Imaging and Deconvolution

CASS Radio Astronomy School
24-28 Sept 2012
Narrabri, NSW, Australia

Urvashi Rau

National Radio Astronomy Observatory
Socorro, NM, USA

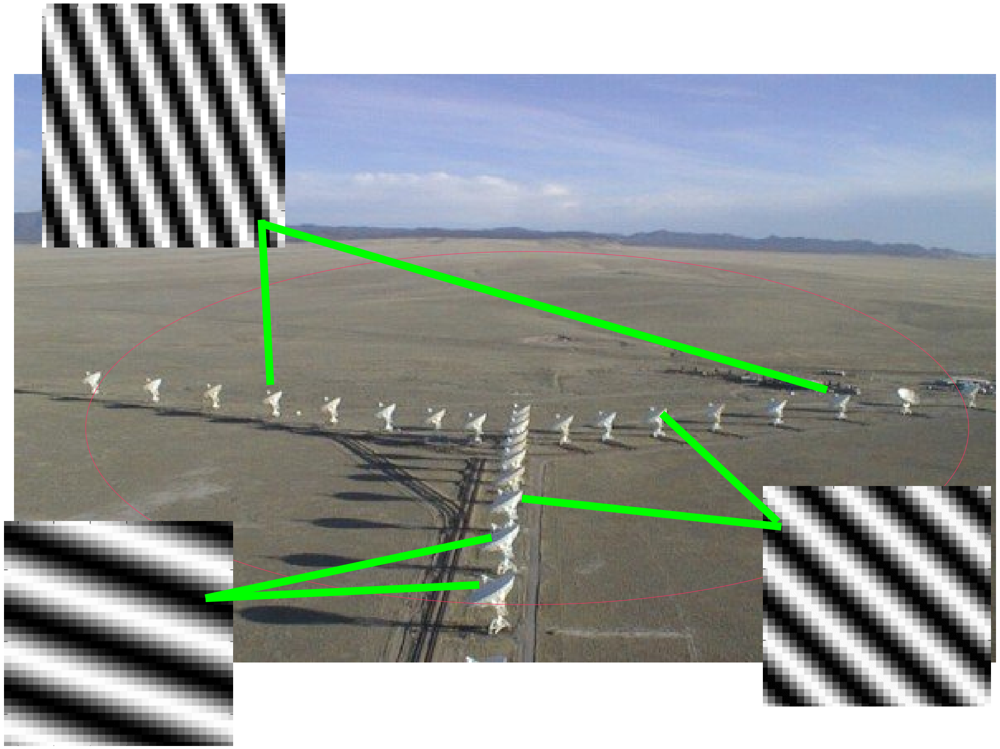
Outline :

- Synthesis Imaging Concepts
- Imaging in Practice
- Image-Reconstruction Algorithms
- An example

The van-Cittert Zernike theorem

" The degree of spatial coherence of the radiation field from a distant spatially incoherent source is proportional to the complex visibility function (spatial Fourier transform) of the intensity distribution across the source. "

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



2D Fourier transform :

Image <--> sum of cosine 'fringes'.

Each antenna-pair measures one 'fringe'.

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

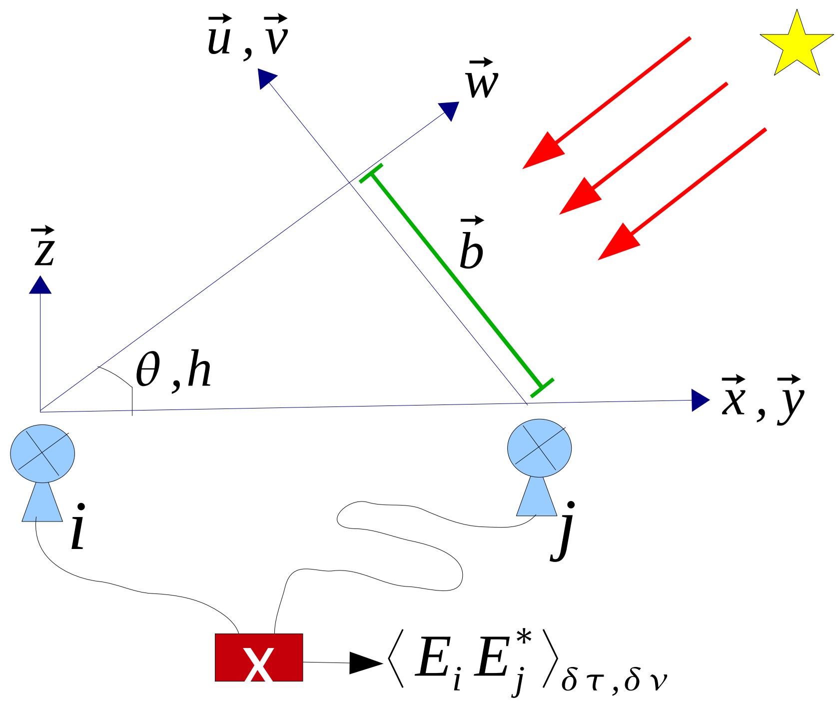
Orientation, Wavelength : Geometry

'spatial frequency'
= coordinates in the spatial Fourier domain.

Think about diffraction patterns through pairs of 'slits'....

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 $\delta u, \delta v$

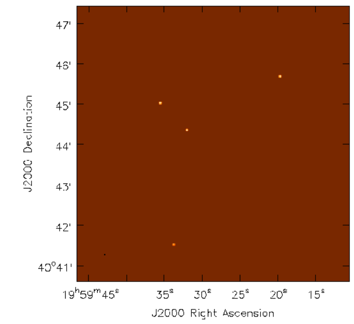
N antennas
N(N-1)/2 antenna-pairs (baselines)

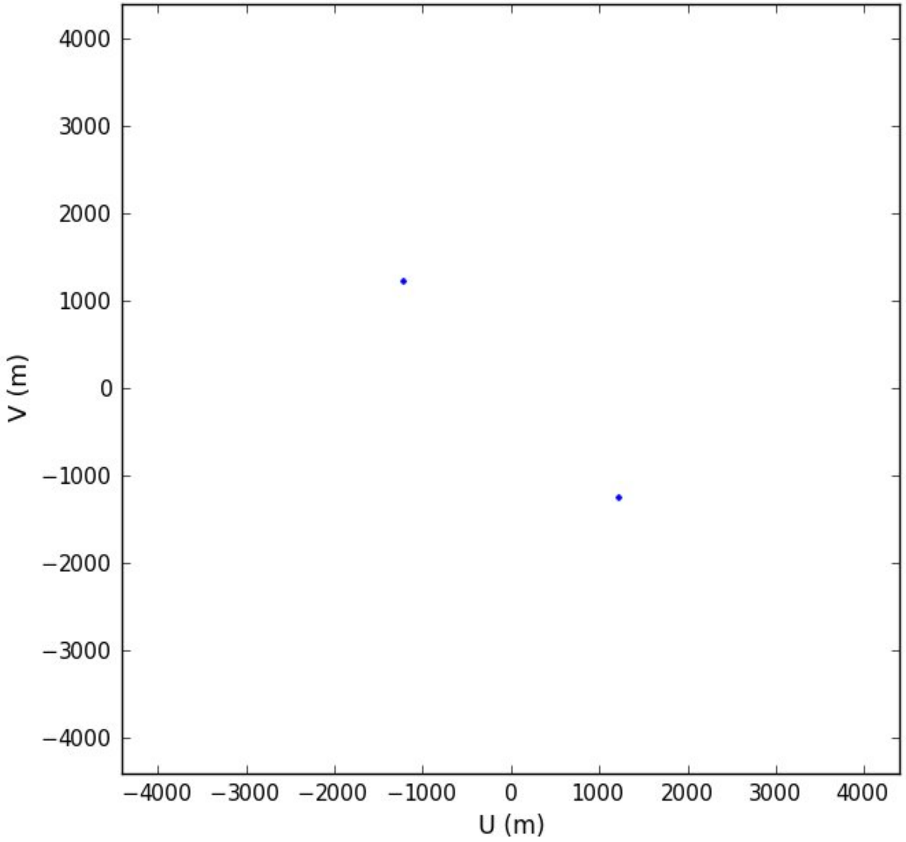
Image is real => Visibility function is Hermitian : $V(u,v) = V^*(-u,-v)$

Spatial Frequency (uv) coverage + Observed Image

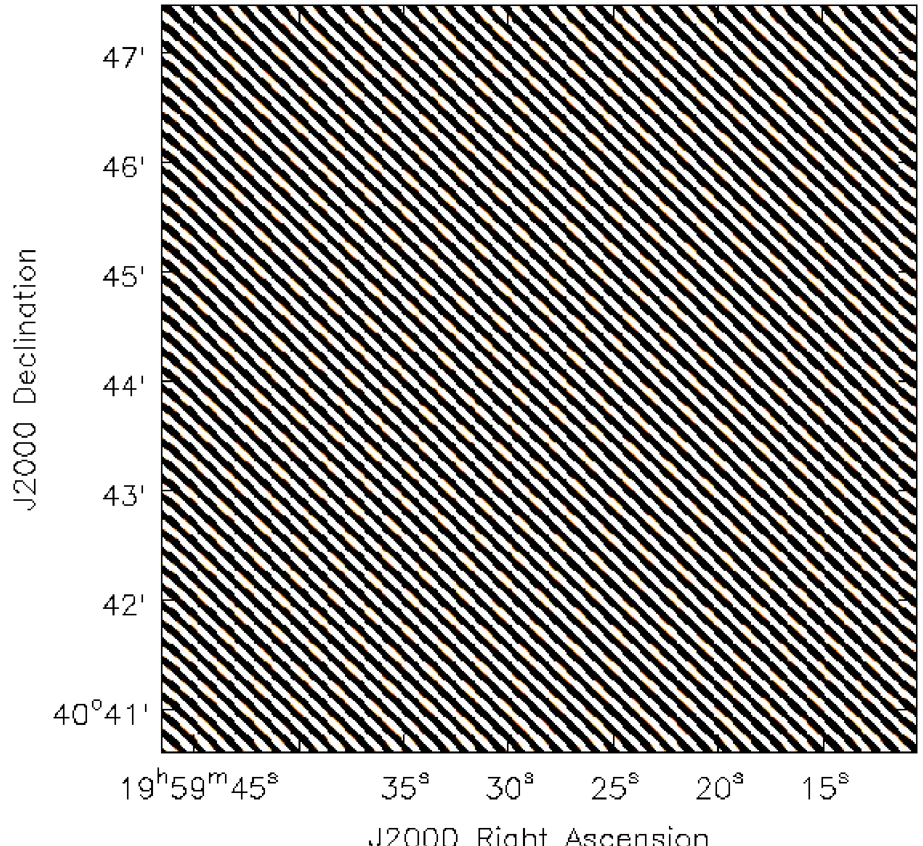
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 2 antennas





$S(u,v)$

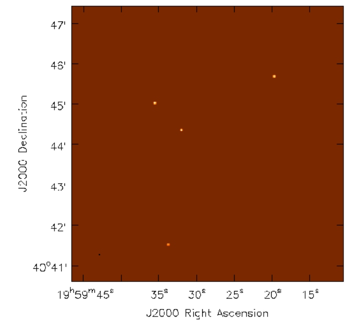


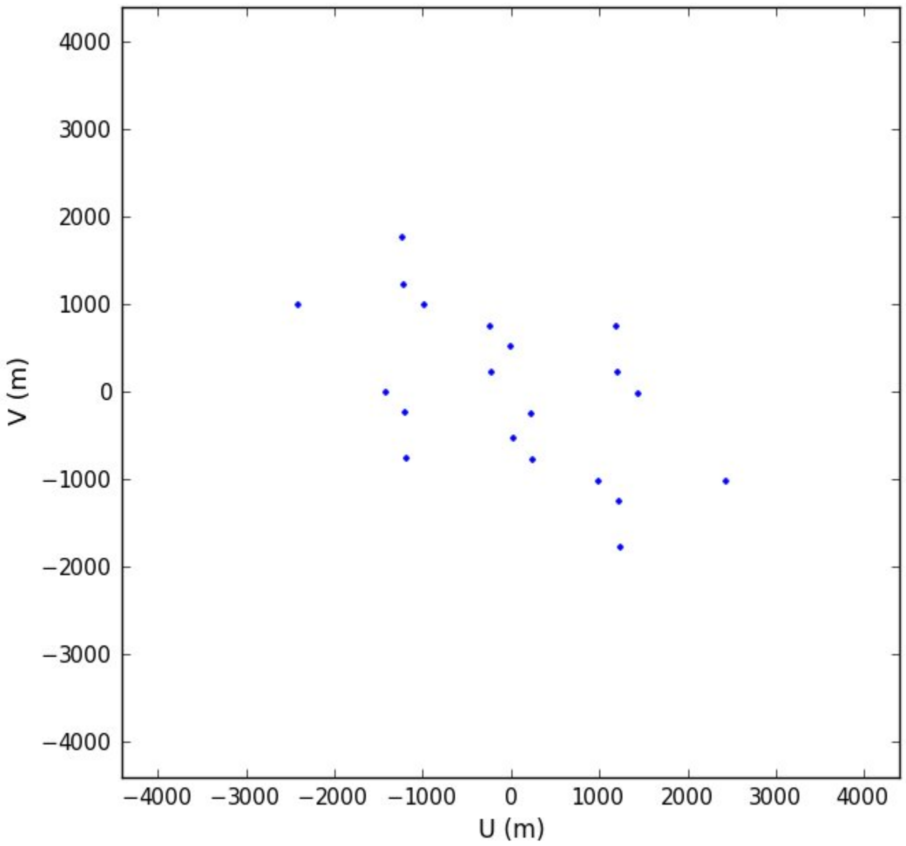
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

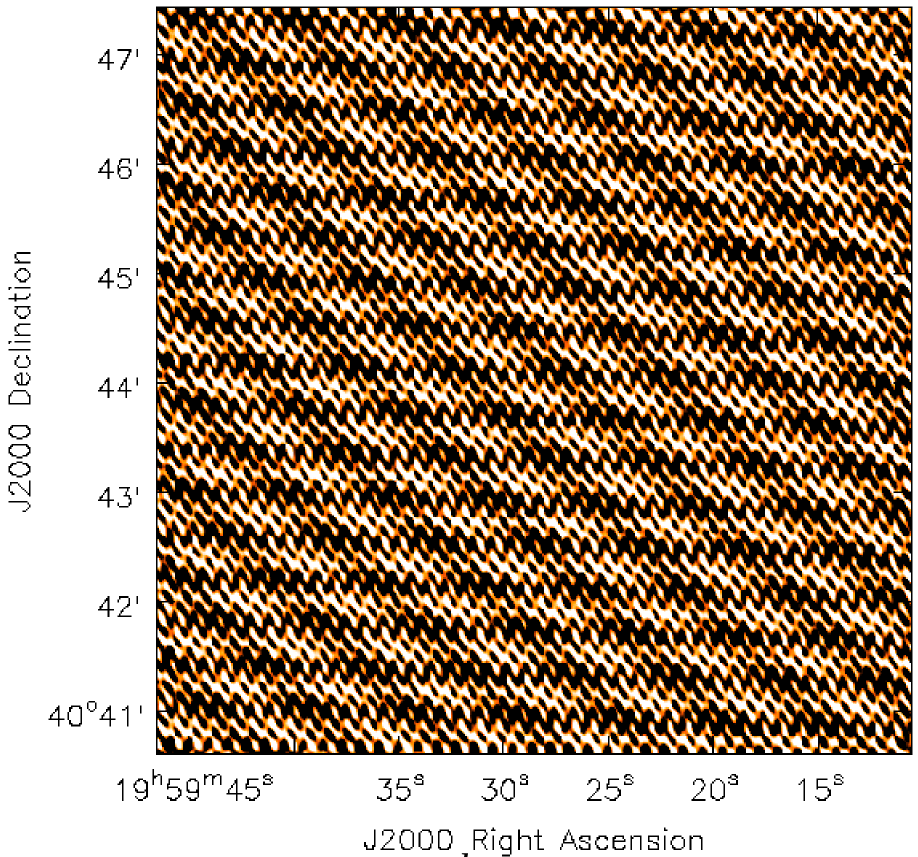
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \left[R(h,\theta) \right] \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 5 antennas





$S(u,v)$

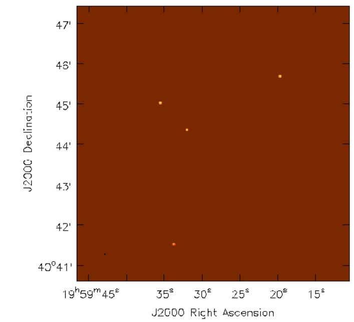


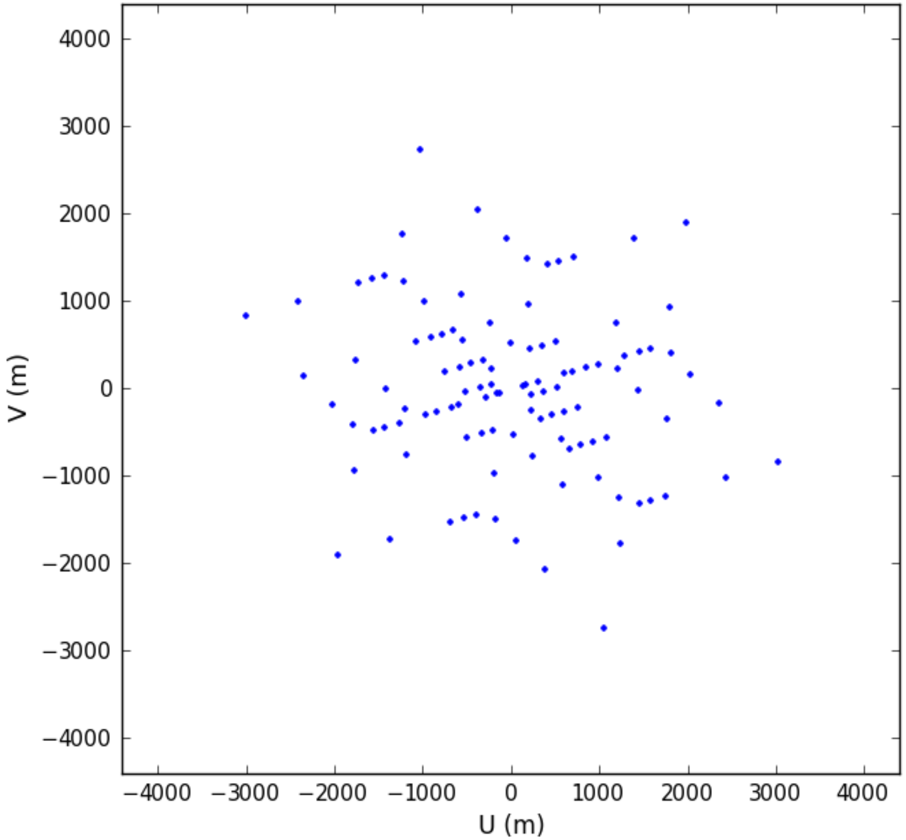
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

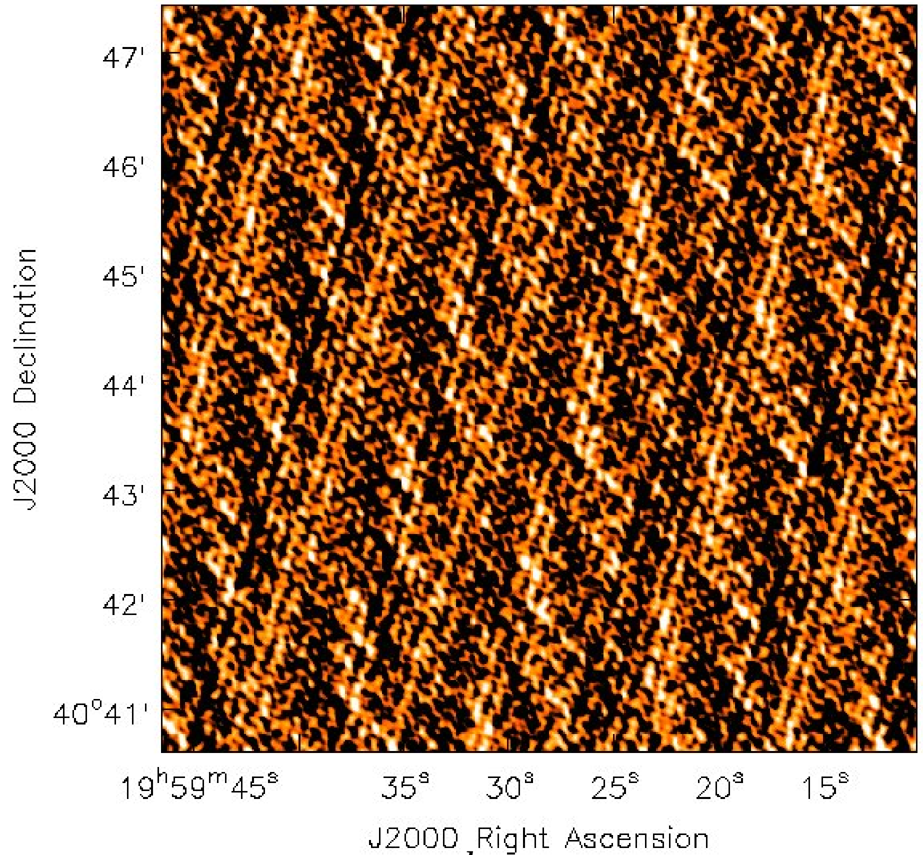
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 11 antennas





$S(u,v)$

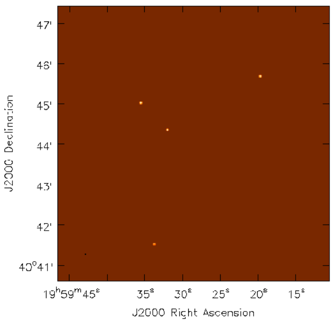


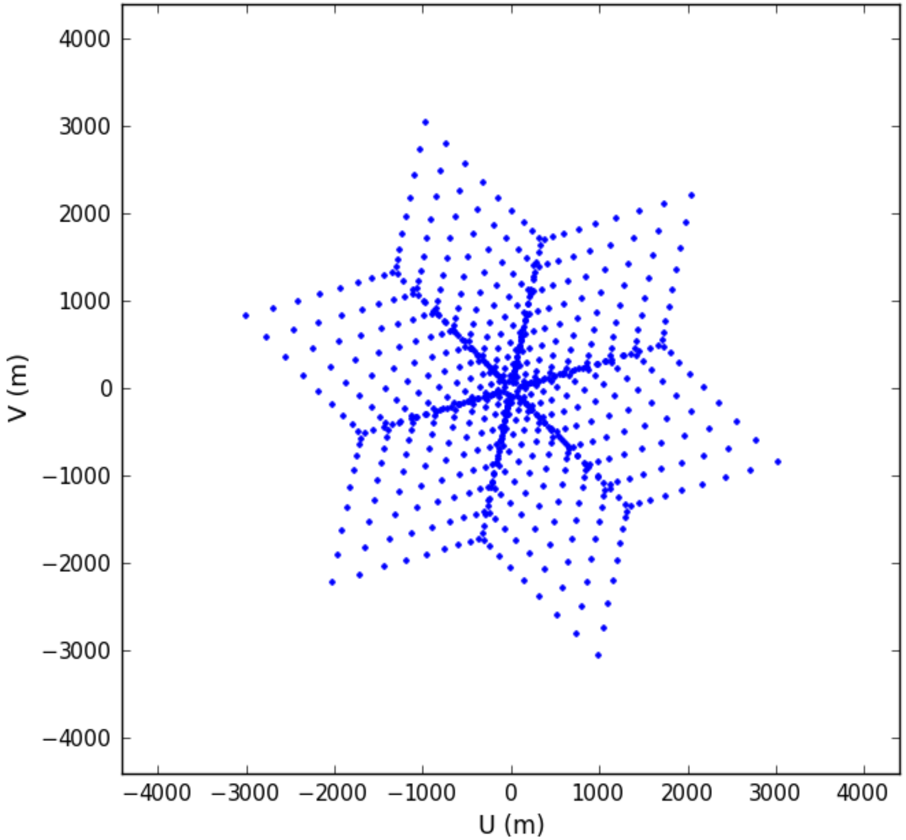
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

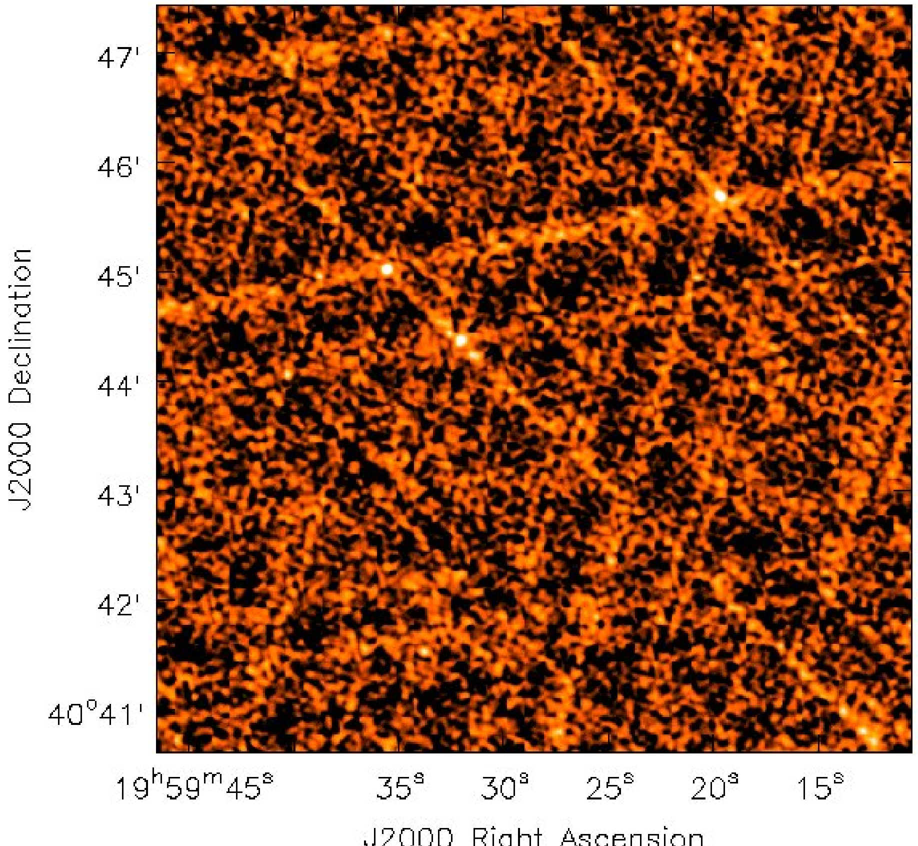
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \left[R(h,\theta) \right] \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

Image of the sky using 27 antennas





$S(u,v)$

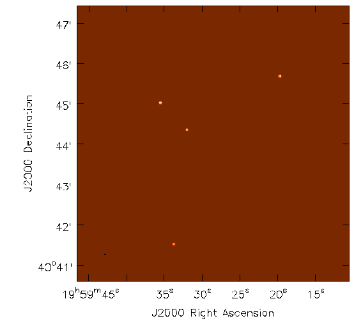


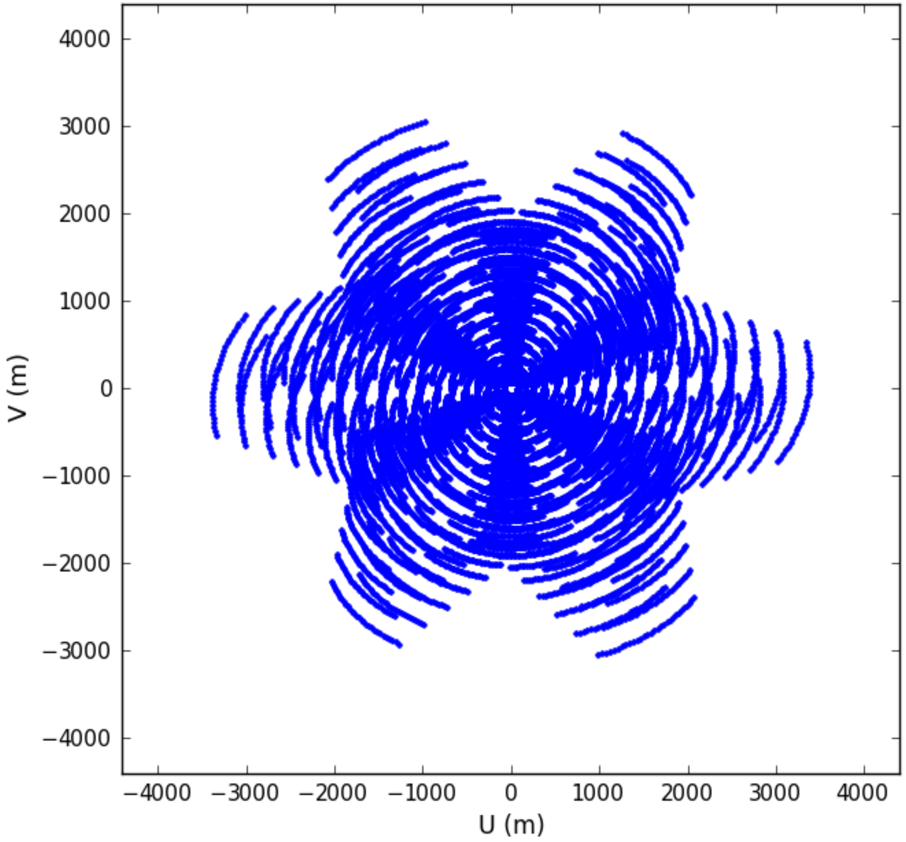
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

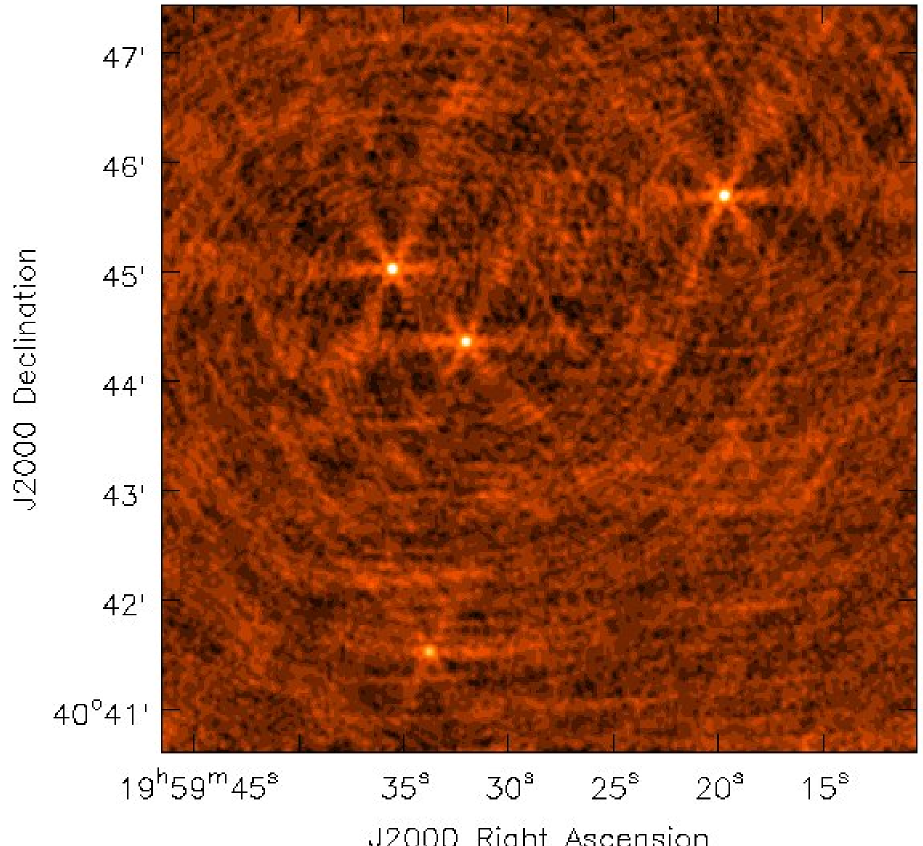
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \left[R(h,\theta) \right] \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

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





$$S(u,v)$$

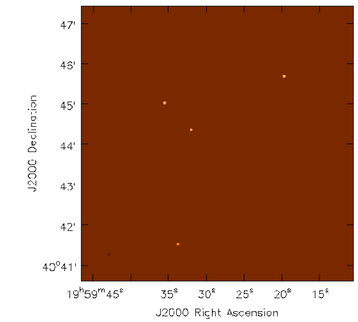


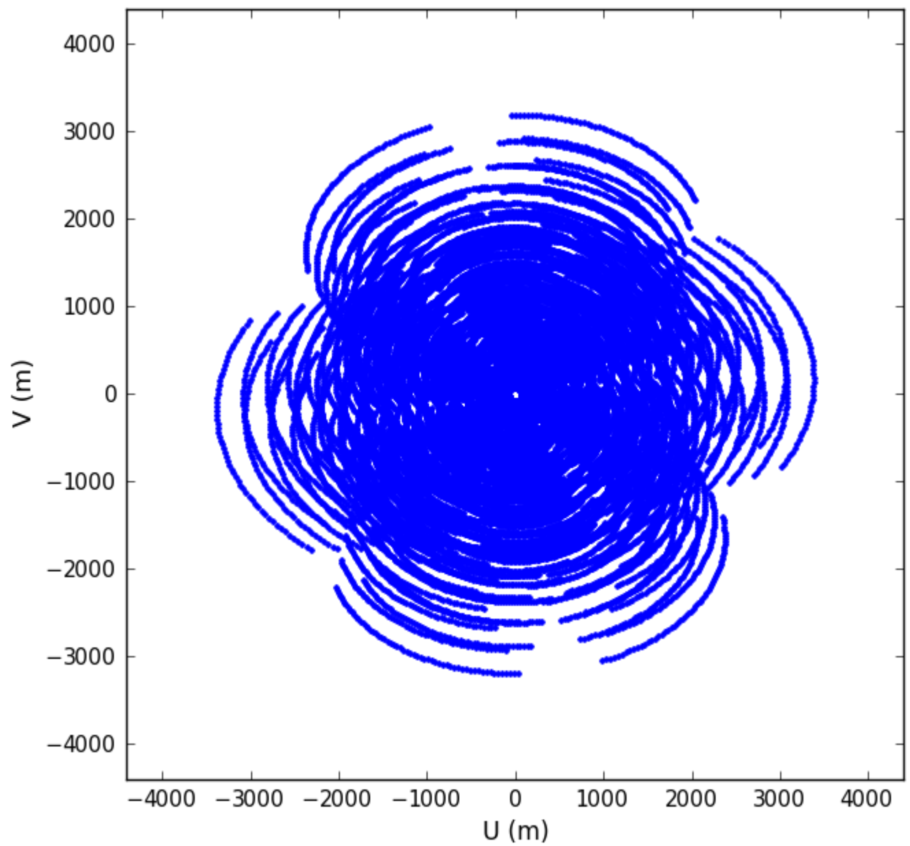
$$I^{obs}(l,m)$$

Spatial Frequency (uv) coverage + Observed Image

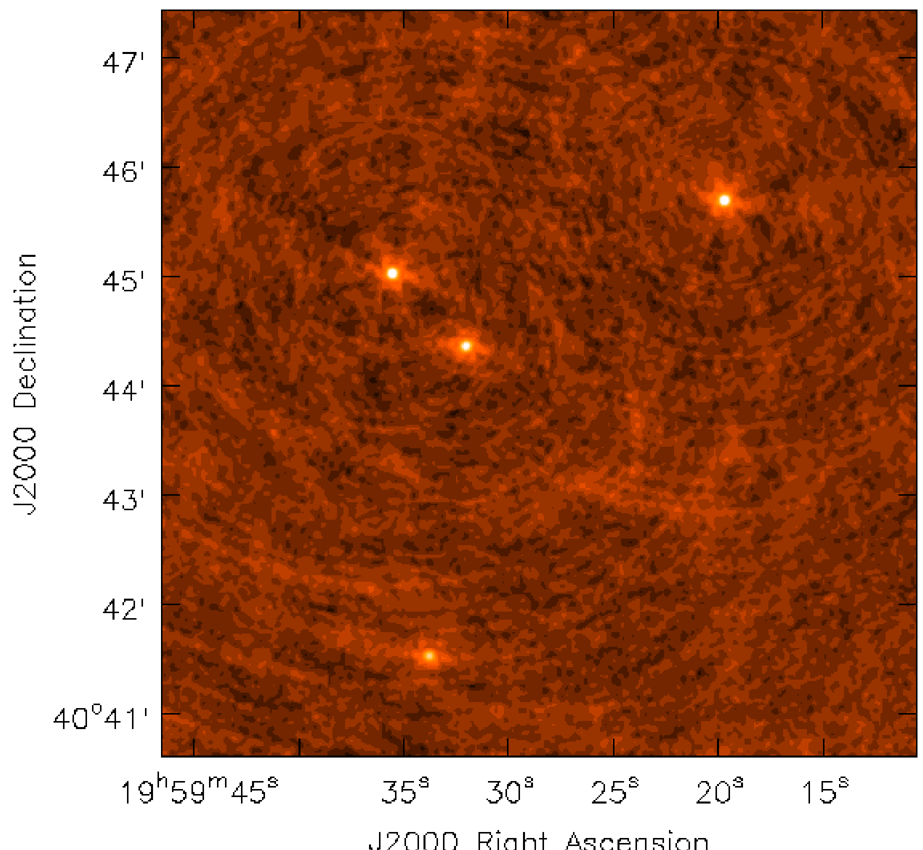
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

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





$S(u,v)$

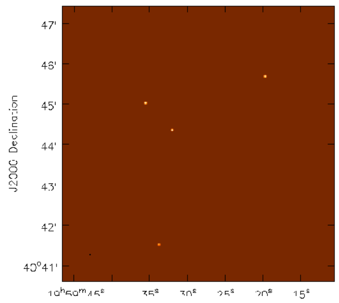


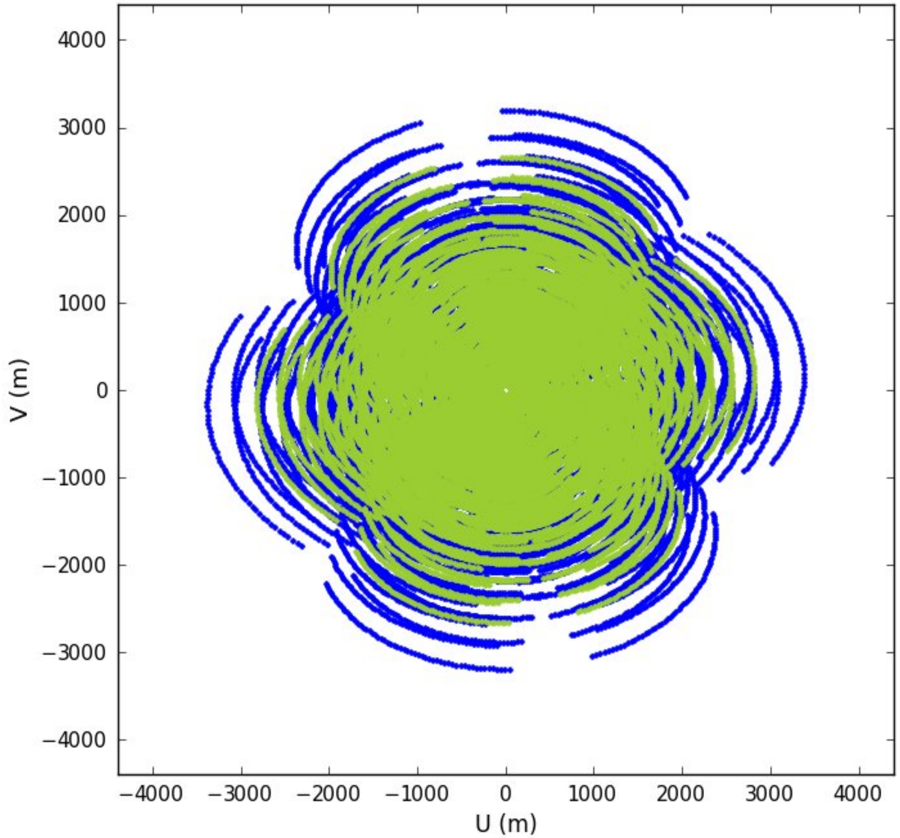
$I^{obs}(l,m)$

Spatial Frequency (uv) coverage + Observed Image

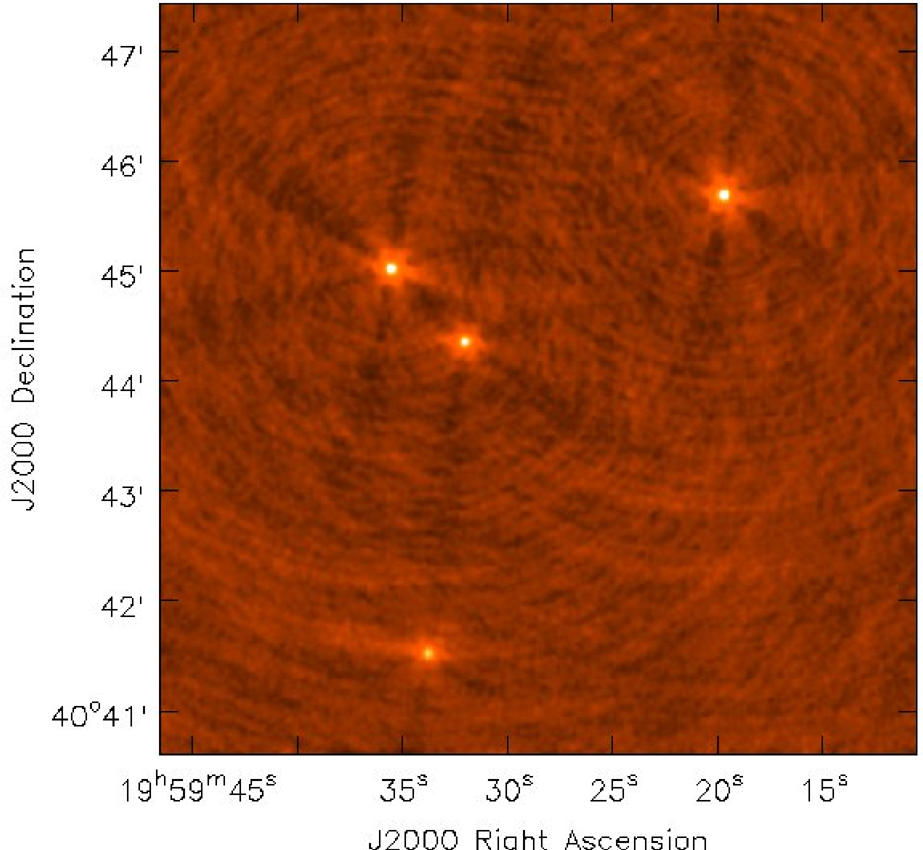
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

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





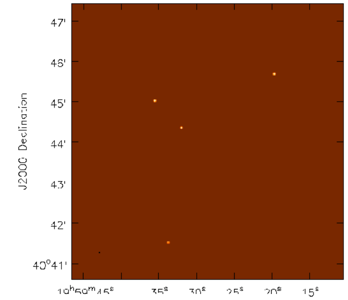
$S(u,v)$

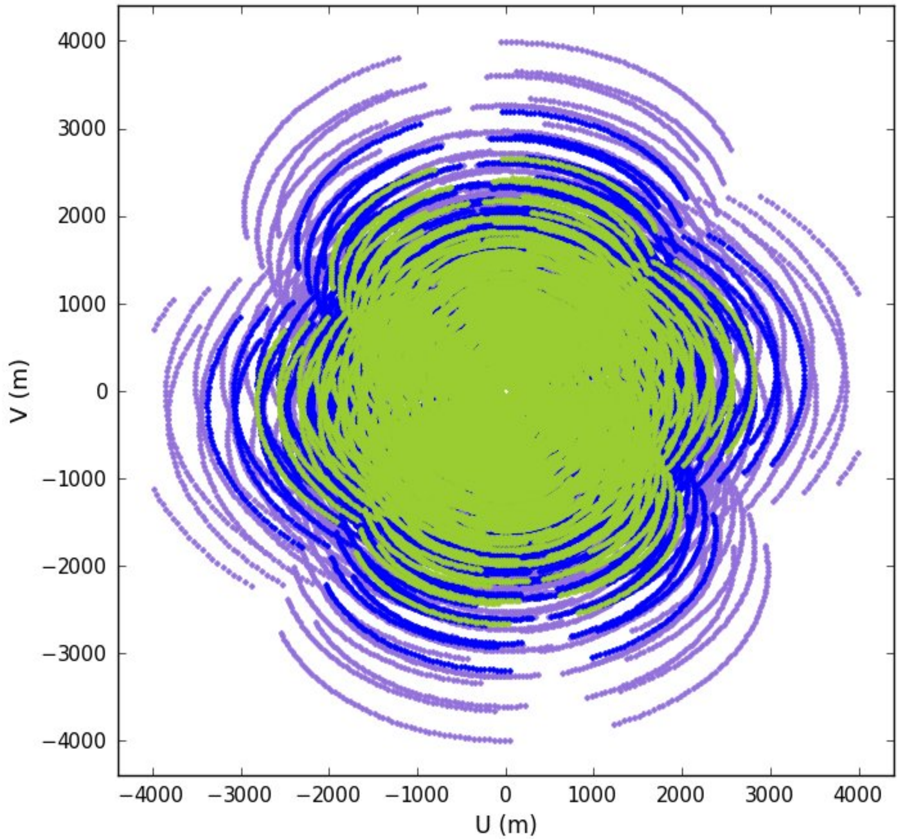


$I^{obs}(l,m)$

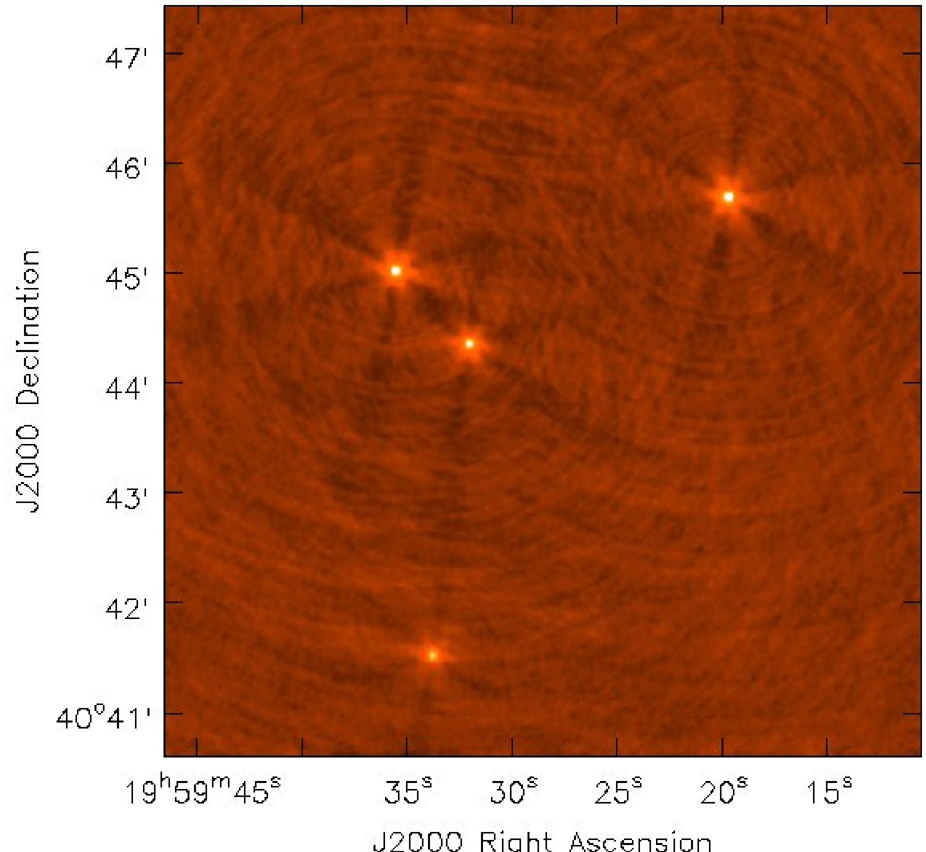
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} R(h,\theta) \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}$$

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





$S(u,v)$



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

Image formed by an interferometer : Convolution Equation

The visibility function is sampled at a finite number of spatial frequencies.

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

$$F^{-1}[V^{obs}(u,v)] = F^{-1}[S(u,v) \cdot V(u,v)]$$

$$I^{obs}(l,m) = F^{-1}[S(u,v)] * F^{-1}[V(u,v)]$$

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

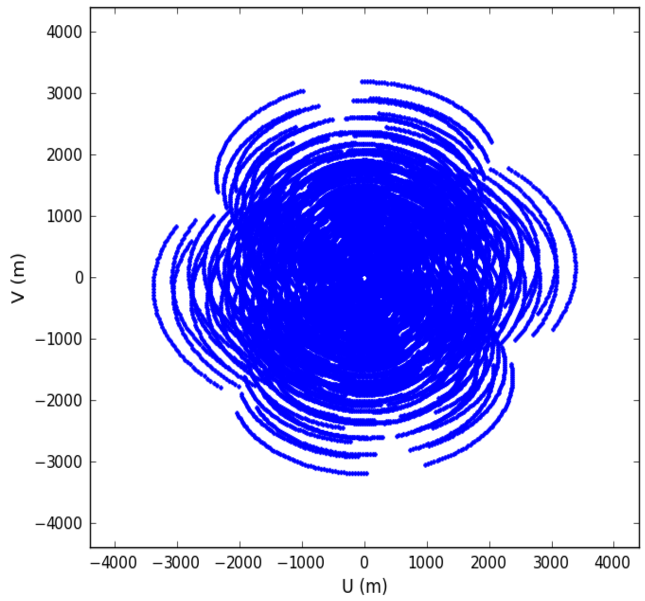
The observed image is a convolution between
the point-spread function and the true sky-brightness distribution

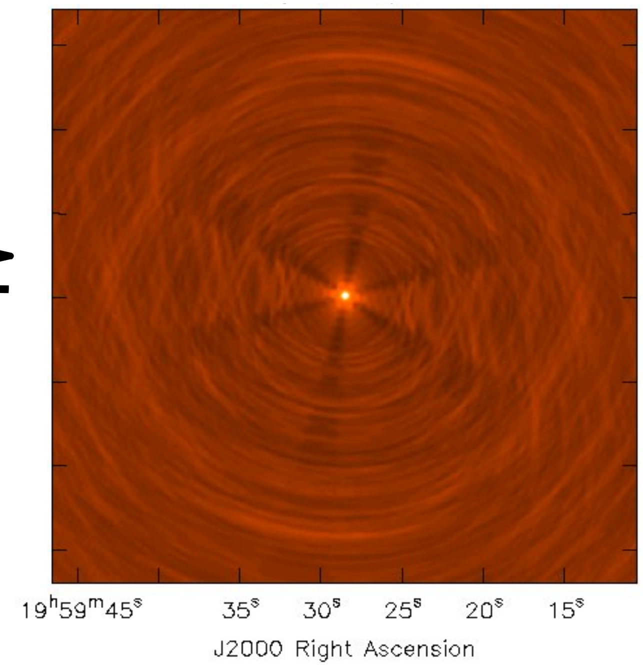
Point Spread Function (PSF)

$$S(u,v)$$

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

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





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)

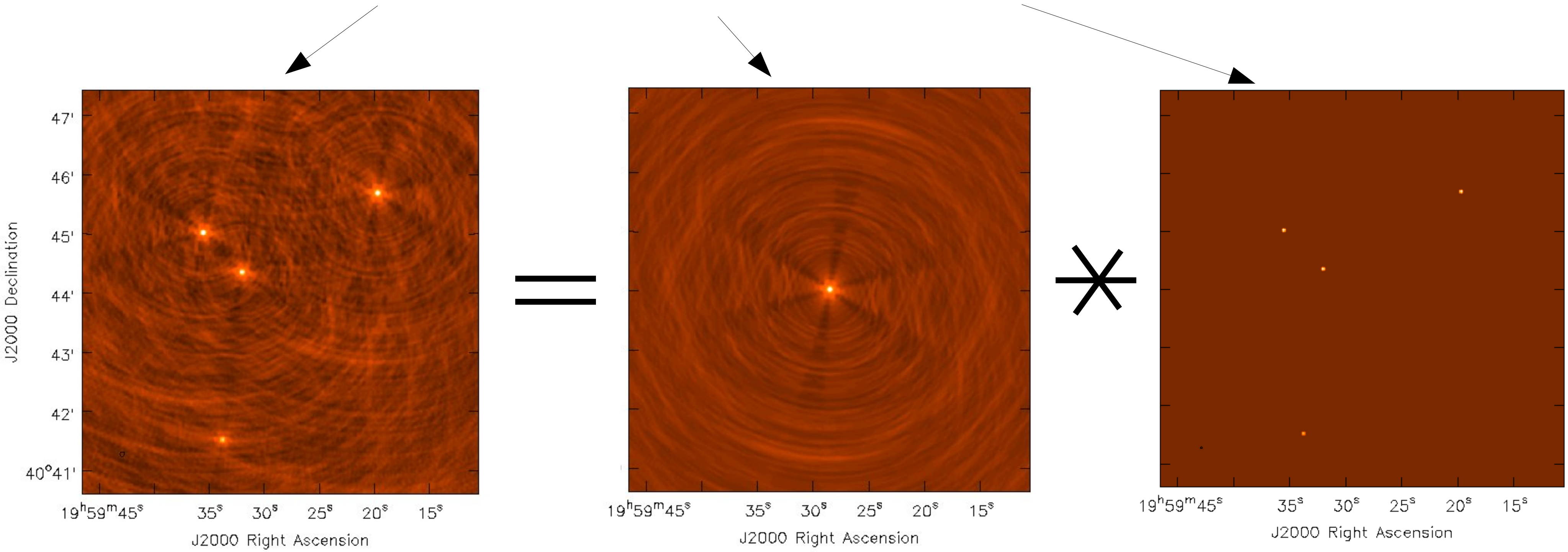
peak sensitivity, (# measurements)

sidelobe levels, (missing spacings)

no total power (central uv-hole)

Image formed by an interferometer : Convolution Equation

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



=> Deconvolution separates the PSF from the sky-brightness distribution

- estimates spatial frequencies in unmeasured regions of the uv-plane

To start..... need to construct the PSF and Observed (dirty) image.....

Imaging in practice

(1) Choose image-coordinates
- Image pixel size : Nyquist-sample the PSF main lobe
- #pixels : Field of view being imaged

(2) The uv-plane is sampled irregularly (along uv-tracks)

But, we use the Fast Fourier Transform (FFT) algorithm to form the image.

=> Need to resample the visibilities onto a regular grid before iFFT.

(3) Measured visibilities contain noise; some uv-ranges sampled more than others.

=> Choose how to 'weight' the visibilities during imaging.

(An image is the FT of a weighted-average of the data)

Make the Image : Grid the weighted visibilities
Make the PSF : Grid the weights

Imaging in practice : Image size, cell-size : image and uv-domains

– 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 $(\delta u, \delta v)$

- Data are recorded with finite $(\delta \tau, \delta \nu)$ => controls the minimum practical $(\delta u, \delta v)$

=> large field-of-view => small uv-grid cells => need high time/freq resolution
=> small image => large uv-grid cells => data averaging => loss of information

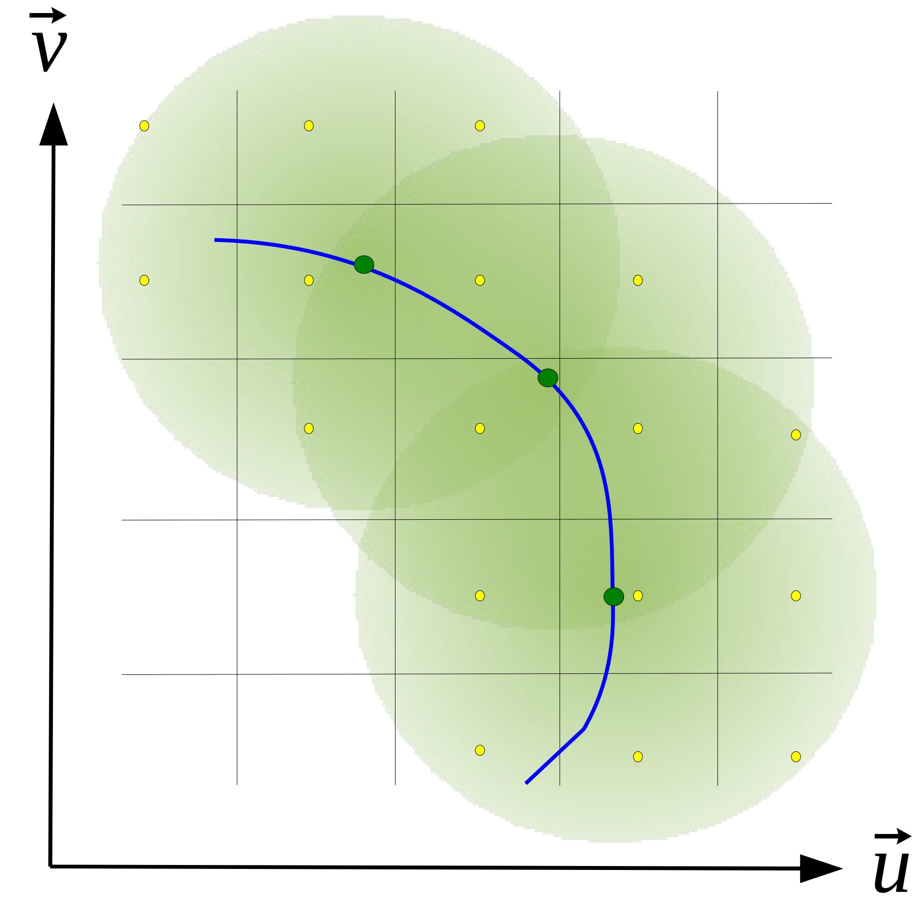
- Antenna primary-beam limits the field-of-view ('slits' of finite width)

Imaging in practice : Gridding + iFFT + Normalization

-- Data are recorded onto a regular grid via convolutional resampling

$$V_G(u,v)=[V(u,v)\cdot S(u,v)]*P_F(u,v)$$

Use Prolate Spheroidal Function
(good anti-aliasing operator)



-- Inverse FFT + take the real part

$$F^{-1}[V_G(u,v)]=I^{obs}(l,m)\cdot P_I(l,m)$$

$$P_I(l,m)=F^{-1}[P_F(u,v)]$$

-- Divide by iFT of the conv. function

$$I^{obs}(l,m)=\frac{F^{-1}[V_G(u,v)]}{P_I(l,m)}$$

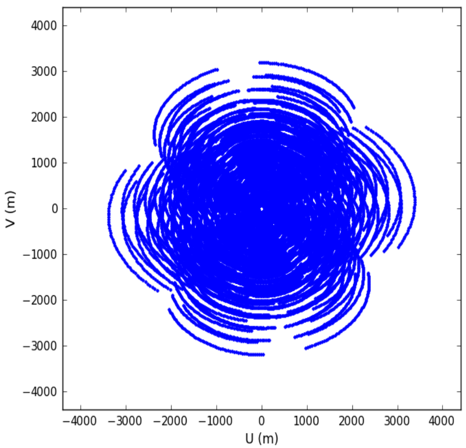
Anti-aliasing operator.....

UV-domain convolution with a few-pixel prolate-spheroidal function = Image-domain multiplication by an image-sized prolate-spheroidal function that damps down aliased power at image edges.

Imaging in practice : Weighting schemes

An Image is a weighted-average of the data.

$$I^{obs}(l,m) = \frac{F^{-1}\left[V(u,v)\cdot S(u,v)\cdot W(u,v)\right]_G}{\sum_{u,v}\left[W(u,v)\right]_G} \quad \text{where} \quad W(u,v) = \left(\frac{1}{\sigma^2}\right)\cdot ImWt$$



Choosing a weighting-scheme
=> modify the imaging properties of the instrument

	Uniform/Robust	Natural/Robust	UV-Taper
	All spatial-frequencies get equal weight	All data points get equal weight	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 : Weighting schemes

$$I^{obs}(l,m)=\frac{F^{-1}[V(u,v)\cdot S(u,v)\cdot W(u,v)]_G}{\sum_{u,v}[W(u,v)]_G}$$

where

$$W(u,v)=\left(\frac{1}{\sigma^2}\right)\cdot ImWt$$

- Natural Weighting : $ImWt=1$
- Uniform Weighting : $ImWt=\dfrac{1}{\rho(u_k,v_k)}$

  where $\rho(u_k,v_k)$ is the weight-density of all uv-points in the k^{th} uv-grid cell

  Super-uniform weighting : Consider density over uv-regions larger than one cell

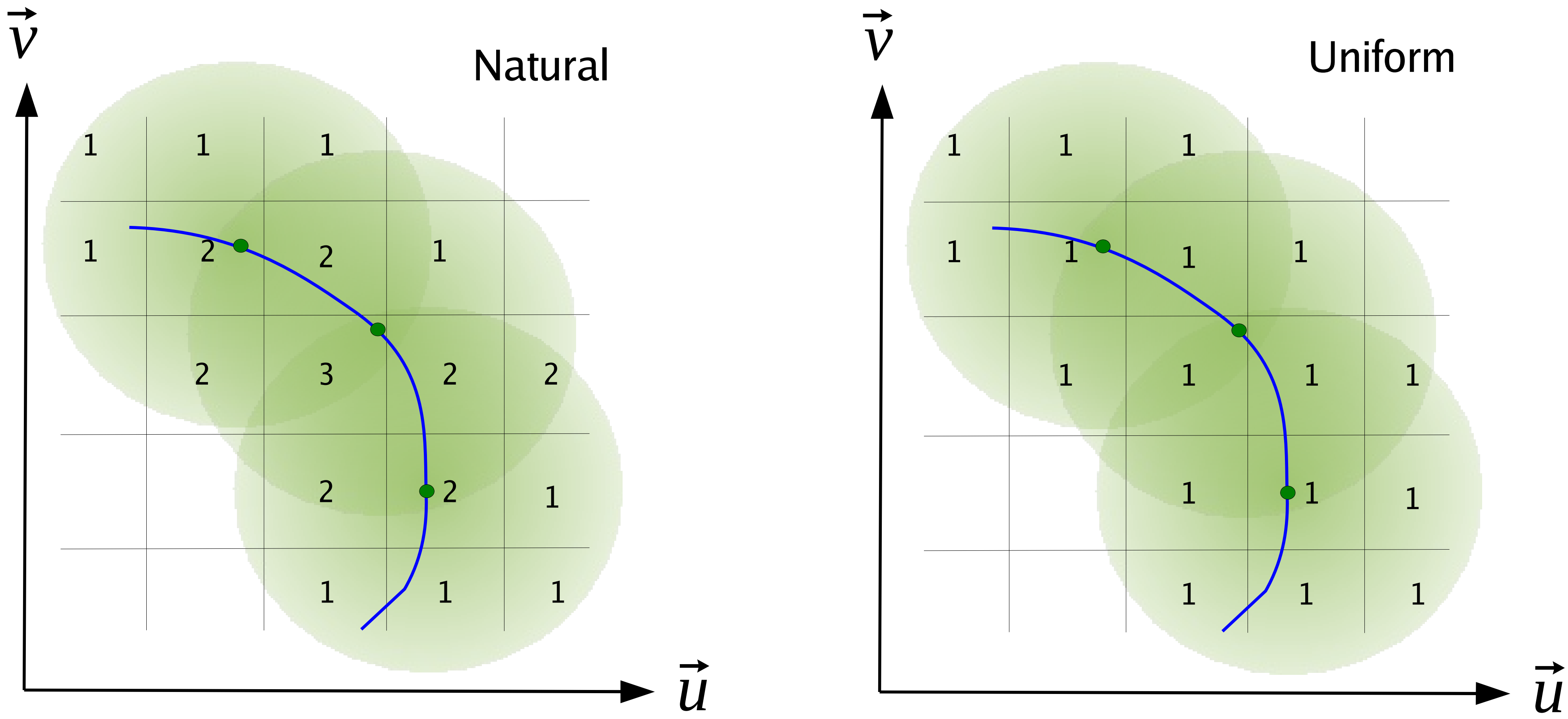
- Robust/Briggs Weighting : $ImWt=\dfrac{1}{[s\cdot\rho(u_k,v_k)+\sigma_k^2]}$

  where s,σ_k^2 control relative emphasis on signal vs noise (uniform vs natural)

- UV-taper : $ImWt=Gaussian(u,v)$ Emphasize shorter spatial frequencies

Imaging in practice : Weighting + Gridding

For illustration, let all visibilities have equal noise, and let $P_F(u,v)$ be a disk of ones.



Natural weighting : Densely-sampled regions the uv-plane get more weight
Uniform weighting : Sparsely-sampled regions get more relative weight
Robust weighting : In-between...

Imaging in practice : Weighting + Gridding + iFFT + Normalization

Putting it all together.......

$$I^{obs}(l,m) = \frac{F^{-1}\left[\left[V(u,v)\cdot S(u,v)\cdot W(u,v)\right] * P_F(u,v)\right]_G}{\left(\sum_{u,v}\left[W(u,v)\right]_G\right)\cdot P_I(l,m)}$$

Image

iFFT

Data

UV-coverage

Weights : $\left(\frac{1}{\sigma^2}\right)\cdot ImWt$

Gridding convolution function

Evaluate on regular grid

Sum of weights

$F^{-1}\left[P_F(u,v)\right]$

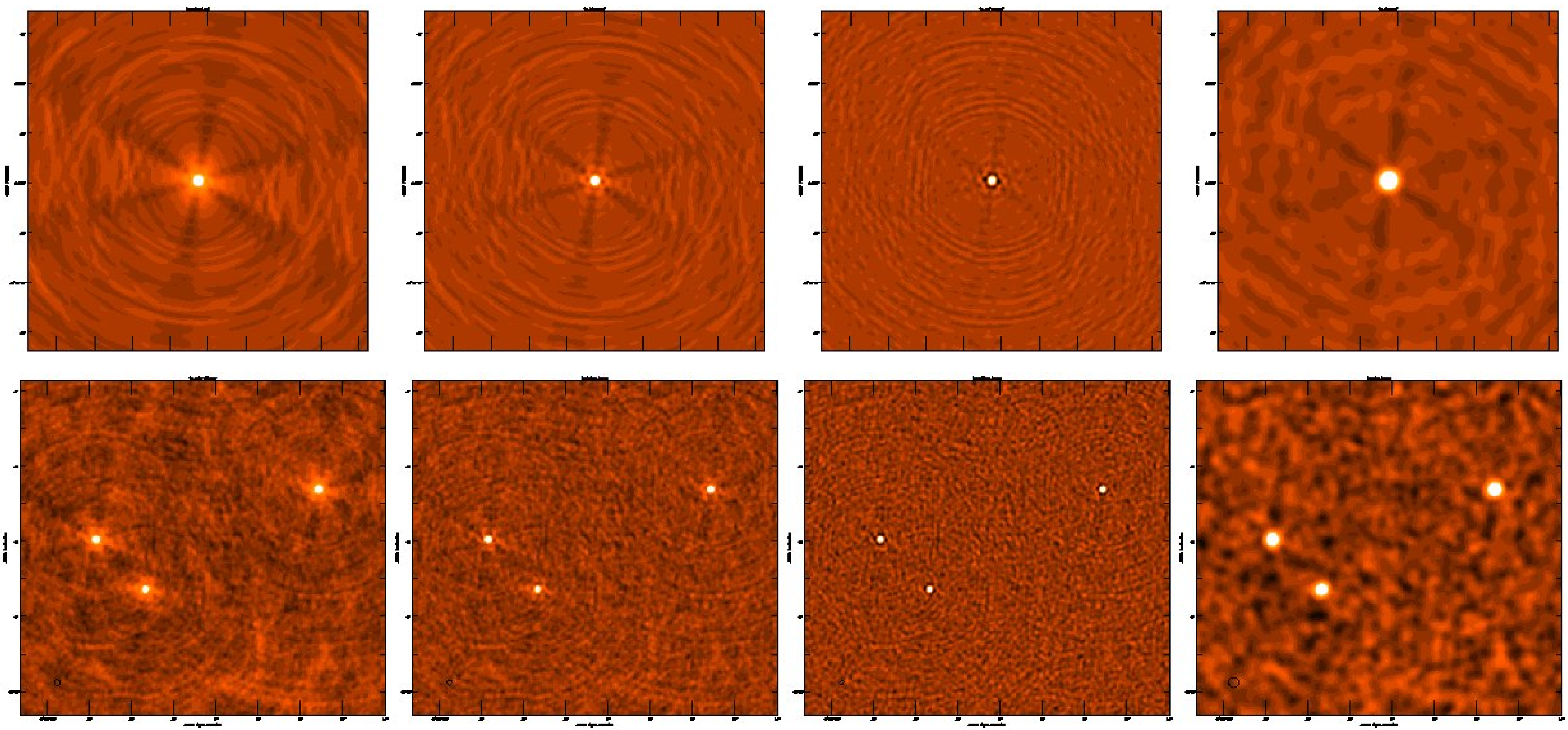
For the image : Grid 'weighted visibilities'

For the PSF : Grid the 'weights' (i.e. set $V(u,v)=1$)

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

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,-0.1 sidelobe	Bm : 8.0arcsec 0.01 sidelobe

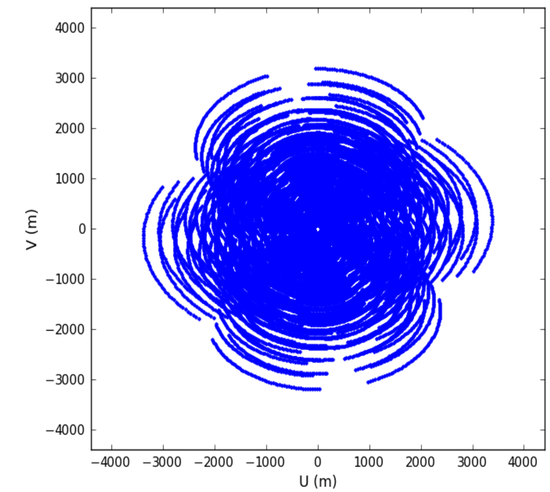


Note the noise-structure. Noise is correlated between pixels by the PSF.

Image Units (Jy/beam)

Image-Reconstruction (Deconvolution) Issues

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



Reconstructing I^{sky} :

Estimate the visibility function in unsampled regions of uv-space, such that it fits the data where it is sampled.

No unique solution. In fact, there are infinite solutions.
----- “invisible distribution “

Constrain the solution by forcing astrophysical plausibility (point-like compact structure, positive intensity, smooth extended emission, etc...)

But,

- There will always be un-resolved structure (max sampled spatial-freq)
  In most cases, it is unphysical to believe structure finer than the PSF beam-width

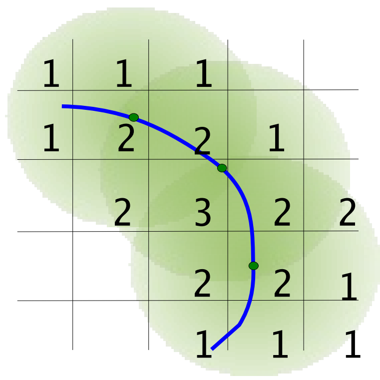
- Total integrated power is never measured
  Reconstruction of largest spatial scales is always an extrapolation

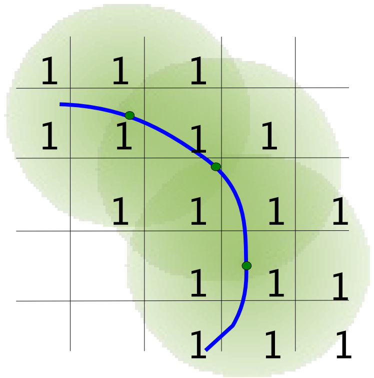
Deconvolution – Linear vs Non-Linear

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

Linear Deconvolution :

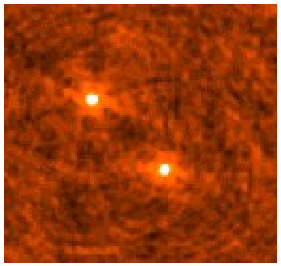
$$I^{sky}(l,m) = F^{-1}\left[\frac{F[I^{obs}(l,m)]}{F[I^{PSF}(l,m)]}\right]$$

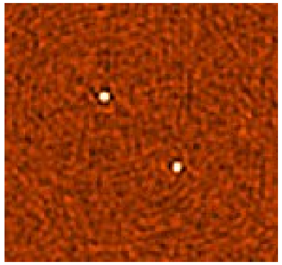




But..... this is just Uniform/Robust weighting !

(Divide gridded visibilities by gridded weights
– can also be done as a Wiener filter)





BUT, linear deconvolution cannot estimate visibilities in unsampled regions of the spatial frequency plane.

Need a non-linear approach, which iteratively fits a model that predicts the visibility function all over the spatial-frequency plane.

Deconvolution – Non-Linear, iterative image-reconstruction

Image Reconstruction : Iteratively fit a sky-model to the observed visibilities.

Measurement Equation : $[A] I^m = V^{obs}$ (using matrix /linear-algebra notation...)

- The operator $[A] = [S][F]$ includes the UV-coverage and FT
- The vector I^m is the sky model (e.g. image-pixels, Gaussian set)

Fit the parameters of I^m via a weighted least-squares optimization :

- Minimize $\chi^2 = [V^{obs} - A I^m]^T W [V^{obs} - A I^m]$ ==> $\frac{\delta \chi^2}{\delta I^m} = 0$

Deconvolution – Non-Linear, iterative image-reconstruction

Image Reconstruction : Iteratively fit a sky-model to the observed visibilities.

Measurement Equation : $[A]I^m = V^{obs}$ (using matrix /linear-algebra notation...)

- The operator $[A]=[S][F]$ includes the UV-coverage and FT
- The vector I^m is the sky model (e.g. image-pixels, Gaussian set)

Fit the parameters of I^m via a weighted least-squares optimization :

- Minimize $\chi^2 = [V^{obs} - A I^m]^T W [V^{obs} - A I^m]$ ==> $\frac{\delta \chi^2}{\delta I^m} = 0$

Normal Equations : $[A^T W A] I^m = [A^T W] V^{obs}$

- This describes an image-domain convolution $I^{psf} * I^m = I^{dirty}$

Iterative Solution : $I^m_{i+1} = I^m_i + g[A^T W A]^+ \left(A^T W \left(V^{obs} - A I^m_i\right)\right)$

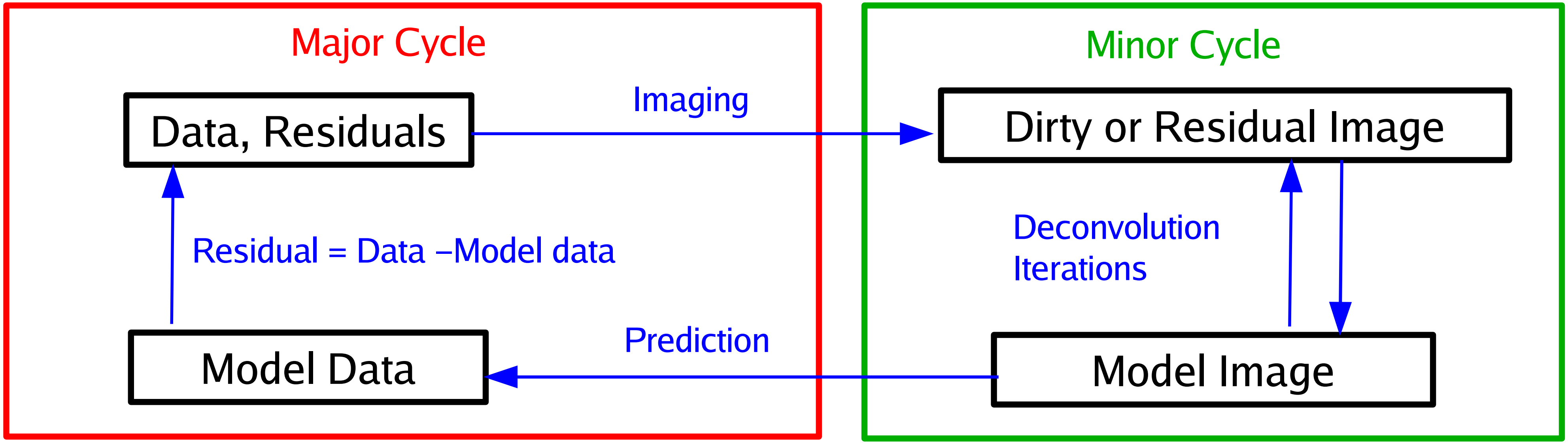
Deconvolution

Imaging (Gridding + iFT)

Prediction (FT + de-Gridding)

Deconvolution – Non-Linear, iterative image-reconstruction

Image Reconstruction : Iteratively fit a sky-model to the observed visibilities.



Normal Equations : $[A^T W A] I^m = [A^T W] V^{obs}$

– This describes an image-domain convolution $I^{psf} * I^m = I^{dirty}$

Iterative Solution : $I^m_{i+1} = I^m_i + g[A^T W A]^+ \left(A^T W \left(V^{obs} - A I^m_i\right)\right)$

Deconvolution

Imaging
(Gridding + iFT)

Prediction
(FT + de-Gridding)

Deconvolution Algorithms + Image Restoration

(Minor cycle) deconvolution algorithms differ in choice of sky-model, optimization scheme, and how they handle coupled parameters.

Classic CLEAN : Point-source sky model, Steepest-descent optimization

Maximum Entropy Method : Point-source sky model with a smoothness constraint.
Steepest-descent optimization with backtracking

Multi-Scale CLEAN : Sky is a linear combination of components of different known shapes/sizes. Steepest-descent optimization

Adaptive-Scale-Pixel CLEAN : Sky is a linear combination of best-fit Gaussians.
BFGS optimization.

- Several adaptations of compressed-sensing reconstruction techniques (R&D)

Output of deconvolution (minor cycle) : A model image (units : Jy/pixel)
A residual image (units : Jy/beam)

Restoration : Convolve model with a 'clean beam' (Gaussian fit to PSF main lobe)
Add in residual image. (units : Jy/beam)

Deconvolution – Hogbom CLEAN

Sky Model : List of delta-functions : $I^{sky} = \sum_x a_x \, \delta(x)$

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

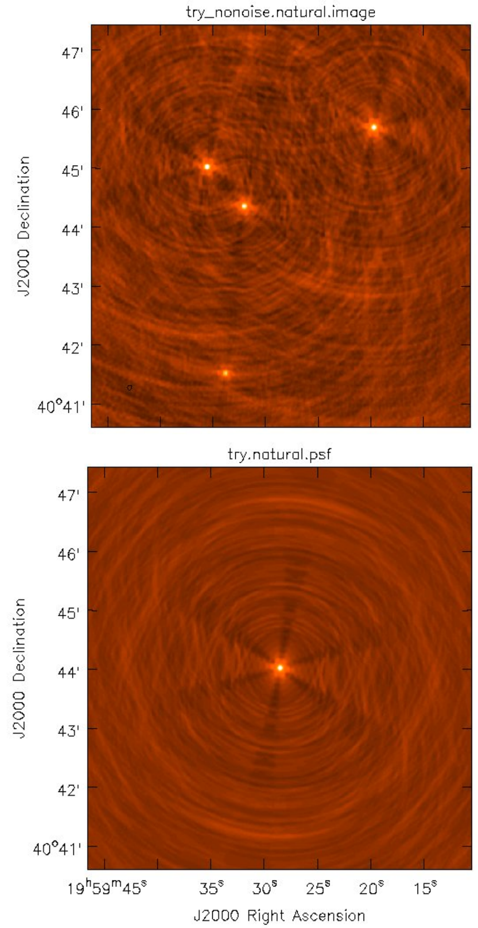
(2) Search for the location of peak amplitude.

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

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

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

(5) Restore the model using a 'clean beam' and adding in final residuals



Variants : Clark CLEAN (use psf patches for updates, calculate residuals using gridded visibilities)
Cotton-Schwab CLEAN (Periodically predict model-visibilities, calculate residual visibilities and re-grid –major and minor cycles)

Deconvolution – CLEAN

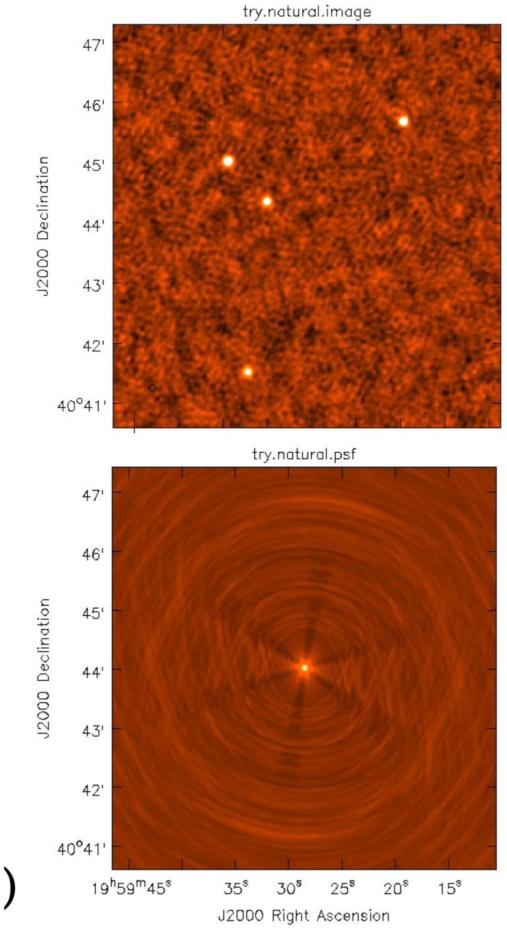
- This is a steepest-descent minimization, where one component is added per iteration.

- Step-size per iteration is controlled by a loop gain ~ 0.1
  - update the model and residuals by a fraction of the peak

- A scale-insensitive algorithm. Good for point-source Dominated fields. Not good for extended emission.

- Stopping criteria are either a maximum number of iterations or a flux-threshold below which components cannot be trusted.

  Flux threshold
  = multiple of noise rms (when noise dominates)
  = fraction/multiple of peak residual (when artifacts dominate)



- Windowing functions (clean boxes) can be used to constrain the source locations

- Point-sources not at pixel centers sometimes cause artifacts (use smaller pixels)

Deconvolution – MEM

Sky Model : List of delta functions $I^{sky} = \sum_x a_x \, \delta(x)$ with a smoothness constraint

A constrained optimization where the function being minimized is :

$$\left(I^m * I^{psf} - I^{obs}\right)^2 - a\, I^{sky} \log\left(\frac{I^{sky}}{I^{prior}}\right)$$

I^{prior} is a 'default image' that biases the reconstruction in the absence of sufficient constraints from the data.

- A 'flat prior' encourages smoothness in the reconstruction (extended emission)
- A low-resolution image (from single-dish observations) provides constraints on extended emission that may fall within the uv-hole.

Useful when the final image is some combination of images (mosaics, single-dish)

Every pixel is a potential degree of freedom –a scale-insensitive algorithm
Convergence-rate is slow for high dynamic-range images.

MEM can also be described using Bayesian analysis and conditional probability

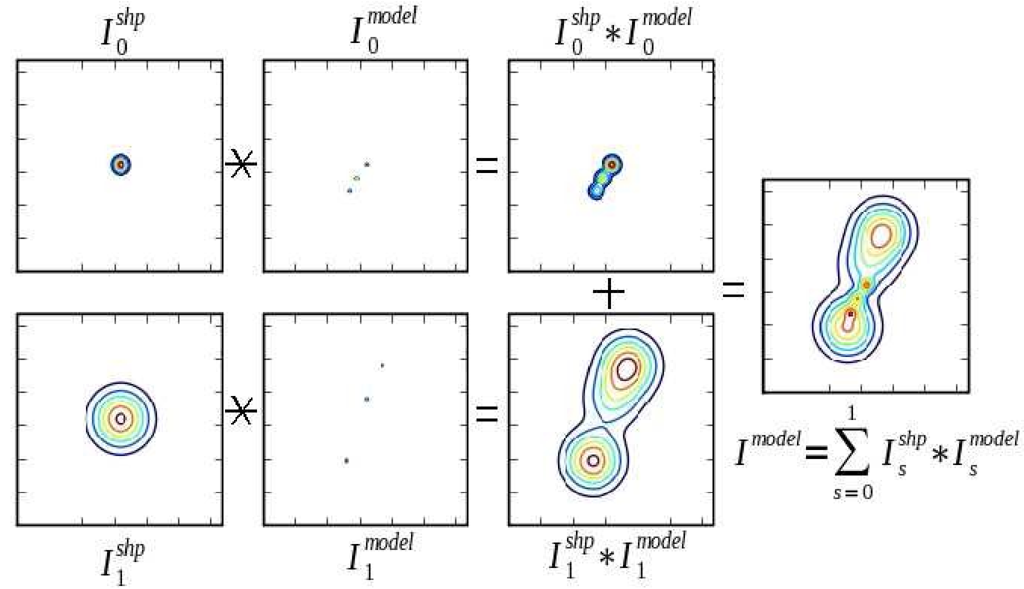
Deconvolution – MS-CLEAN

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

$$I^{sky} = \sum_s [I_s^{shp} * I_s^m]$$

where I_s^{shp} is a blob of size 's'

$$I_s^m = \sum_i a_{s,i} \delta(l - l_{s,i})$$



A scale-sensitive algorithm

(1) Choose a set of scale sizes –define a basis set

(2) Calculate diry/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 with Major and Minor cycles

Deconvolution – MS-CLEAN

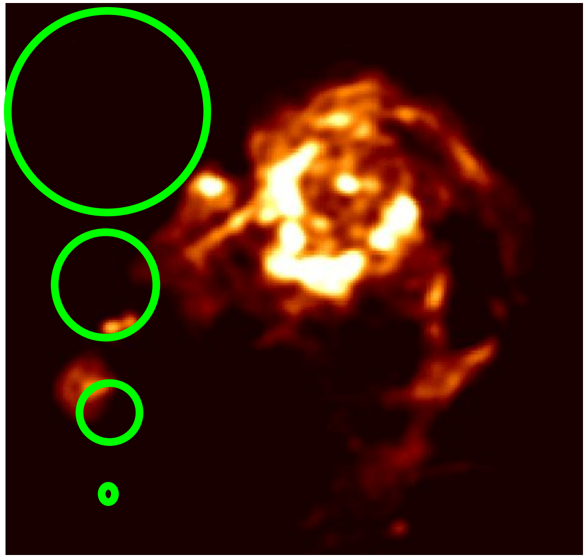
- Efficient representation of both compact and extended structure (sparse basis)

- MS-CLEAN naturally detects and removes the scale with maximum power

- The implementation efficiency of CLEAN is retained (model and residual updates)

- This method could use any basis set to bias the reconstruction (shapelets)

  - Inverted truncated paraboloids of a few fixed sizes.
  - Choose sizes that match dominant image patterns.
  - Always include a delta-function



- Can use higher loop-gains than with CLEAN because the model is more accurate

- Multi-Resolution CLEAN : A variant of MS-CLEAN where minor cycles are performed separately on images tapered to different resolutions.

Deconvolution – Adaptive Scale Pixel (ASP) CLEAN

Sky Model : List of Gaussians $I^{sky} = \sum_c a_c e^{-\frac{(x-x_c)^2}{\sigma^2}}$

(1) Calculate the dirty image, smooth to a few scales.

(2) Find the peak across scales to identify a good initial guess of a_c, x_c, σ_c for a new component.

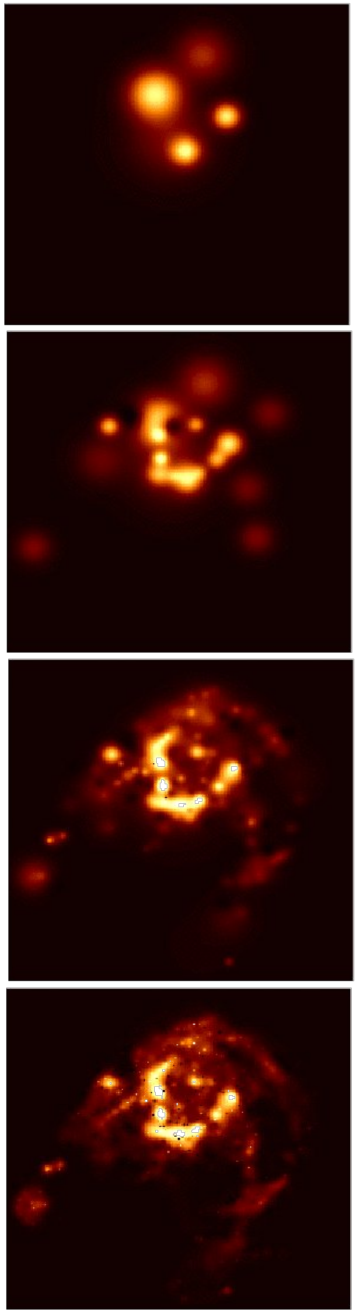
(3) Add this component to a list.

(4) Choose a subset of components most likely to have a significant impact on convergence. Re-fit Gaussian parameters for new and old components together.

(5) Subtract the contribution of all updated components from the dirty image.

Repeat steps (2)-(5) until a stopping criterion is reached.

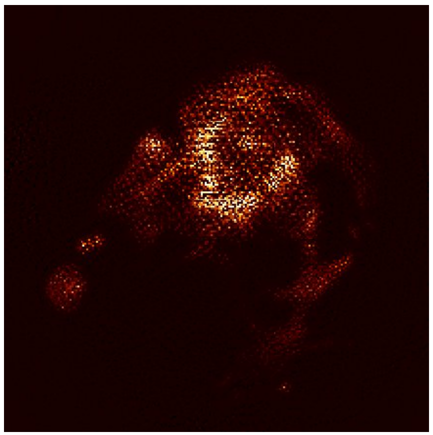
Adaptive Scale sizes leads to better reconstruction than MS-Clean, and more noise-like residuals.

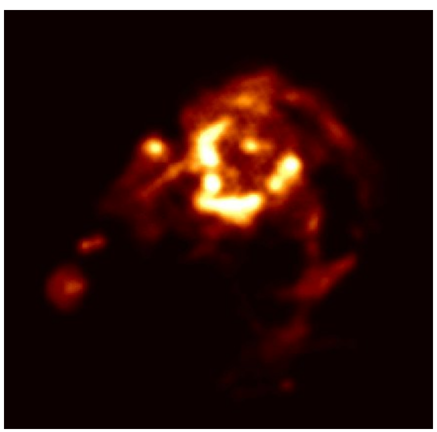


Deconvolution – Comparison of Algorithms

CLEAN

Minimize L2 (assume sparsity in the image)

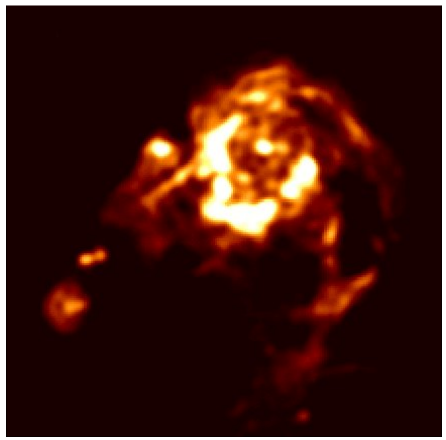


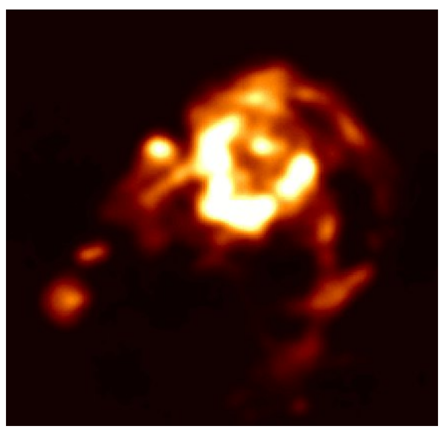


(Hogbom 1974, Clark 1980, Schwab & Cotton 1983)

MEM

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

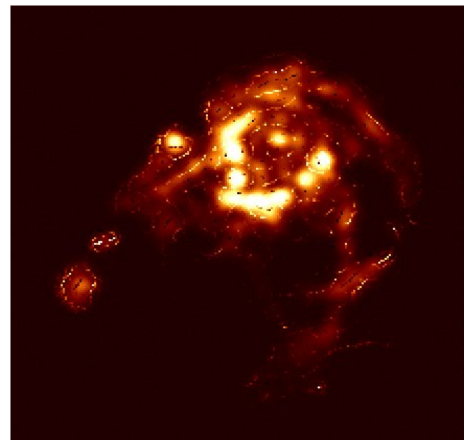


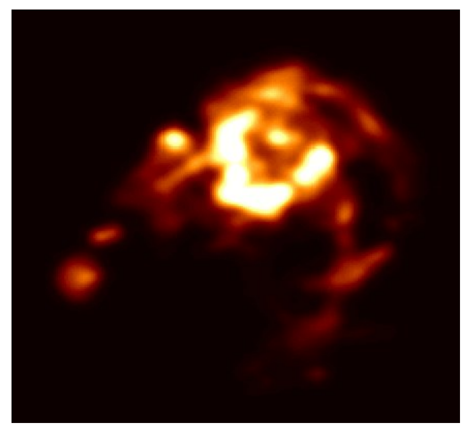


(Cornwell & Evans, 1985)

MS-CLEAN

Minimize L2 (assume a set of spatial scales)

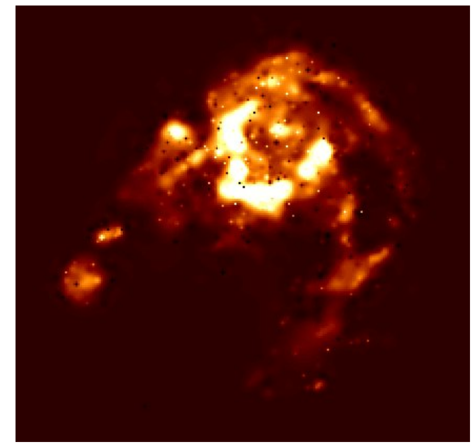


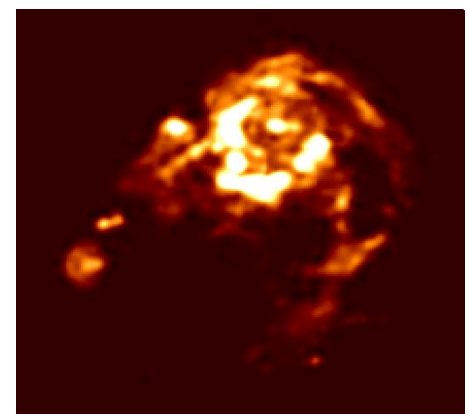


(Cornwell, 2008)

ASP

Minimize L2 with TV-based subspace searches





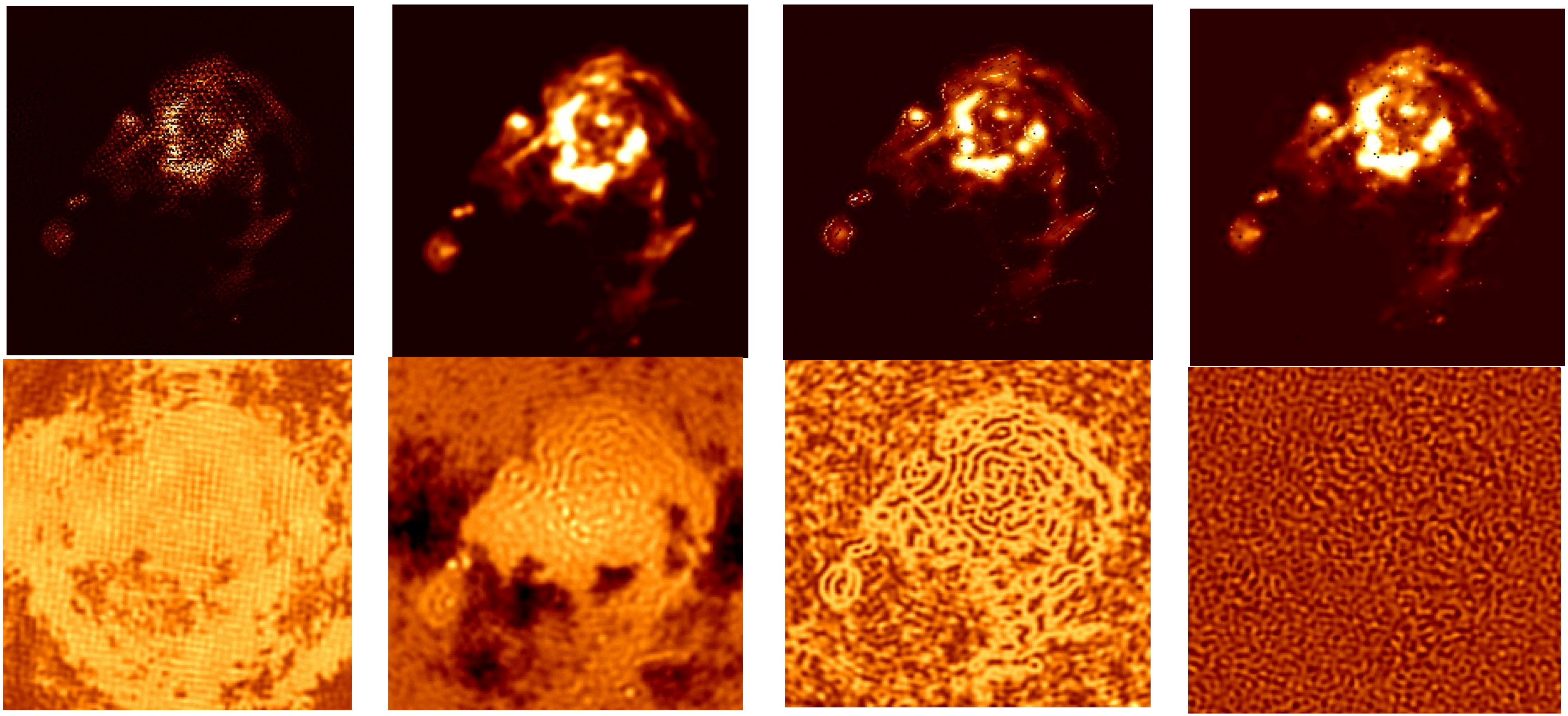
(Bhatnagar & Cornwell 2004)

I^m

I^{out}

Deconvolution – Comparison of Algorithms

CLEAN	MEM	MS-CLEAN	ASP
Minimize L2 (assume sparsity in the image)	Minimize L2 subject to an entropy-based prior (e.g. smoothness)	Minimize L2 (assume a set of spatial scales)	Minimize L2 with TV-based subspace searches



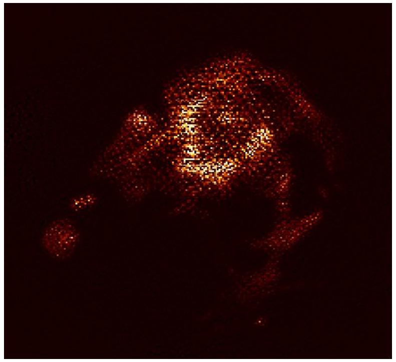
I^m

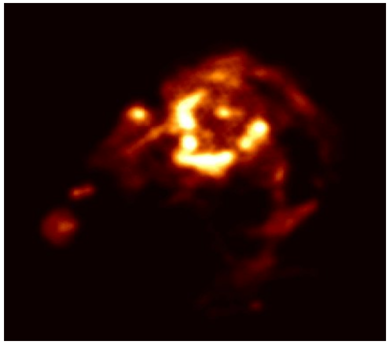
I^{res}

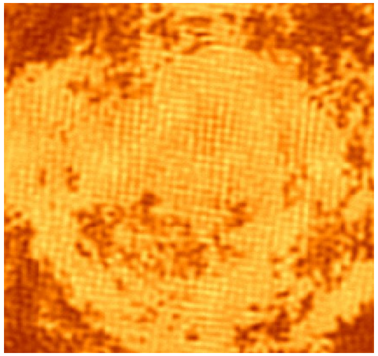
(Hogbom 1974, Clark 1980, Schwab & Cotton 1983)	(Cornwell & Evans, 1985)	(Cornwell, 2008)	(Bhatnagar & Cornwell 2004)

Image Quality

Model image : Raw components added together
Restored image : Image for astrophysical interpretation
Residual image : Noise, left-over (undeconvolved) flux, artifacts







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)

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 :

$$\frac{I^m * I^{beam}}{I^m * I^{beam} - I^{restored}}$$

Iterations of Self-calibration and Imaging

Measurement Equation : $$V_{ij}^{obs}(\nu,t) = M_{ij}(\nu,t)\, S_{ij}(\nu,t) \iint I(l,m)\, e^{2\pi i(ul+vm)}\, dl\, dm$$

$$M_{ij}(\nu,t) \rightarrow g_i g_j^*$$

In general, $M_{ij}(\nu,t)$ and $I(l,m)$ are both unknown.

In standard calibration, a known source is observed and $M_{ij}(\nu,t)$ is calculated

$M_{ij}(\nu,t)$ is interpolated/extrapolated to target time/freq ranges and divided-out before imaging the target source.

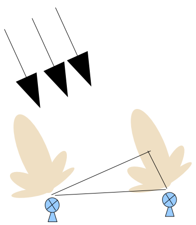
Sometimes, these gain solutions are not good-enough (image artifacts persist)
(instrumental variability between calibrator scans, calibrator data loss...)

=> Iterate a few times between solving for $M_{ij}(\nu,t)$ and for $I(l,m)$.

Self-Calibration is often required for high dynamic-range imaging........

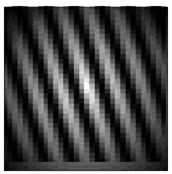
A more-realistic Measurement Equation

$$V_{ij}^{obs}(\nu,t) = M_{ij}(\nu,t)\, S_{ij}(\nu,t) \iiint M_{ij}^{S}(l,m,\nu,t)\, I(l,m,\nu,t)\, e^{2\pi i(ul+vm+w(n-1))}\, dl\, dm\, dn$$



Primary Beam effects

- A dish is not an infinitely-small 'slit'



- Rotation while tracking, pointing offsets, frequency-scaling

Sky-brightness varies with time and freq.

- All sources have spectral structure

- Some vary with time

W-Term

-Non-coplanar baselines

-Sky curvature

- Increased imaging sensitivity (lower Tsys, larger bandwidth, larger collecting-area)

=> Artifacts that were earlier below the noise, are visible in the image.
- direction, frequency and time-dependence of the sky and instrument.

- Need wider fields-of-view from a single observation (pressure on telescope time...)

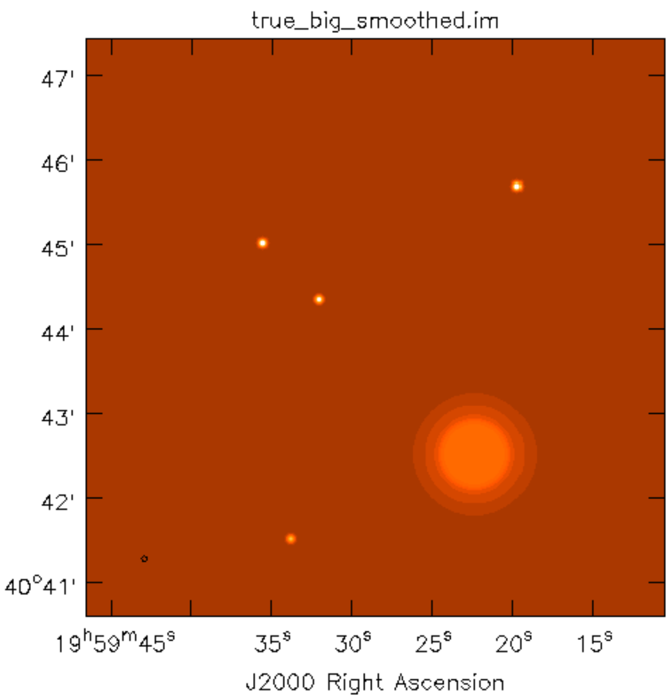
See later talks on high dynamic-range imaging, error recognition, wideband imaging, wide-field imaging and mosaicing (including single-dish data).

Example Imaging Problem – Simulated data

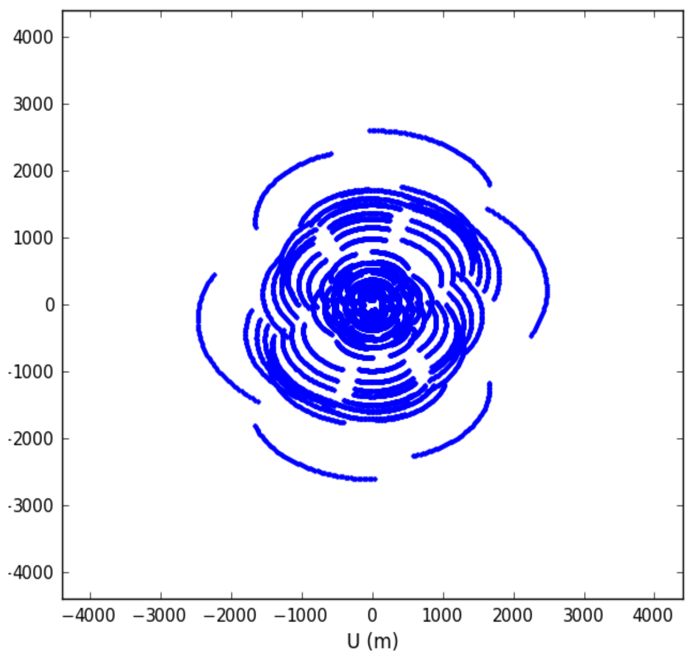
Simulated 5 GHz observation with a 13-antenna array over 5 hours

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

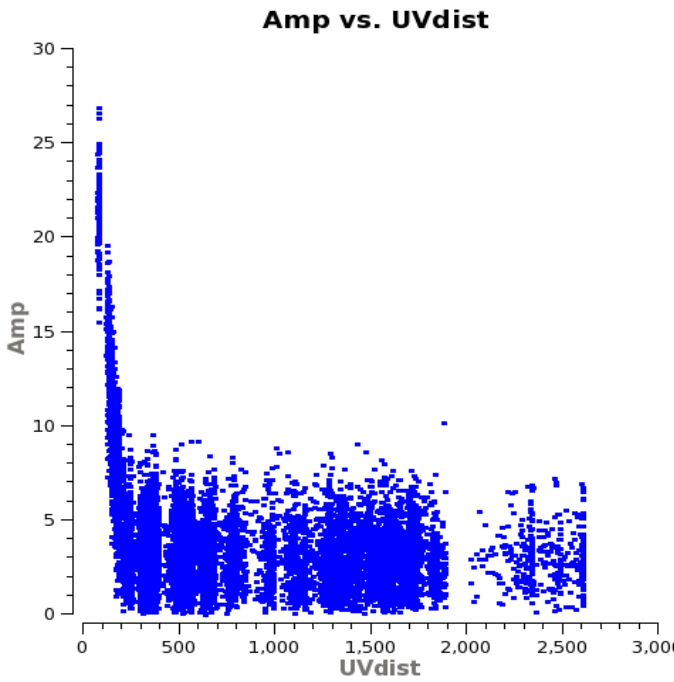
Angular resolution : 5 arcsec (Max baseline of 2500m at 5.0 GHz)
Sky brightness has compact and extended structure (partially-sampled).
Peak brightness : 1 Jy => Target dynamic range = 50



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



$S(u,v)$



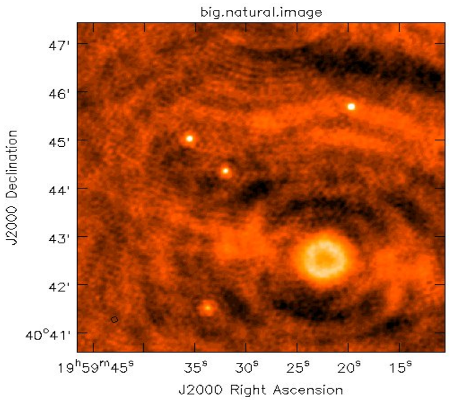
$V^{sky}(u,v) \cdot S(u,v)$

Example Imaging Problem – First try....

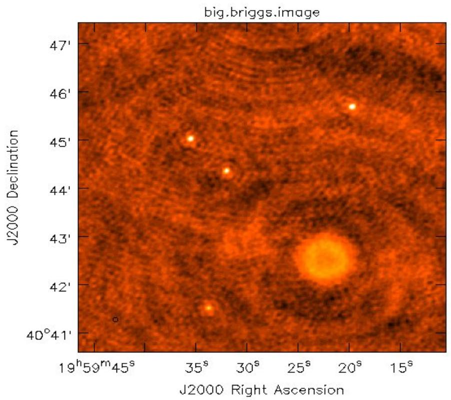
Quick deconvolution with different weighting schemes :

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

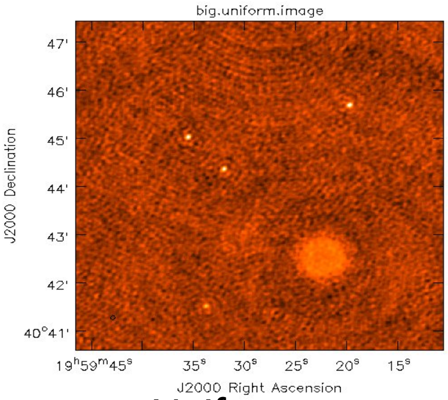
MS-CLEAN : NIter=100, scales=[0,6,40], gain=0.3, robust=0.7



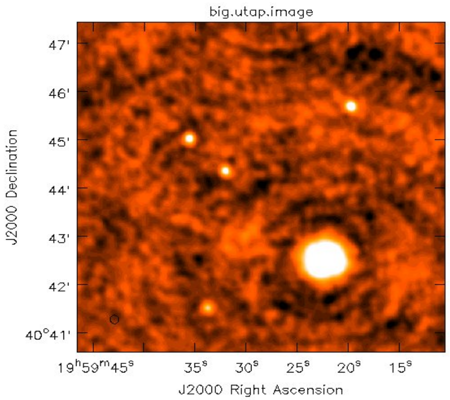
Natural
High sidelobes



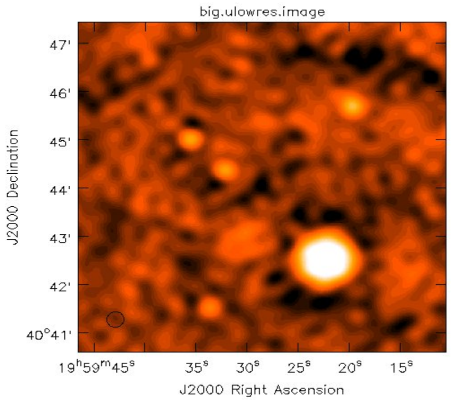
Robust = 0.7



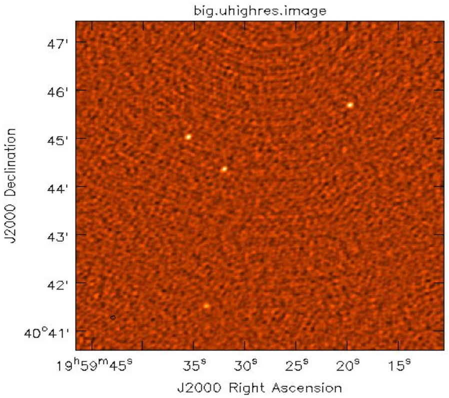
Uniform
Low sensitivity to extended emission



Uniform with a uv-taper for 9 arcsec



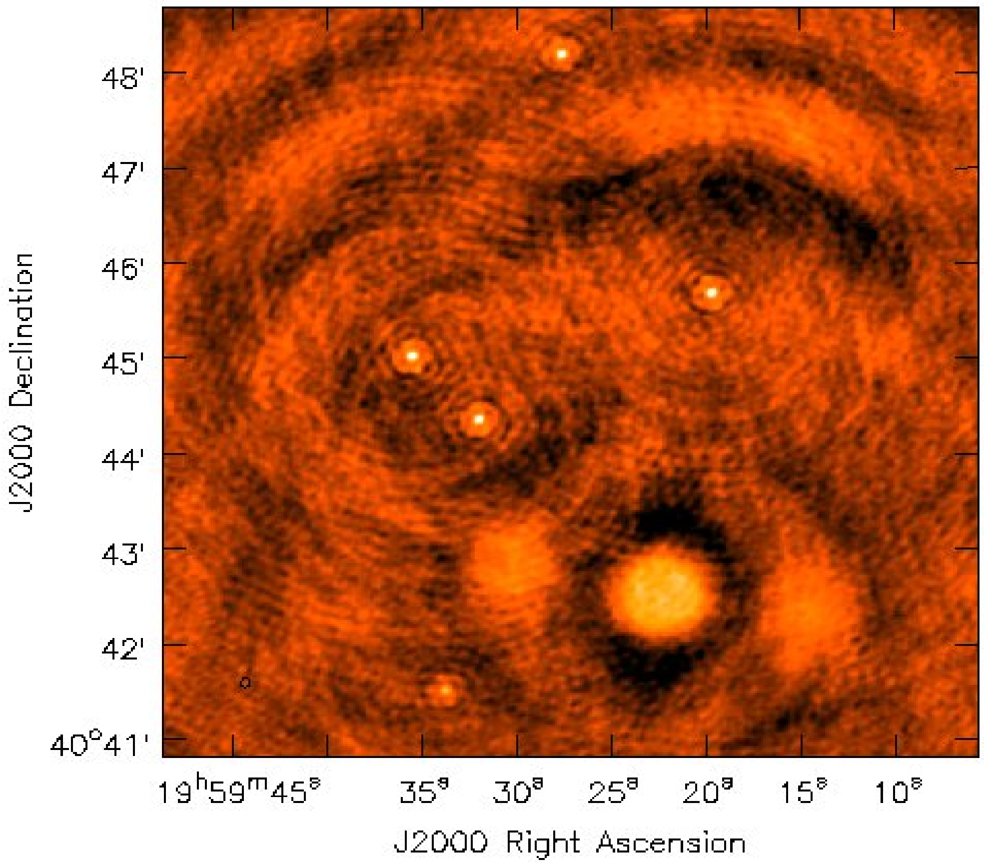
Uniform with only SHORT Baselines < 500m



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

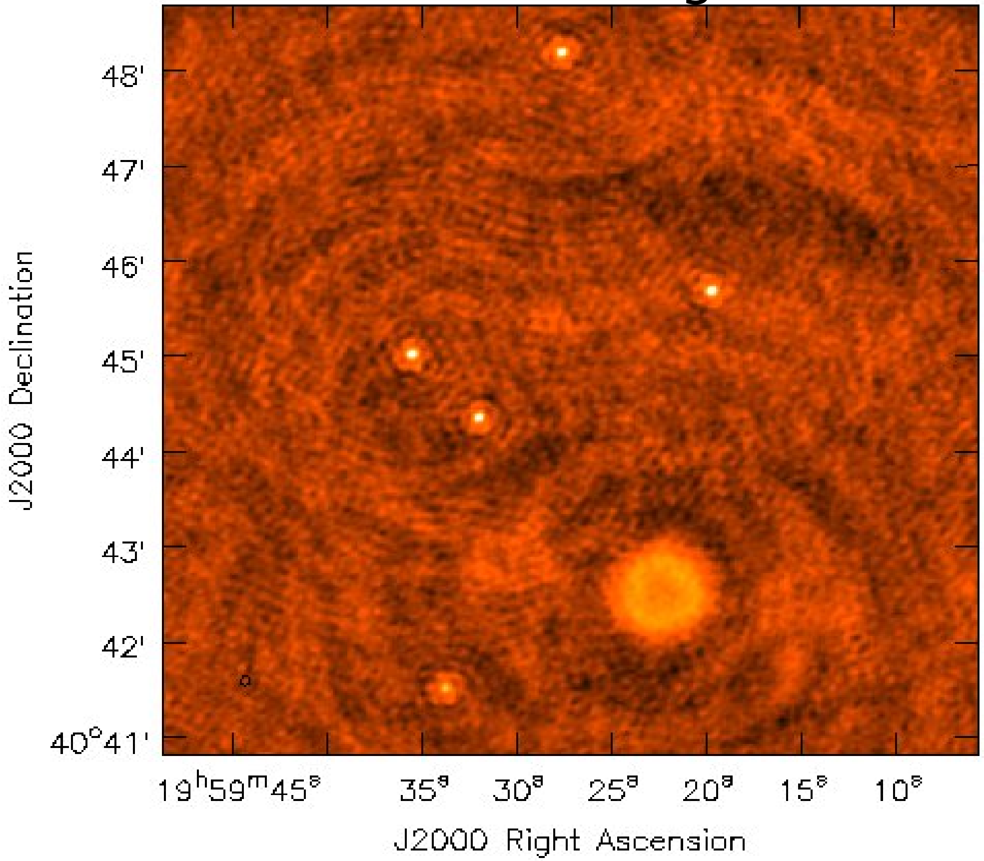
Loop gain = 0.2
(go slow, because of insufficient data-constraints for the extended emission)

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

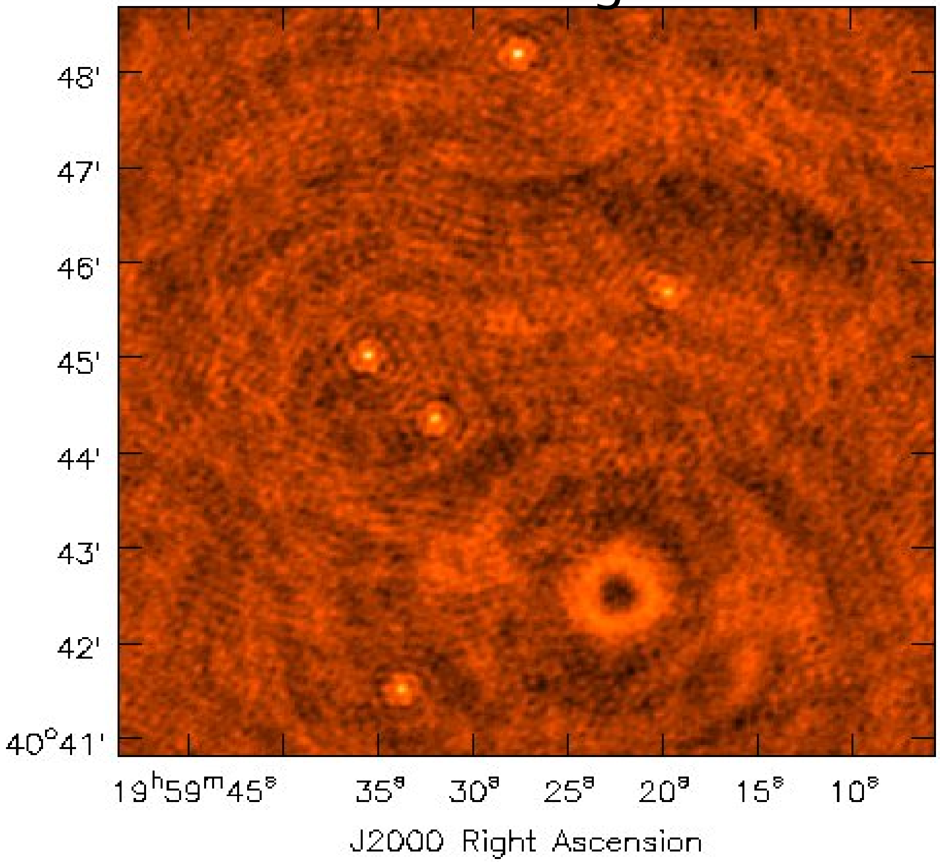
Example Imaging Problem – Second try...

After 100 iterations.

Restored Image



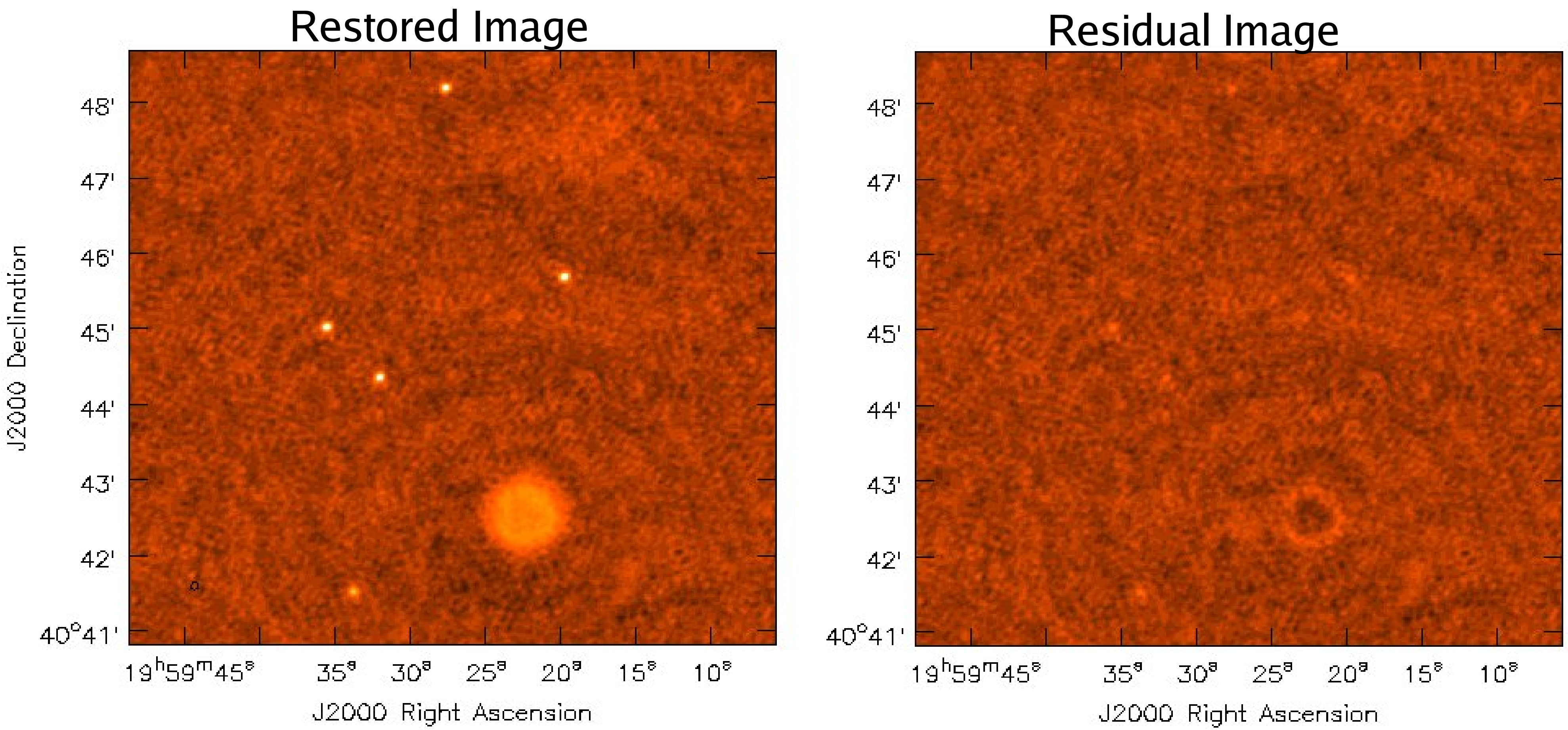
Residual Image



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

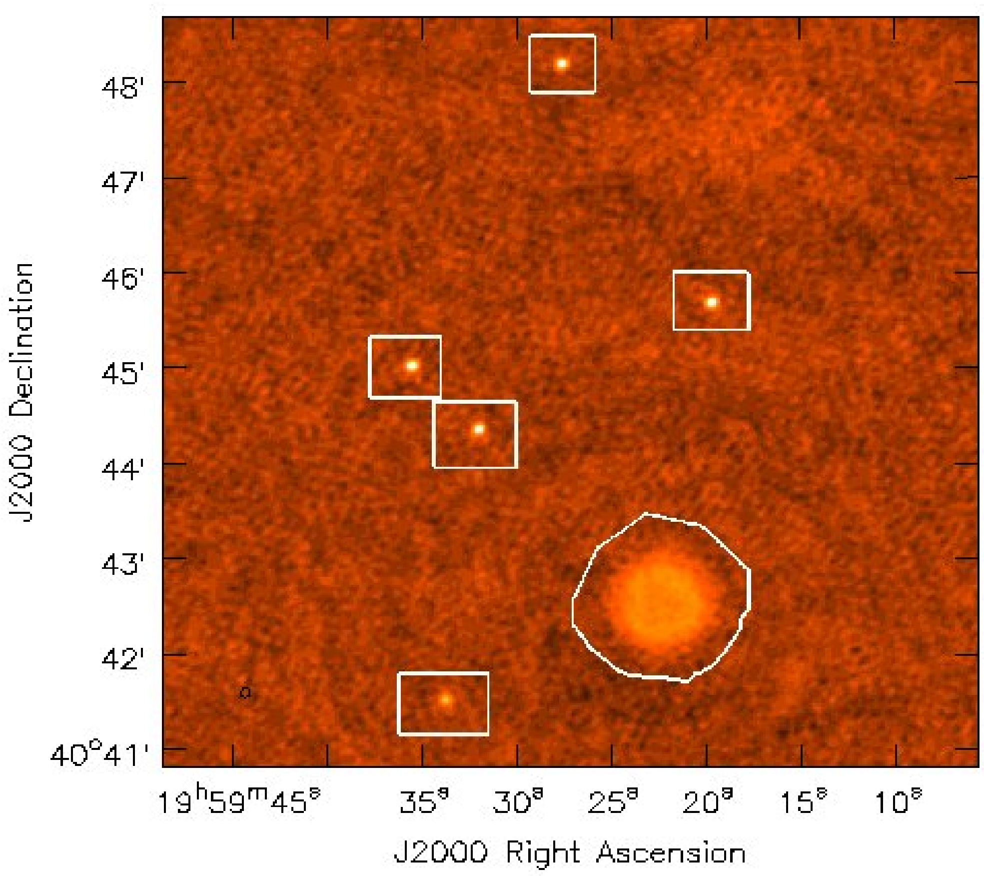


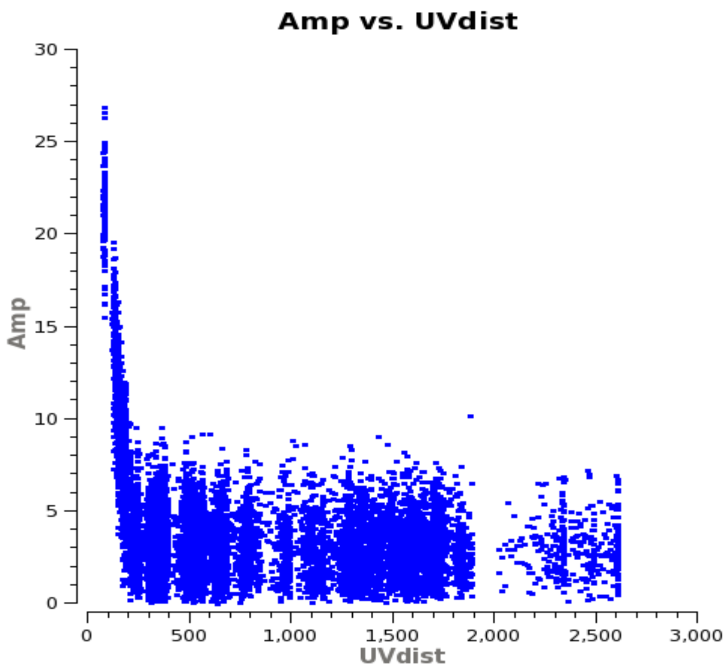
Peak artifacts : 0.07 Jy/beam.   Off-source RMS : 0.02 Jy/beam
Peak brightness : 1 Jy/beam   =>   Dynamic Range : 14 ~ 50

– 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.



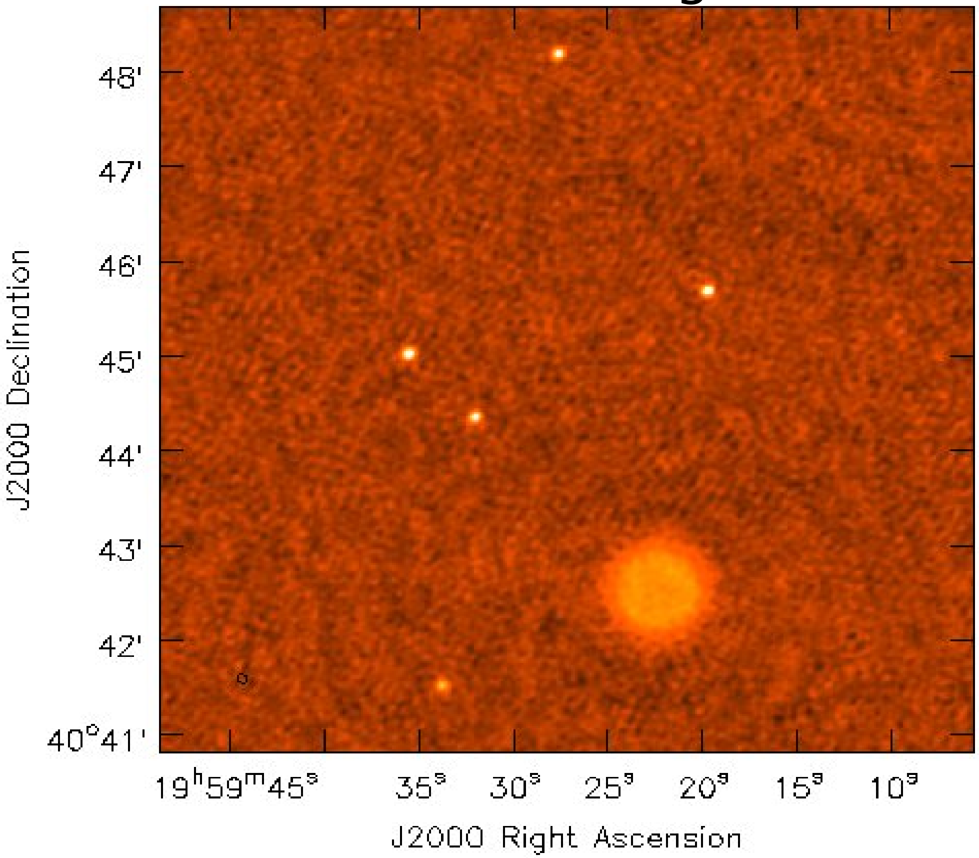


In general, point sources do not require boxes.
Extended emission needs it only if data constraints are insufficient.

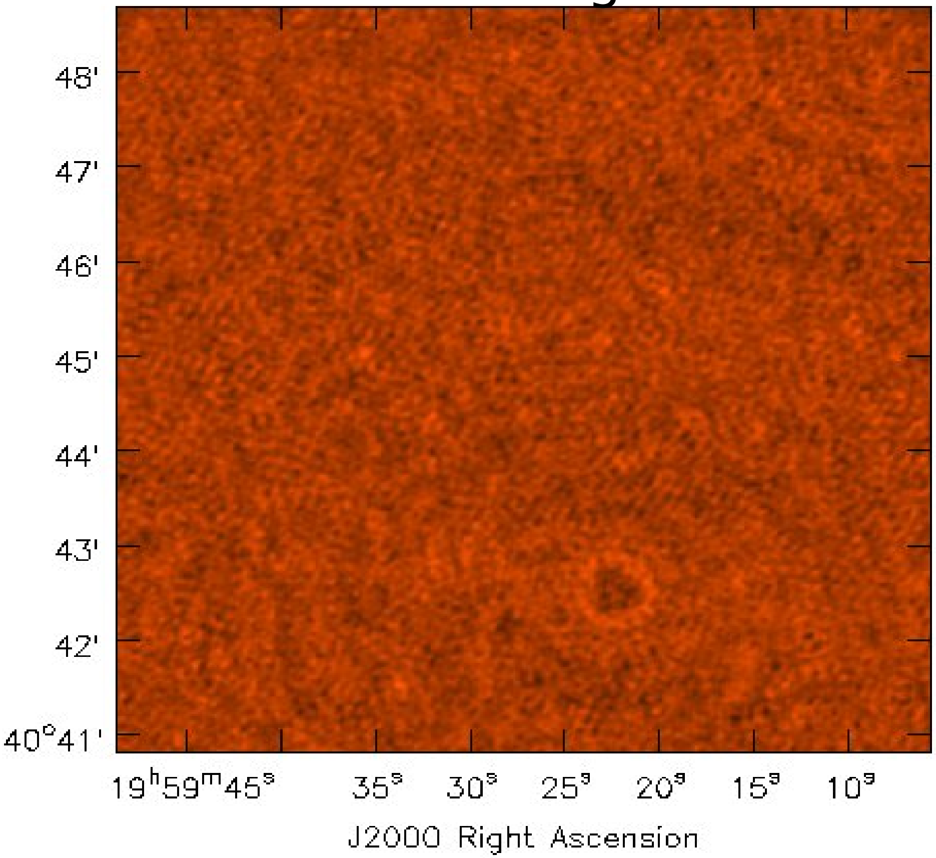
Example Imaging Problem – Third try...

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

References

Synthesis Imaging in Radio Astronomy II,
Astronomical Society of the Pacific conference series, Vol 180, 1999
Eds. Taylor, G.B. ; Carilli, C.L. ; Perley, R.A.

Online array simulator :   VRI, the Virtual Radio Interferometer
Type vri in searchbox on ATNF website
http://www.narrabri.atnf.csiro.au/astronomy/vri.html
Lets you experiment with Fourier transforms and ATCA configurations

CASA 'simdata' task : Simulate, image and deconvolve

CASS radio astronomy school lectures :
http://www.atnf.csiro.au/research/radio-school/2010/programme.html

NRAO synthesis Imaging school lectures :
http://www.aoc.nrao.edu/events/synthesis/2012/lectures.shtml

Acknowledgements

Mark Wieringa 2006 Lecture (ATNF)
Sanjay Bhatnagar 2008 Lecture (NRAO)
Emil Lenc 2010 Lecture (ATNF)
David Wilner 2012 Lecture (NRAO)