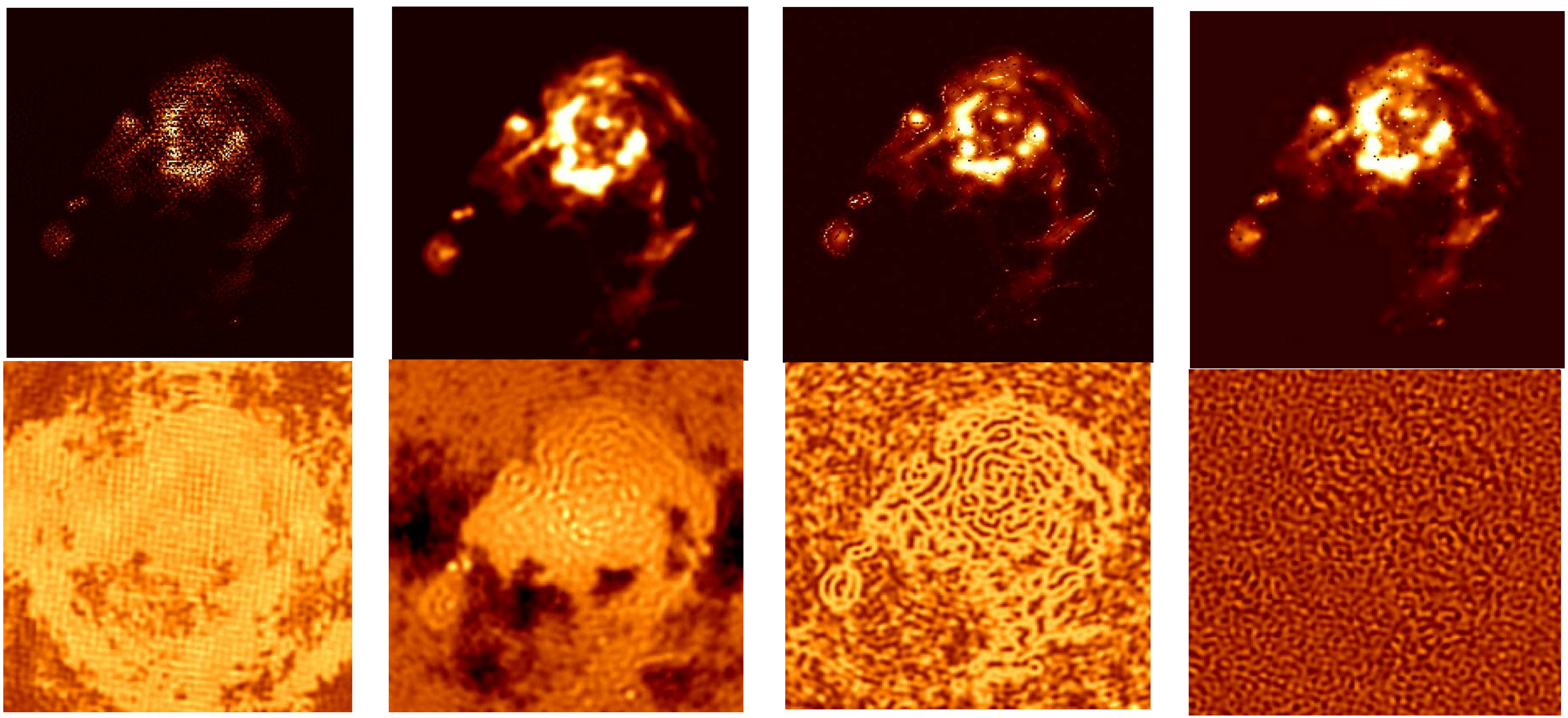
Minimize L2 subject to
an entropy-based prior
(e.g. smoothness)

MS-CLEAN

Minimize L2
(assume a set of
spatial scales)

ASP

Minimize L2 with
TV-based subspace
searches



I^m

I^{res}

Image Quality

Noise in the image : Measured from restored or residual images

– With perfect reconstruction,
  The ideal noise level is :

$$RMS \propto \frac{0.12\,\frac{T_{sys}}{\eta_a}}{\sqrt{N_{ant}(N_{ant}-1)\cdot\delta\tau\cdot\delta\nu\cdot N_{pol}}}$$

– In reality, measure the RMS of residual pixel amplitudes

Dynamic Range : Measured from the restored image

– Standard : Ratio of peak brightness to RMS noise
  in a region devoid of emission.

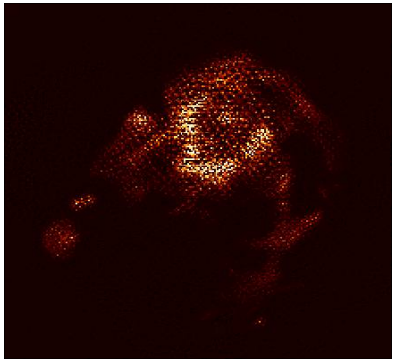
– More truthful : Ratio of peak brightness to peak error (residual)

Image Fidelity : Correctness of the reconstruction

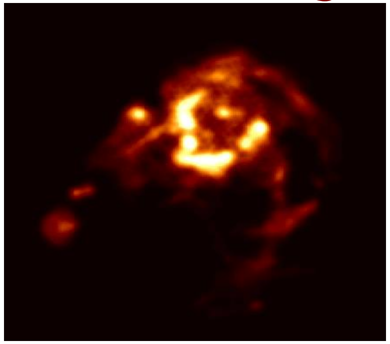
- remember the infinite possibilities that fit the data perfectly ?
- useful only if a comparison image exists.

Inverse of relative error : $$\frac{I^m * I^{beam}}{I^m * I^{beam} - I^{restored}}$$

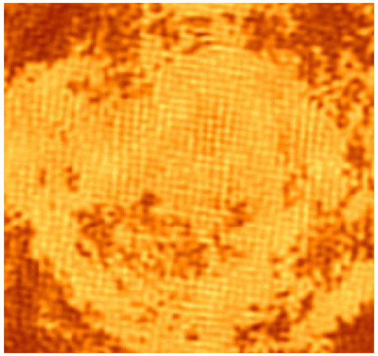
Model image



Restored image



Residual image



How can you control the quality of image reconstruction ?

(1) Iterations and stopping criterion

'niter' : maximum number of iterations / components
'threshold' : don't search for flux below this level

– minor cycles can be inaccurate, so periodically trigger major cycles

(2) Using masks

Need masks only if the deconvolution is "hard".

=> Bad PSFs with high sidelobes
=> Leftover bad data causing stripes or ripples
=> Extended emission with sharp edges
=> Extended emission that is seen only by very few baselines

Can draw them interactively (start small, and grow them) or supply final mask.

(3) Self-Calibration

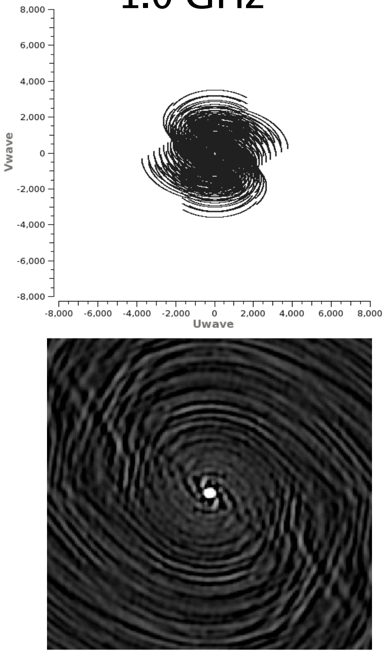
Use your current best estimate of the sky (i.e. the model image)
to get new antenna gain solutions. Apply them. Image again and repeat.

Wide-band Imaging – Sensitivity and Multi-Frequency Synthesis

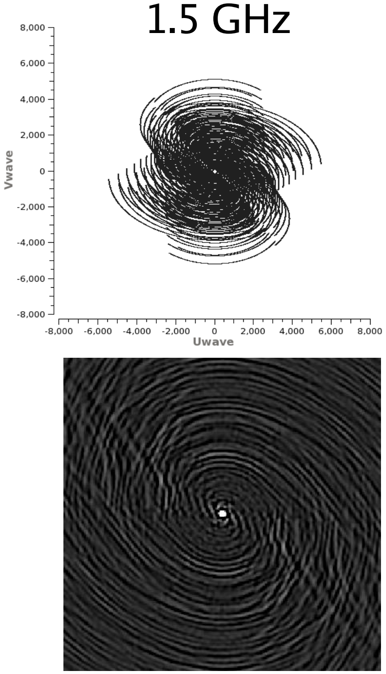
Frequency Range :	(1 –2 GHz)	(4 –8 GHz)	(8 –12 GHz)
Bandwidth : $\nu_{max} - \nu_{min}$	1 GHz	4 GHz	4 GHz
Bandwidth Ratio : $\nu_{max} : \nu_{min}$	2 : 1	2 : 1	1.5 : 1
Fractional Bandwidth : $(\nu_{max} - \nu_{min}) / \nu_{mid}$	66%	66%	40%

UV-coverage / imaging properties change with frequency

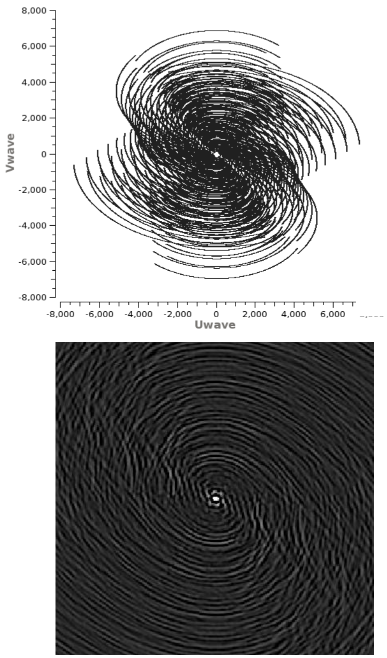
1.0 GHz



1.5 GHz



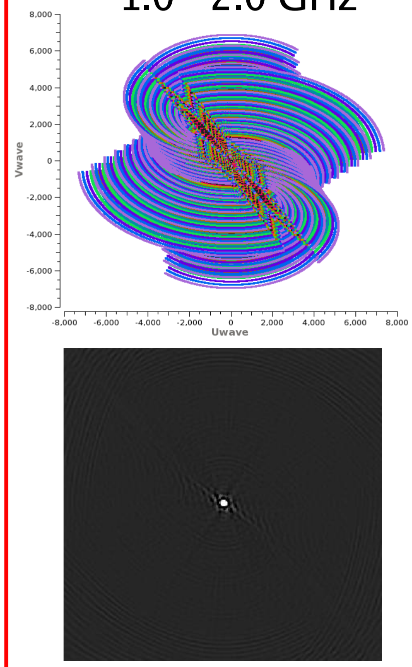
2.0 GHz



$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\,\nu}{c}$$

=> combine multi-frequency measurements during imaging

1.0 - 2.0 GHz



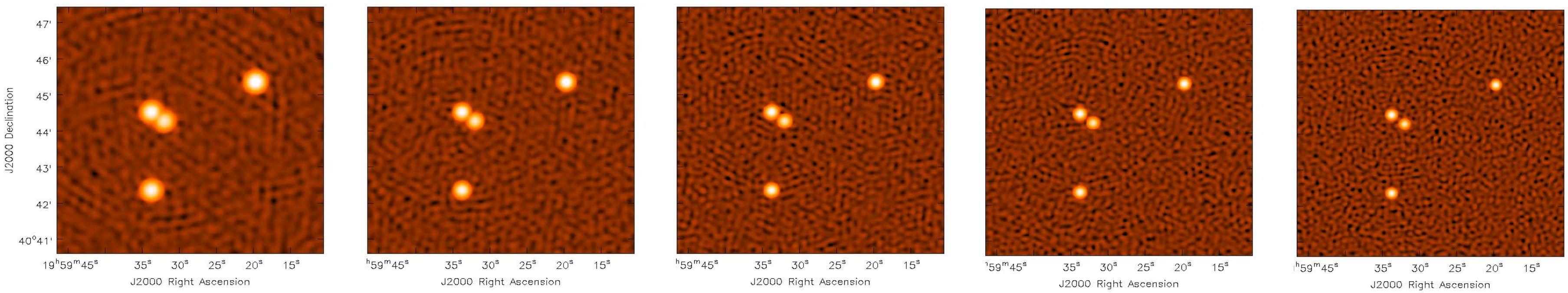
Sky Brightness can also change with frequency → model intensity and spectrum

Spectral Cube (vs) MFS imaging

Simulation : 3 flat-spectrum sources + 1 steep-spectrum source (1-2 GHz VLA observation)

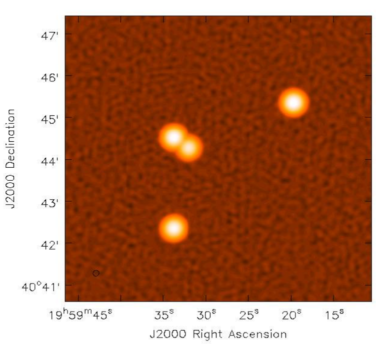
Images made at different frequencies between 1 and 2 GHz (limited to narrow-band sensitivity)

1 GHz

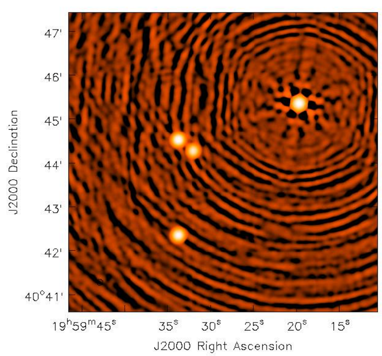


2 GHz

Add all single-frequency images (after smoothing to a low resolution)

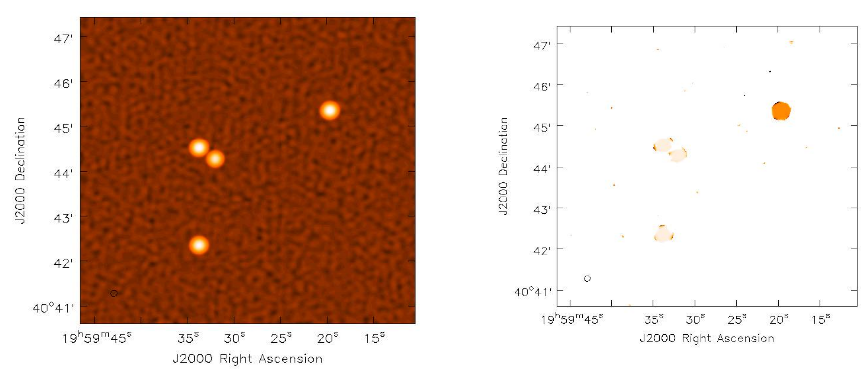


Use wideband UV-coverage, but ignore spectrum
(MFS, nterms=1)



Use wideband UV-coverage
+ Model and fit for spectra too
(MT-MFS, nterms > 1)

Output : Intensity and Spectral-Index



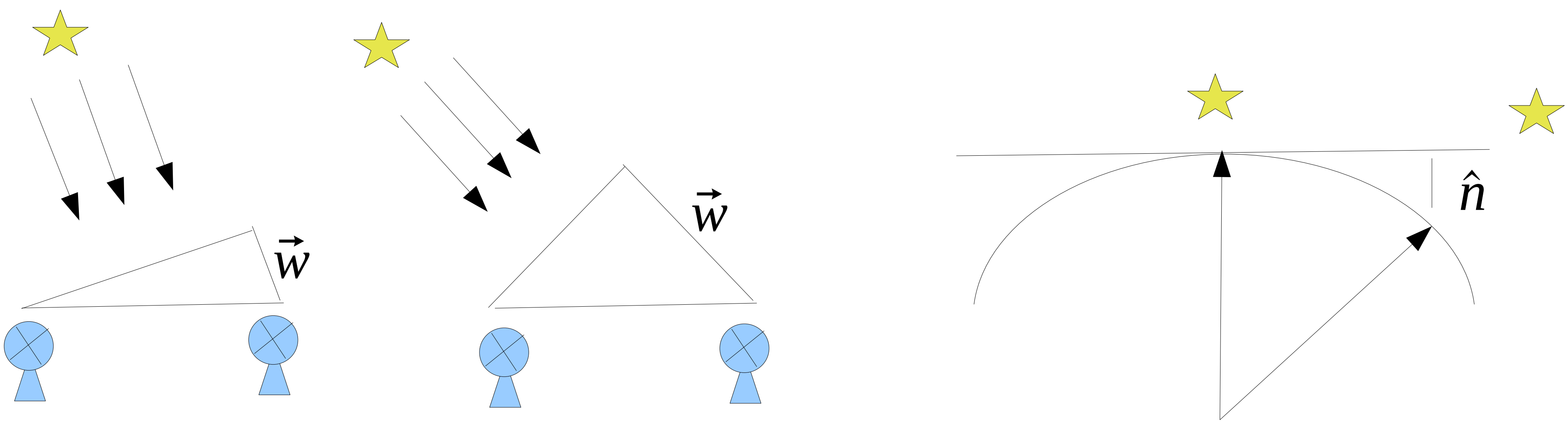
Wide-Field Imaging – W-term

$$V^{obs}(u,v) = S(u,v) \iint I(l,m)\, e^{2\pi i(ul+vm)}\, dl\, dm$$

$$V^{obs}(u,v) = S(u,v) \iiint I(l,m)\, e^{2\pi i(ul+vm+w(n-1))}\, dl\, dm\, dn$$

The ' w ' component of a baseline can be large, away from the image phase center

The ' n ' component of a source can be large, away from the image phase center



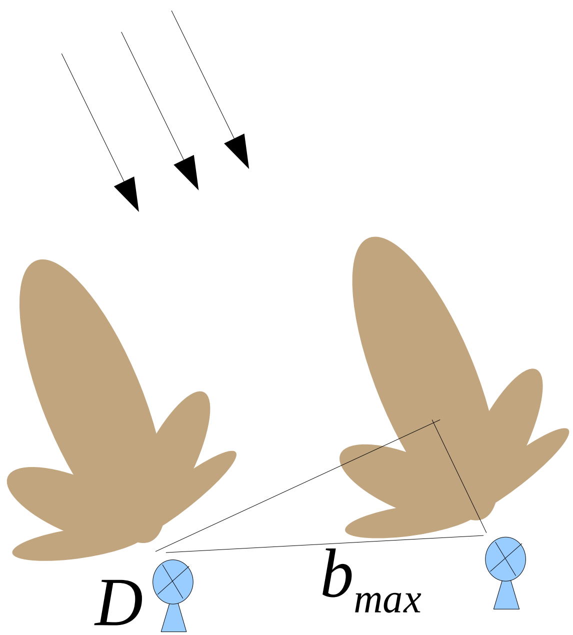
There are algorithms to account for this : Image Faceting, W-Projection.

Wide-Field Imaging – Primary Beams

Each antenna has a limited field of view => Primary Beam (gain) pattern

=> Sky is (approx) multiplied by PB, before being sampled by the interferometer

$$I^{obs}(l,m) \approx I^{PSF}(l,m) * \left[P^{sky}(l,m) \cdot I^{sky}(l,m) \right]$$

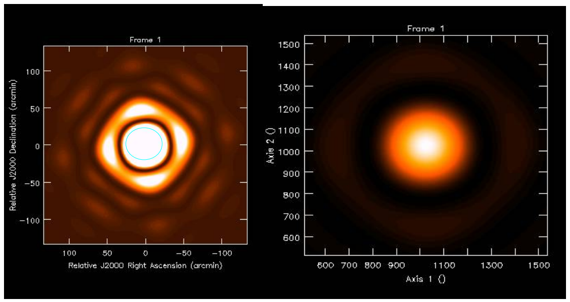


The antenna field of view :
D = antenna diameter

$$\lambda / D$$

Compare with angular resolution of the interferometer :

$$\lambda / b_{max}$$



But, in reality, P changes with time, freq, pol and antenna....

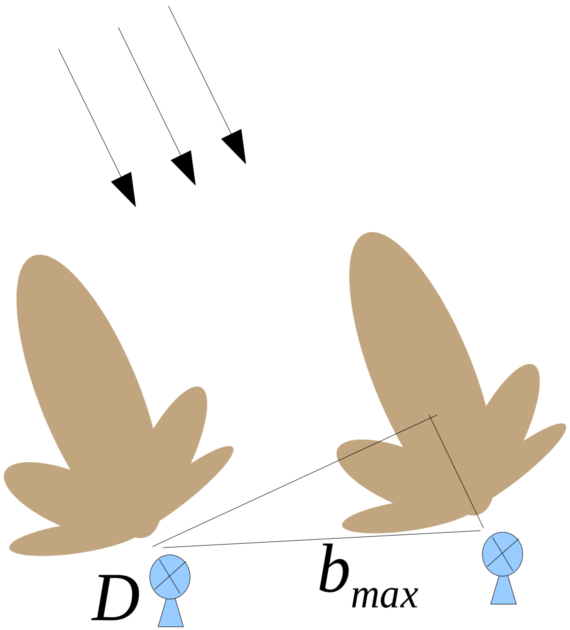
=> Ignoring such effects limits dynamic range to 10^4
=> More-accurate method to account for this : A-Projection

Wide-Field Imaging – Primary Beams

Each antenna has a limited field of view => Primary Beam (gain) pattern

=> Sky is (approx) multiplied by PB, before being sampled by the interferometer

$$I^{obs}(l,m) \approx I^{PSF}(l,m) * \left[P^{sky}(l,m) \cdot I^{sky}(l,m) \right]$$

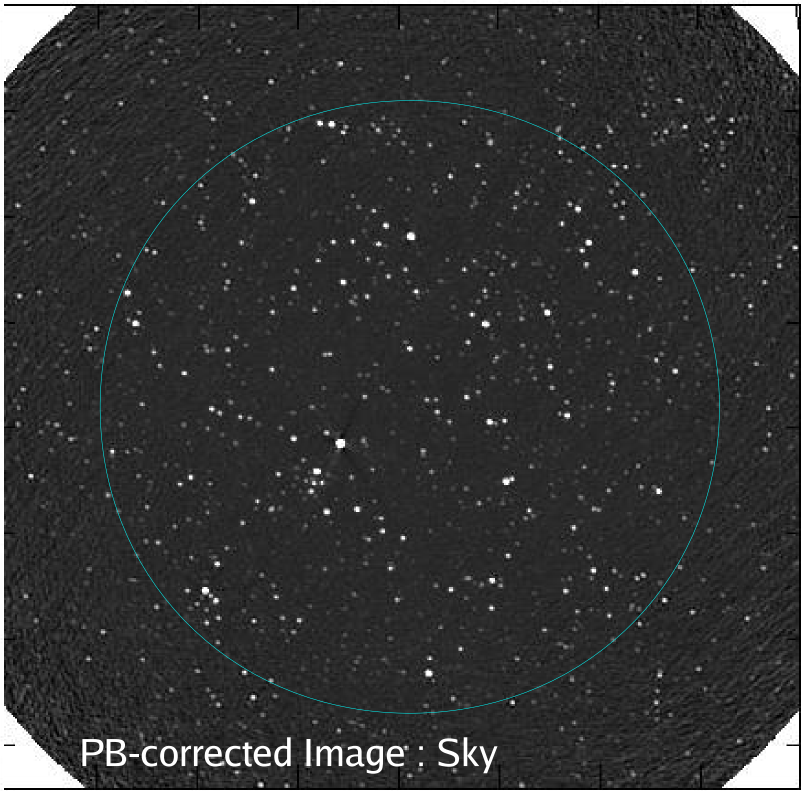


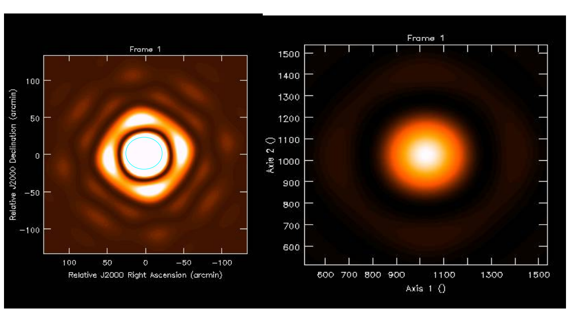
The antenna field of view :
D = antenna diameter

$$\lambda / D$$

Compare with angular resolution of the interferometer :

$$\lambda / b_{max}$$



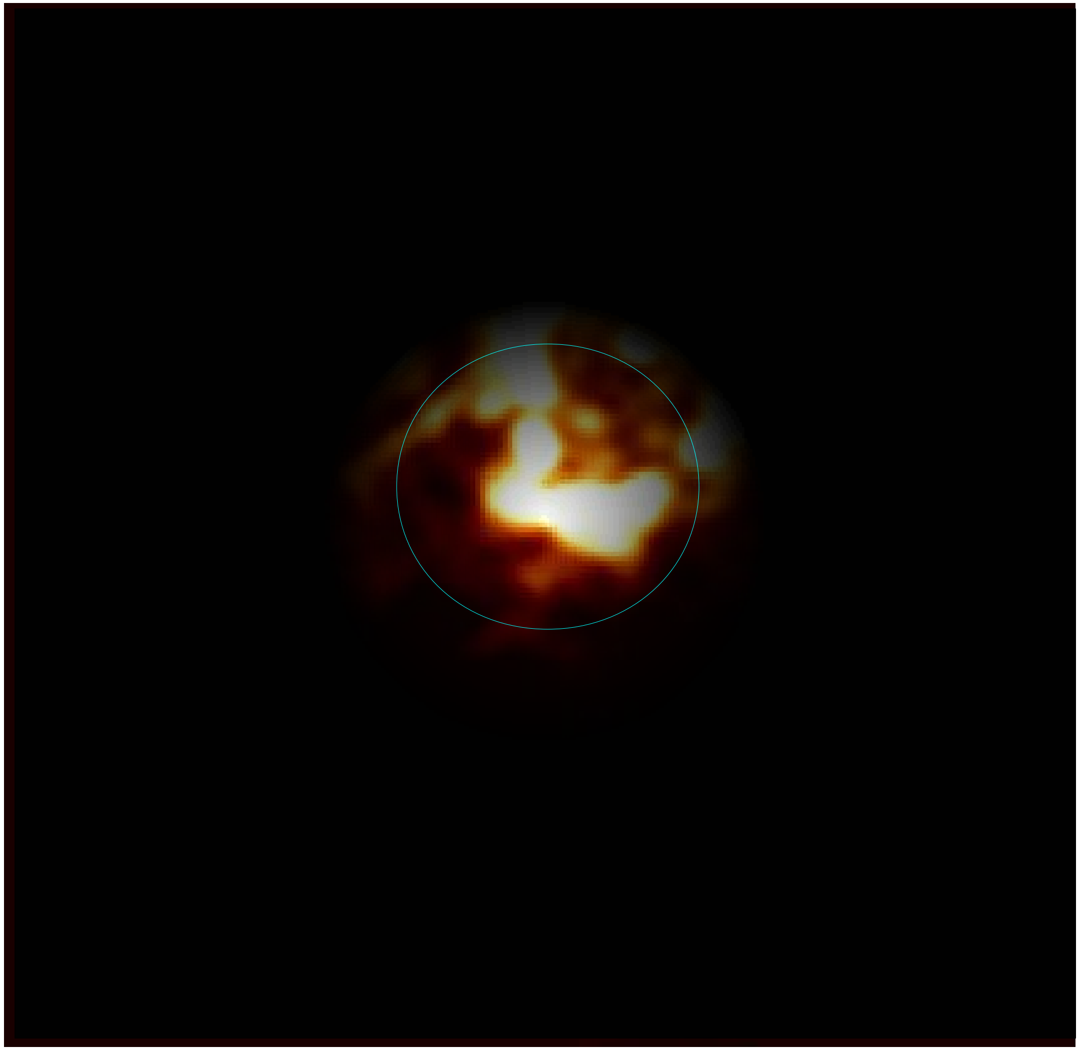


But, in reality, P changes with time, freq, pol and antenna....

=> Ignoring such effects limits dynamic range to 10^4
=> More-accurate method to account for this : A-Projection

Wide-field Imaging -- Mosaics

Combine data from multiple pointings to form one large image.



One Pointing sees only part of the source

Combine pointings either before or after deconvolution.

Stitched mosaic :

-- Deconvolve each pointing separately
-- Divide each image by PB
-- Combine as a weighted avg

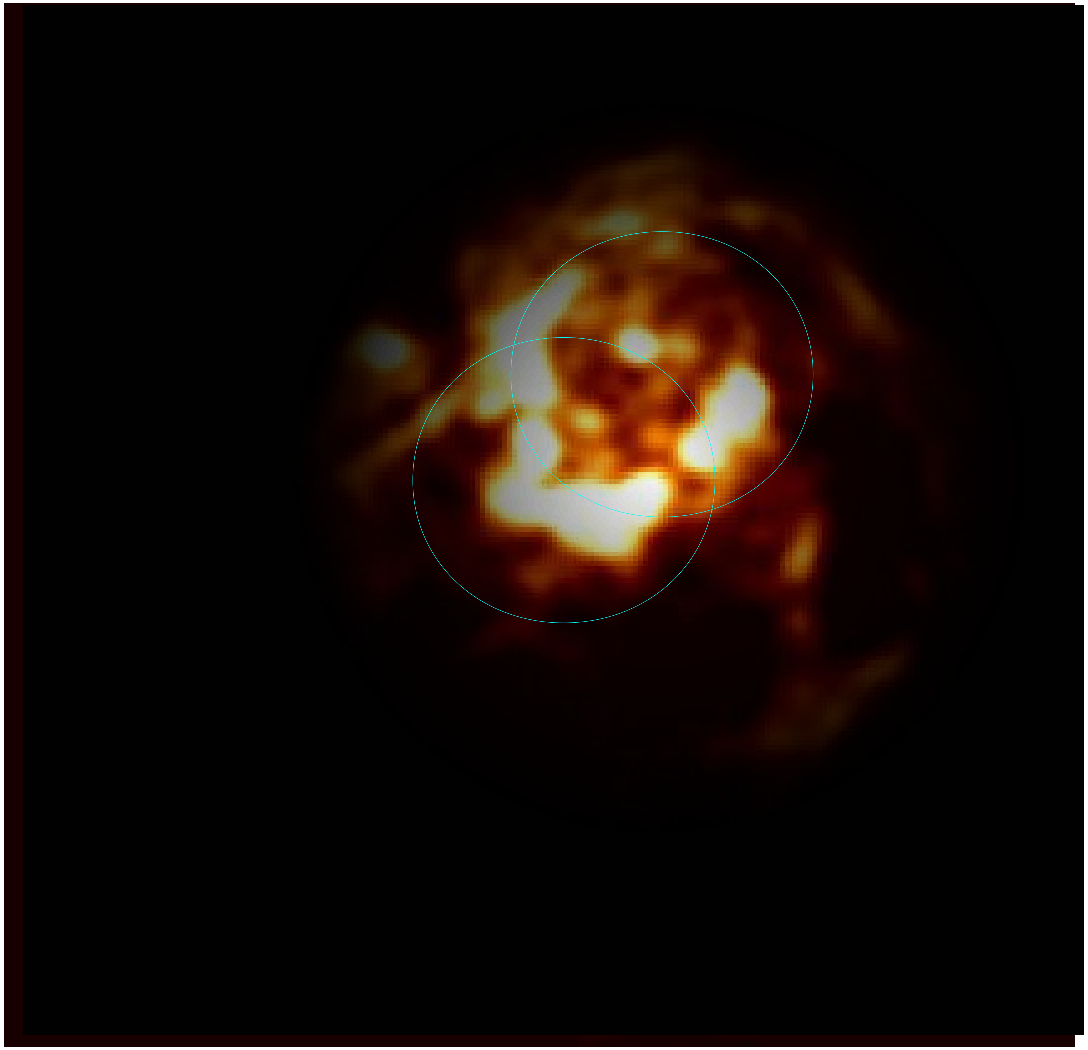
Joint mosaic :

-- Combine observed images as a weighted average
(or)
Grid all data onto one UV-grid, and then iFFT

-- Deconvolve as one large image

Wide-field Imaging -- Mosaics

Combine data from multiple pointings to form one large image.



Two Pointings see more.....

Combine pointings either before or after deconvolution.

Stitched mosaic :

-- Deconvolve each pointing separately
-- Divide each image by PB
-- Combine as a weighted avg

Joint mosaic :

-- Combine observed images as a weighted average

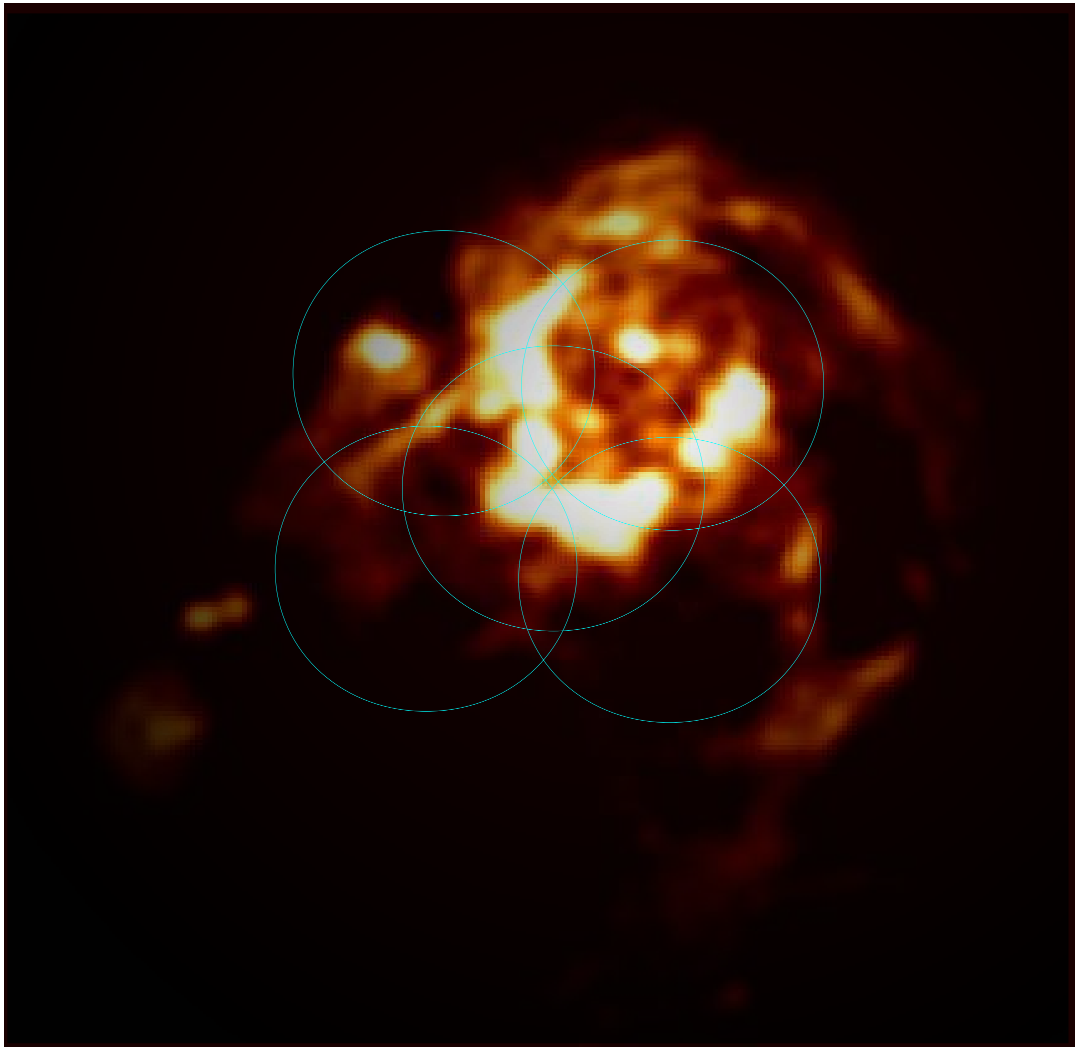
(or)

Grid all data onto one UV-grid, and then iFFT

-- Deconvolve as one large image

Wide-field Imaging -- Mosaics

Combine data from multiple pointings to form one large image.



Use many pointings to cover the source with approximately uniform sensitivity

Combine pointings either before or after deconvolution.

Stitched mosaic :

-- Deconvolve each pointing separately
-- Divide each image by PB
-- Combine as a weighted avg

Joint mosaic :

-- Combine observed images as a weighted average
(or)
Grid all data onto one UV-grid, and then iFFT

-- Deconvolve as one large image

Some points to remember …

How does an interferometer form an image ?

- Each antenna pair measures one 2D fringe. Many antenna pairs => Fourier series

How do you make a raw image from interferometer data ?

- Assign weights to visibilities, grid them, take a Fourier transform

How do you choose the cell-size and image size for imaging ?

- Cell size = (angular resolution / 3).   Image size = field-of-view / cell size

What does the raw observed image represent ?

- Observed Sky is the convolution of the true sky and the PSF

How do you get a model of the sky ?

- Solve the convolution equation via algorithms like Clean, MS-Clean, MT-Clean...

Some points to remember …

How do you measure image quality ?

- RMS noise, Peak residual, Dynamic range, Image fidelity

How does wide-band data affect the imaging process ?

- Increased sensitivity, but the imaging properties and sky change with frequency

How do you image wide-band data ?

- Make a Cube of images, or Multi-Frequency-Synthesis with a spectral fit.

What is an antenna primary beam and what is its effect on an image ?

- Antenna power pattern. It multiplies with the sky, before convolution with the PSF

What is the w-term problem ?

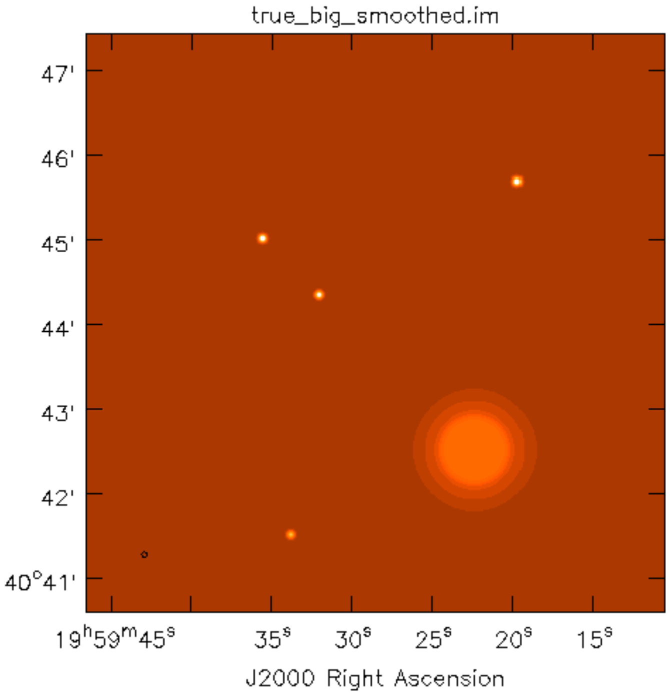
- 2D Fourier transform approximations are invalid far away from the image center

Example Imaging Problem – Simulated data

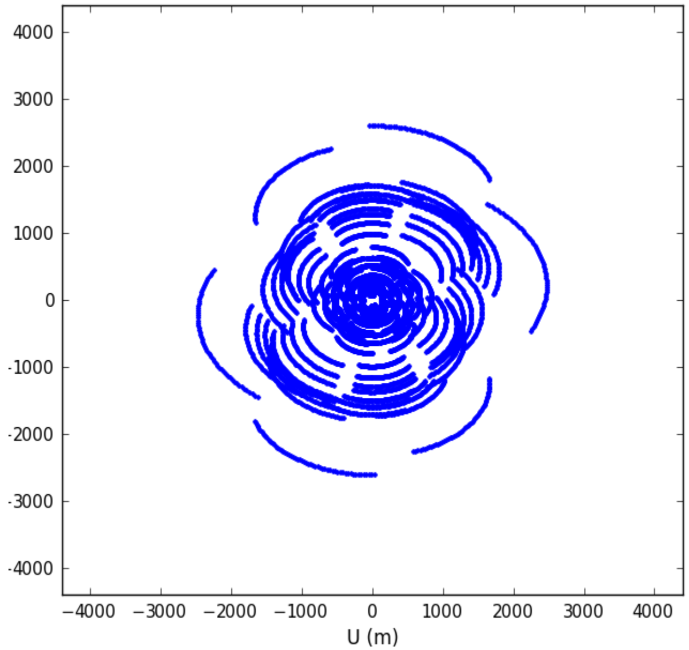
Simulated 5 GHz observation with a 13-antenna array over 5 hours

N visibilities : 9360. Visibility noise : 2 Jy   =>   Theoretical image RMS : 0.02 Jy

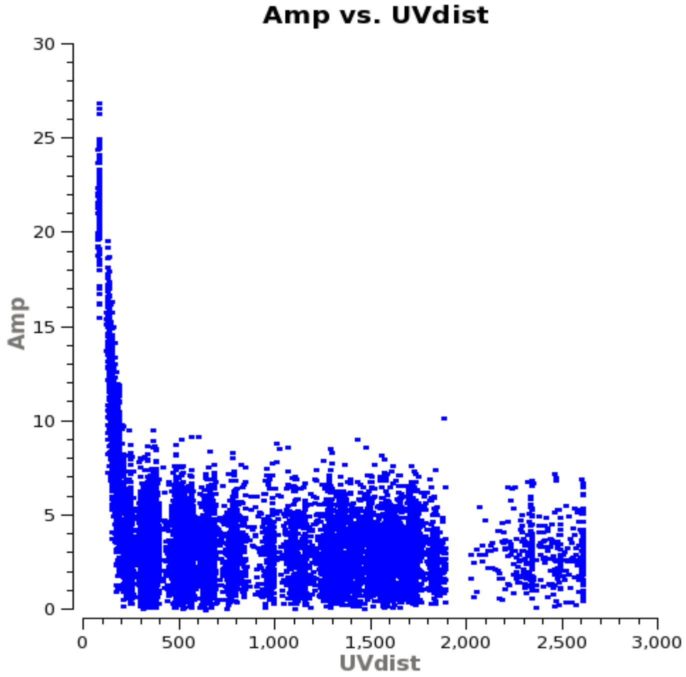
Angular resolution : 5 arcsec   (Max baseline of 2500m at 5.0 GHz)
Sky brightness has compact and extended structure (partially-sampled).
Peak brightness : 1 Jy => Target dynamic range = 50



$I^{sky}(l,m)$



$S(u,v)$



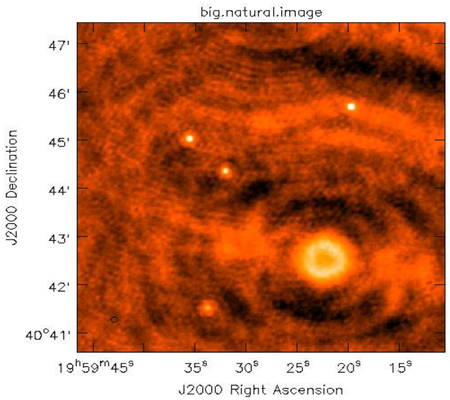
$V^{sky}(u,v) \cdot S(u,v)$

Example Imaging Problem – First try....

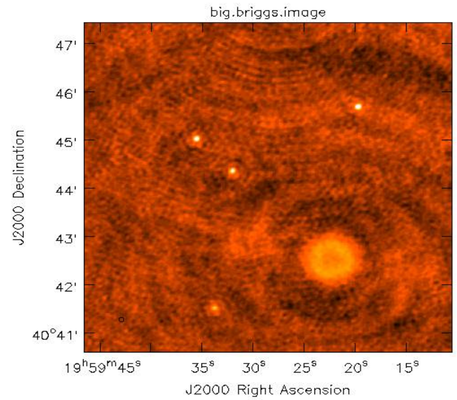
Quick deconvolution with different weighting schemes :

Image FOV : 7 arcmin (512 pixels at 0.8 arcsec pixel size)

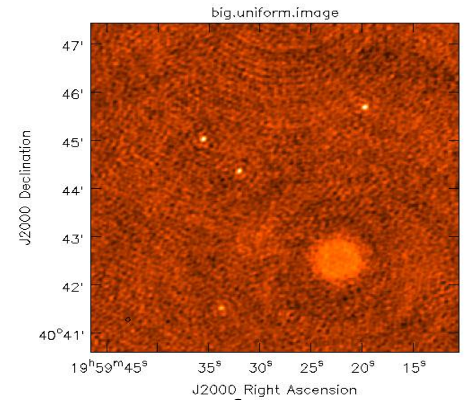
MS-CLEAN : NIter=100, scales=[0,6,40], gain=0.3, robust=0.7



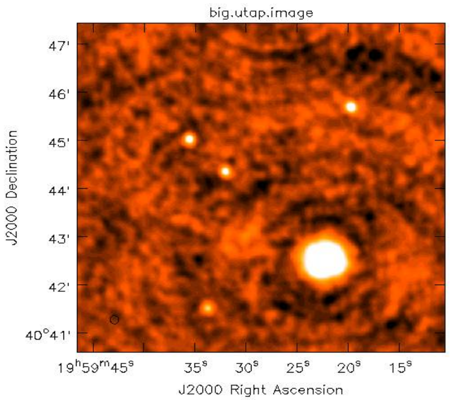
Natural
High sidelobes



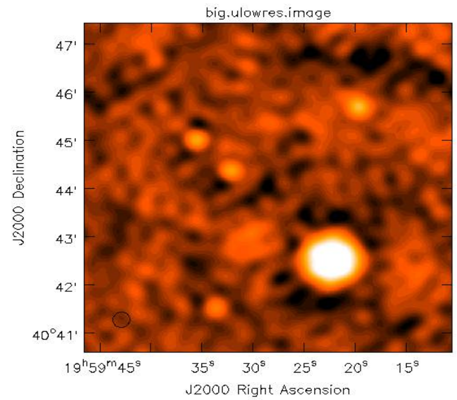
Robust = 0.7



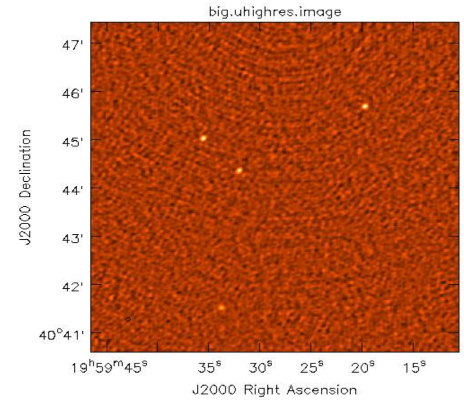
Uniform
Low sensitivity to extended emission



Uniform with a uv-taper for 9 arcsec



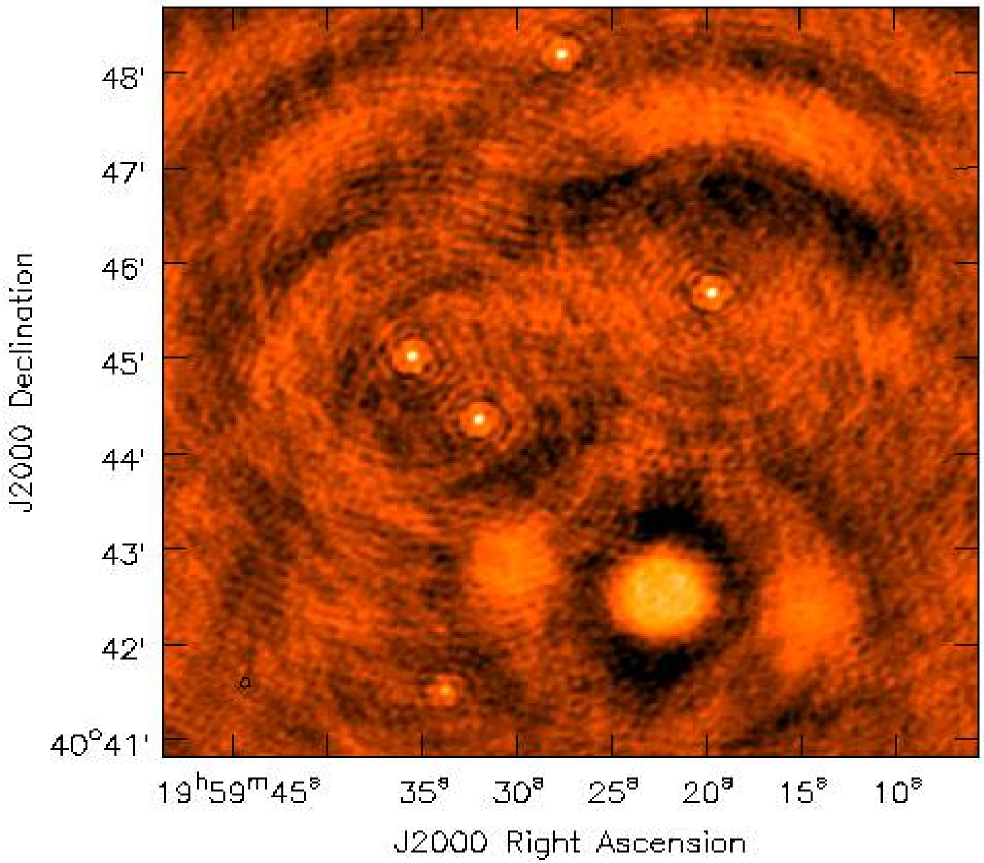
Uniform with only SHORT Baselines < 500m



Uniform with only LONG Baselines > 500m
(Extended structure disappears)

Example Imaging Problem – Second try...

Make a larger image (700 pixels at 0.8 arcsec cell size)



N Iter = 0 (dirty image)

Pick scales = [0,6,16,30,42,60]
Weighting : Robust=0.7

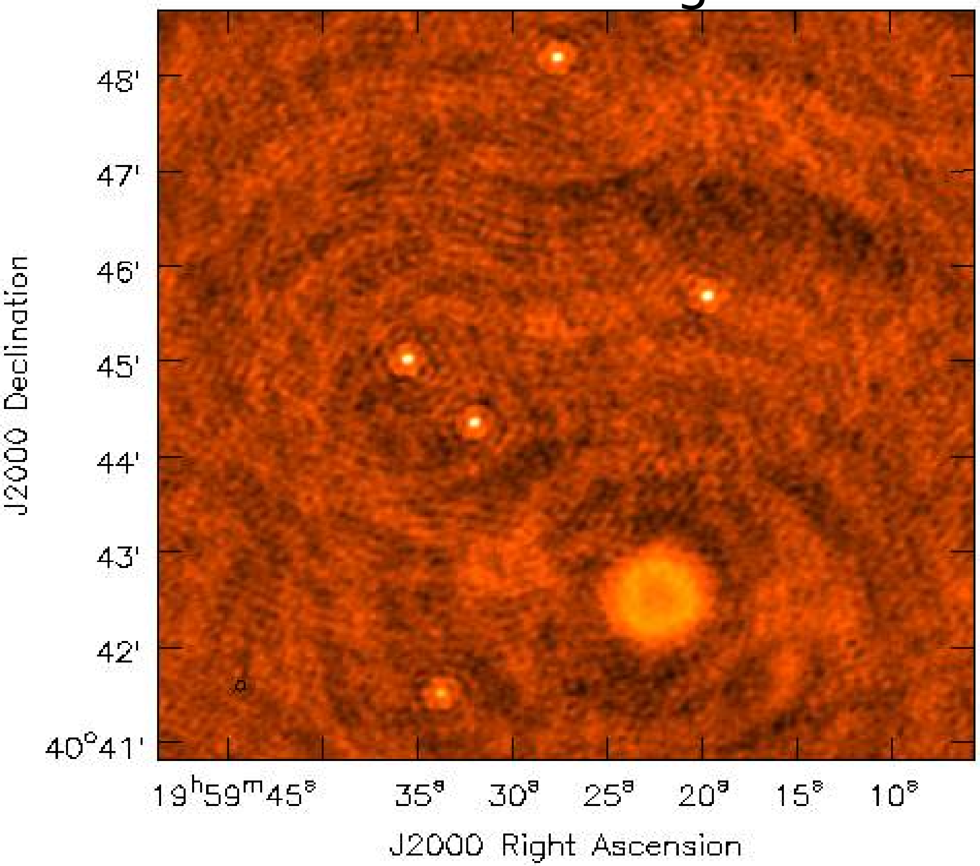
Loop gain = 0.2
(go slow, because of insufficient data-constraints for the extended emission)

Peak sidelobe structure : 0.2 Jy/beam. Off-source RMS : 0.1 Jy/beam
Peak brightness : 1 Jy/beam => Dynamic Range : 10 ~ 20

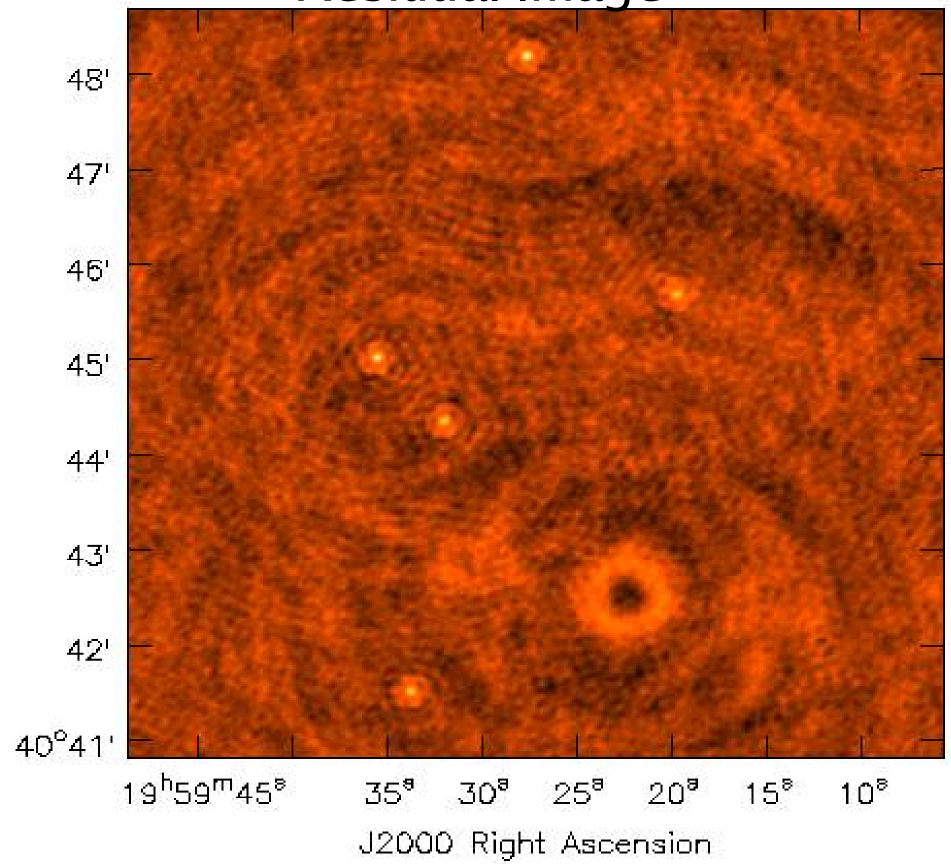
Example Imaging Problem – Second try...

After 100 iterations.

Restored Image



Residual Image

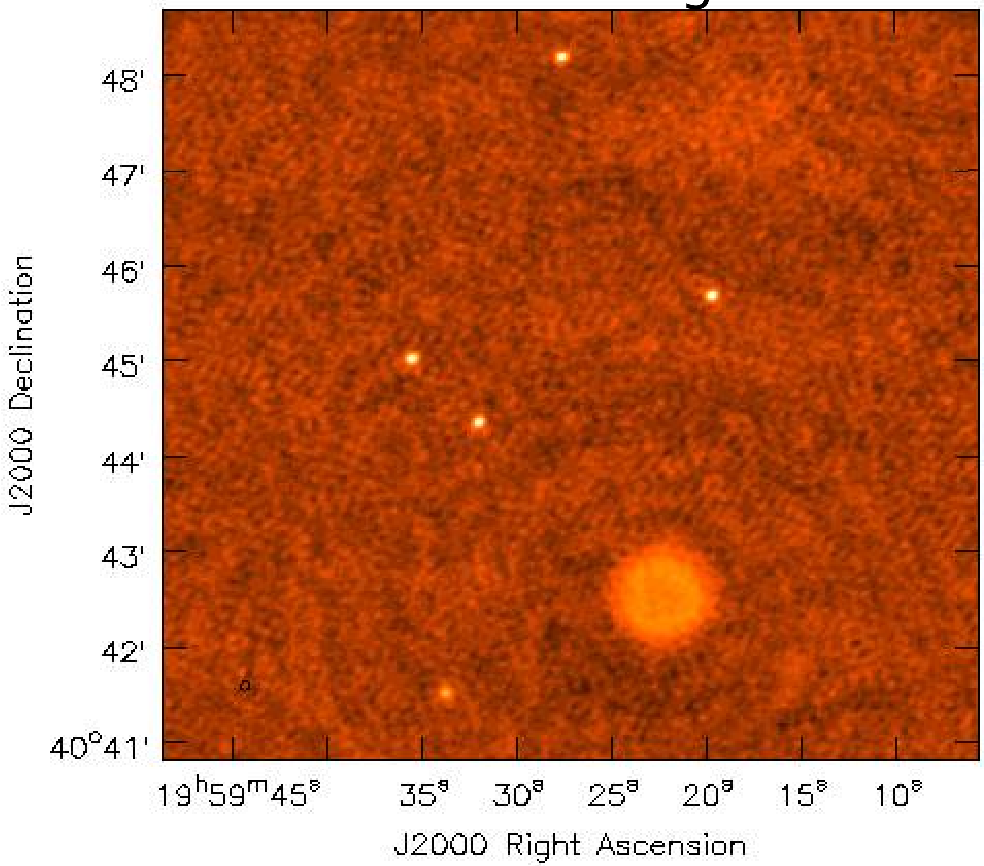


Peak sidelobe structure : 0.1 Jy/beam. Off-source RMS : 0.05 Jy/beam
Peak brightness : 1 Jy/beam => Dynamic Range : 10 ~ 20

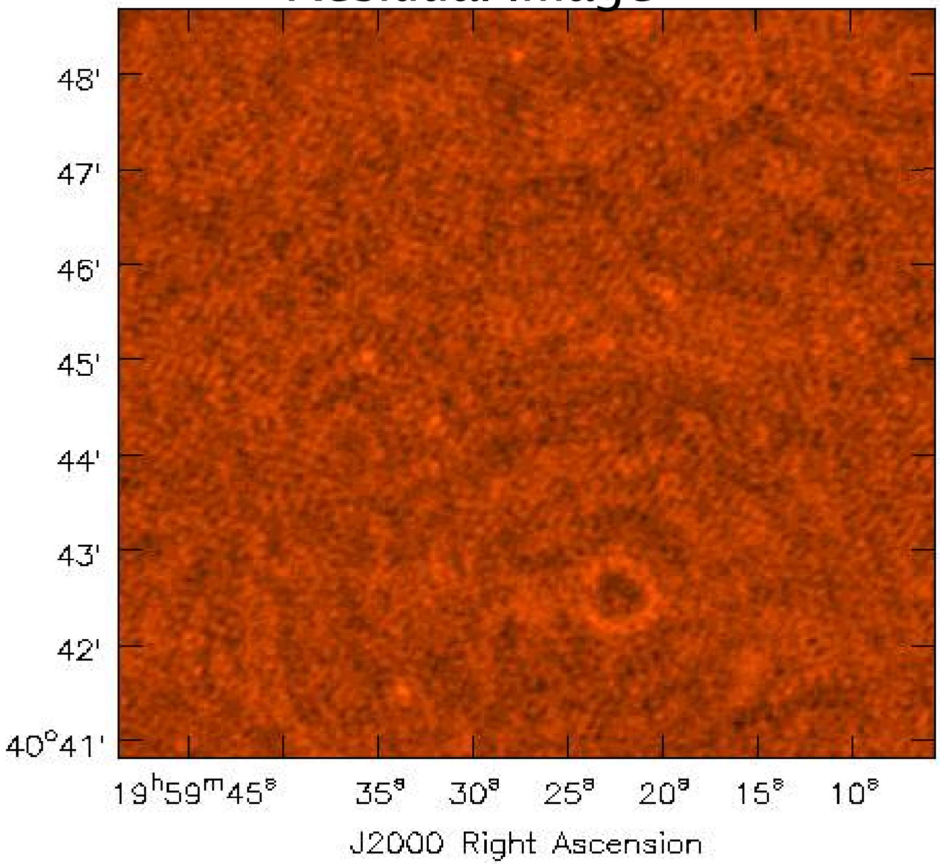
Example Imaging Problem – Second try...

After 500 iterations. Almost OK. Spurious extended flux in the upper-left. No counterpart in the residual image => large scales unconstrained by the data

Restored Image



Residual Image

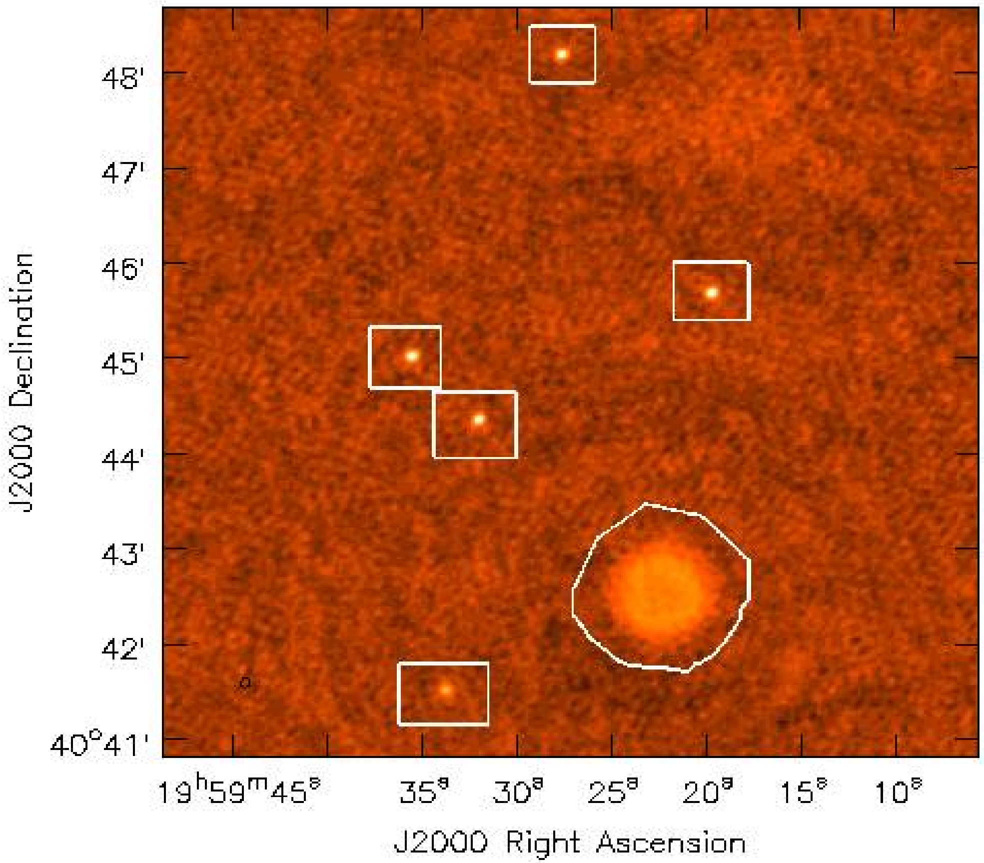


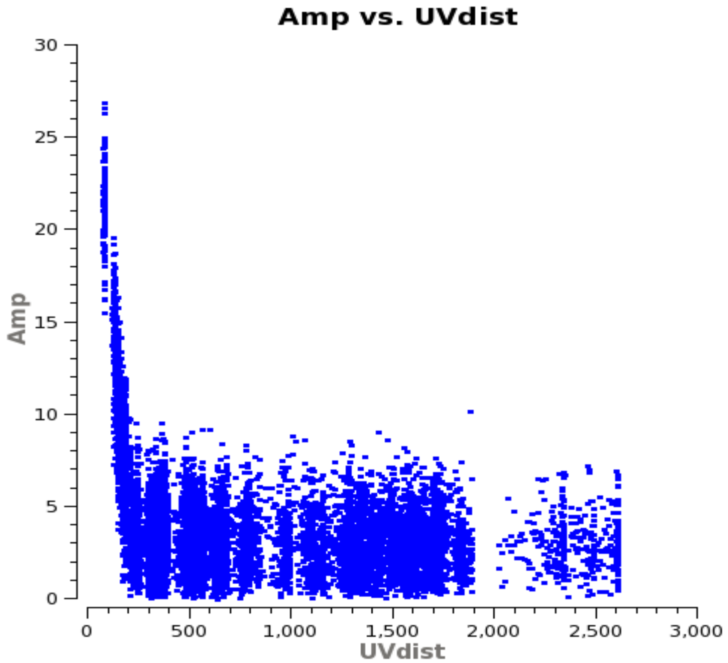
Peak artifacts : 0.07 Jy/beam.   Off-source RMS : 0.02 Jy/beam
Peak brightness : 1 Jy/beam   =>   Dynamic Range : 14 ~ 50

– Reached theoretical off-source RMS of 0.02 Jy/beam. But peak residual is still high.

Example Imaging Problem – Using masks

Build 'CLEAN boxes' or masks and restart. This will force extended emission to be centered within the allowed regions only.



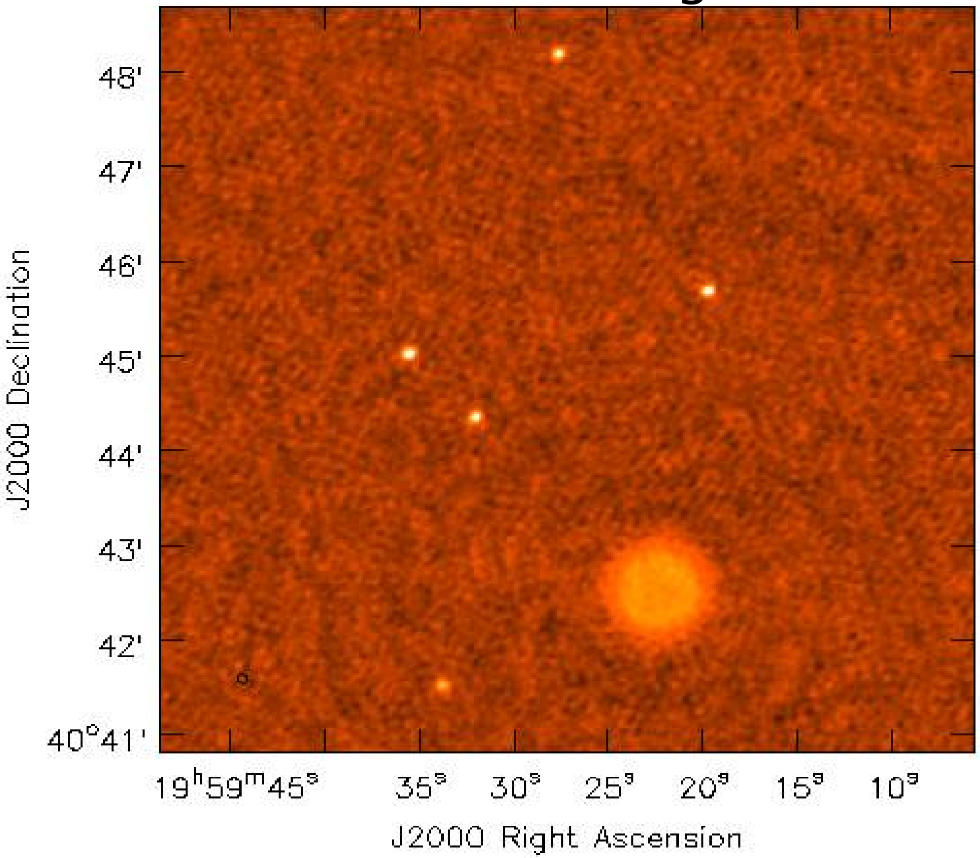


In general, point sources do not require boxes.
Extended emission needs it only if data constraints are insufficient.

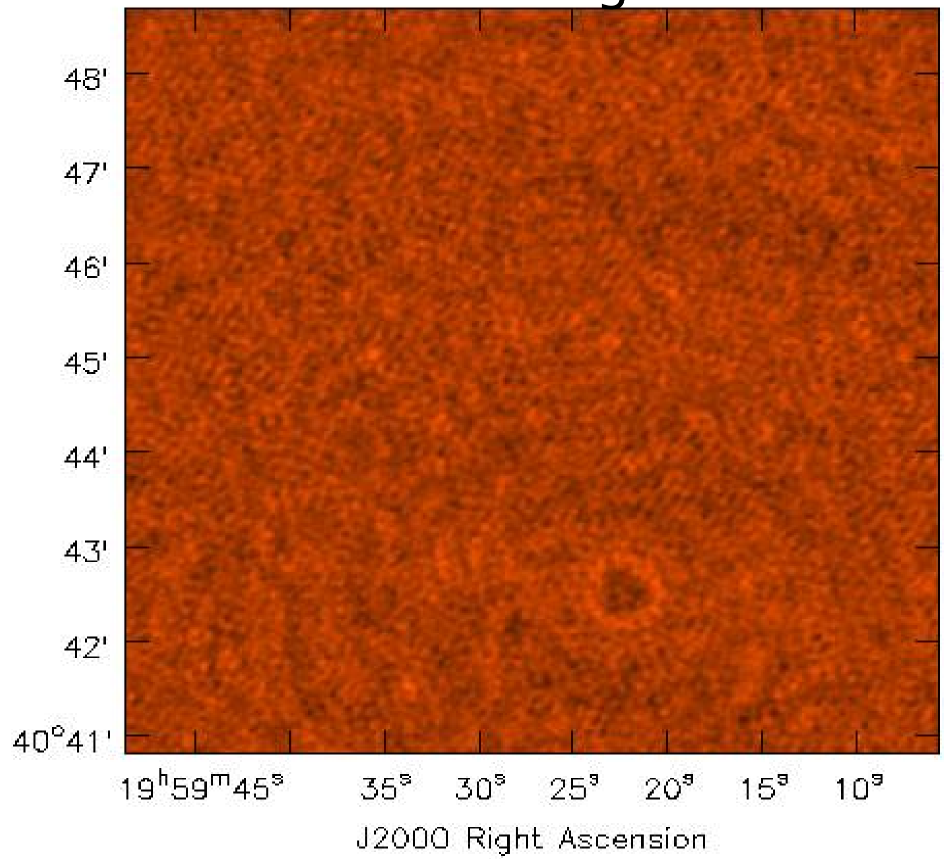
Example Imaging Problem – Third try...

After 300 iterations (compared to 500 earlier) –Reached theoretical rms and dynamic-range !
(in practice, this is not so easy....)

Restored Image



Residual Image



Peak sidelobe structure : 0.04 Jy/beam. Off-source RMS : 0.02 Jy/beam

Peak brightness : 1 Jy/beam => Dynamic Range : 25 ~ 50