Imaging and Deconvolution

Joshua Marvil, NRAO

20th Synthesis Imaging Workshop

16 May 2024

New Mexico Tech
SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY
The University of New Mexico
NSF
NRAO
ALMA

Key Concepts

Visibility – a measurement of one Fourier component of the sky brightness distribution.

Gridding – combine all the measured visibilities onto a UV grid

Imaging – take the inverse Fourier transform of the gridded visibilities

Point spread function – an image artifact due to incomplete UV sampling

Deconvolution – the removal of the PSF from the image; ***'cleaning'***

Visibility ←→ Sky relationship

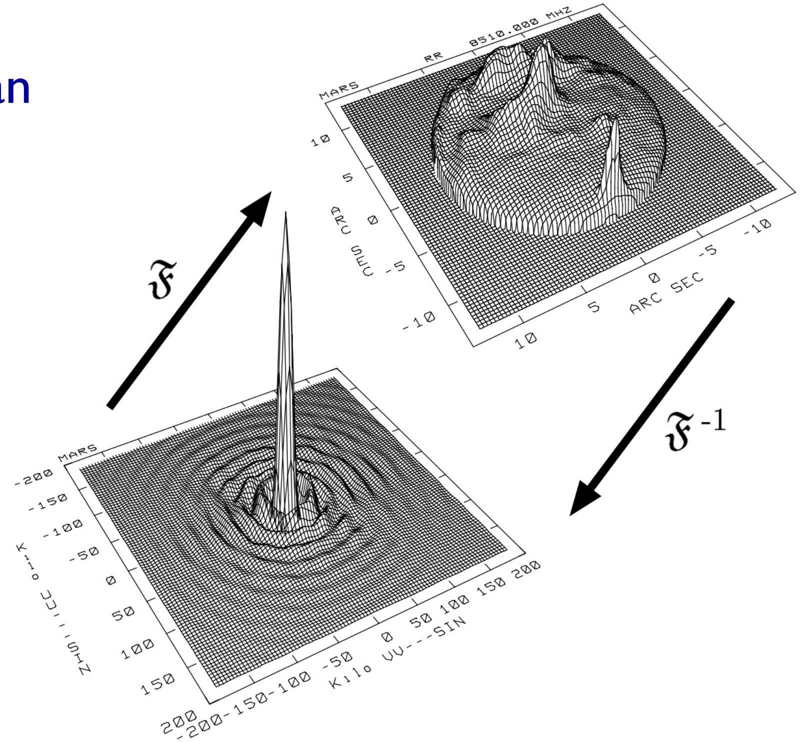
An interferometer directly measures components of the Fourier transform of the intensity distribution

After several simplifying assumptions, we can relate our measurements to the sky brightness:

$$V_\nu(u,v) = \iint I_\nu(l,m) e^{-2\pi i(ul+vm)}\, dl\, dm$$

which can be directly inverted as:

$$I_\nu(l,m) = \iint V_\nu(u,v) e^{2\pi i(ul+vm)}\, du\, dv$$

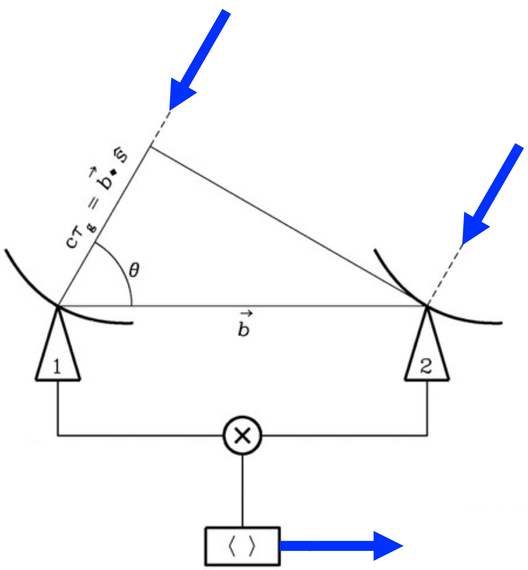


(Figure from Taylor, Carilli, & Perley)

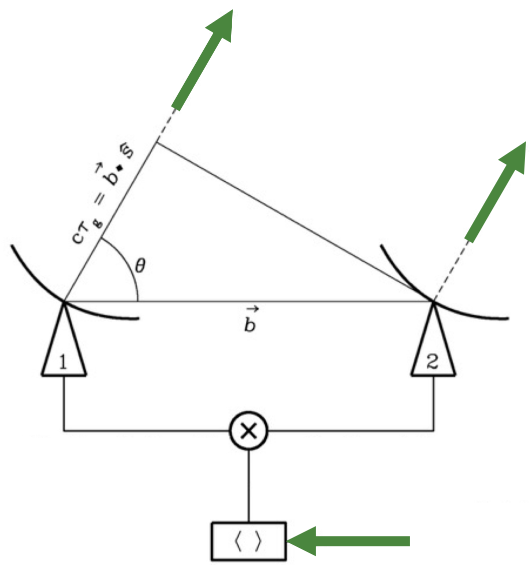
Visibility ←→ Sky relationship

The (weak) ***reciprocity*** theorem:

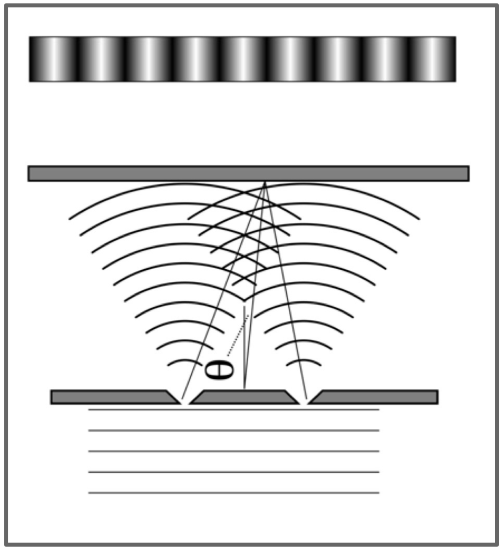
- Maxwell's equations are valid when time is reversed
- a system can be treated either as a receiver or a transmitter



receiver



transmitter

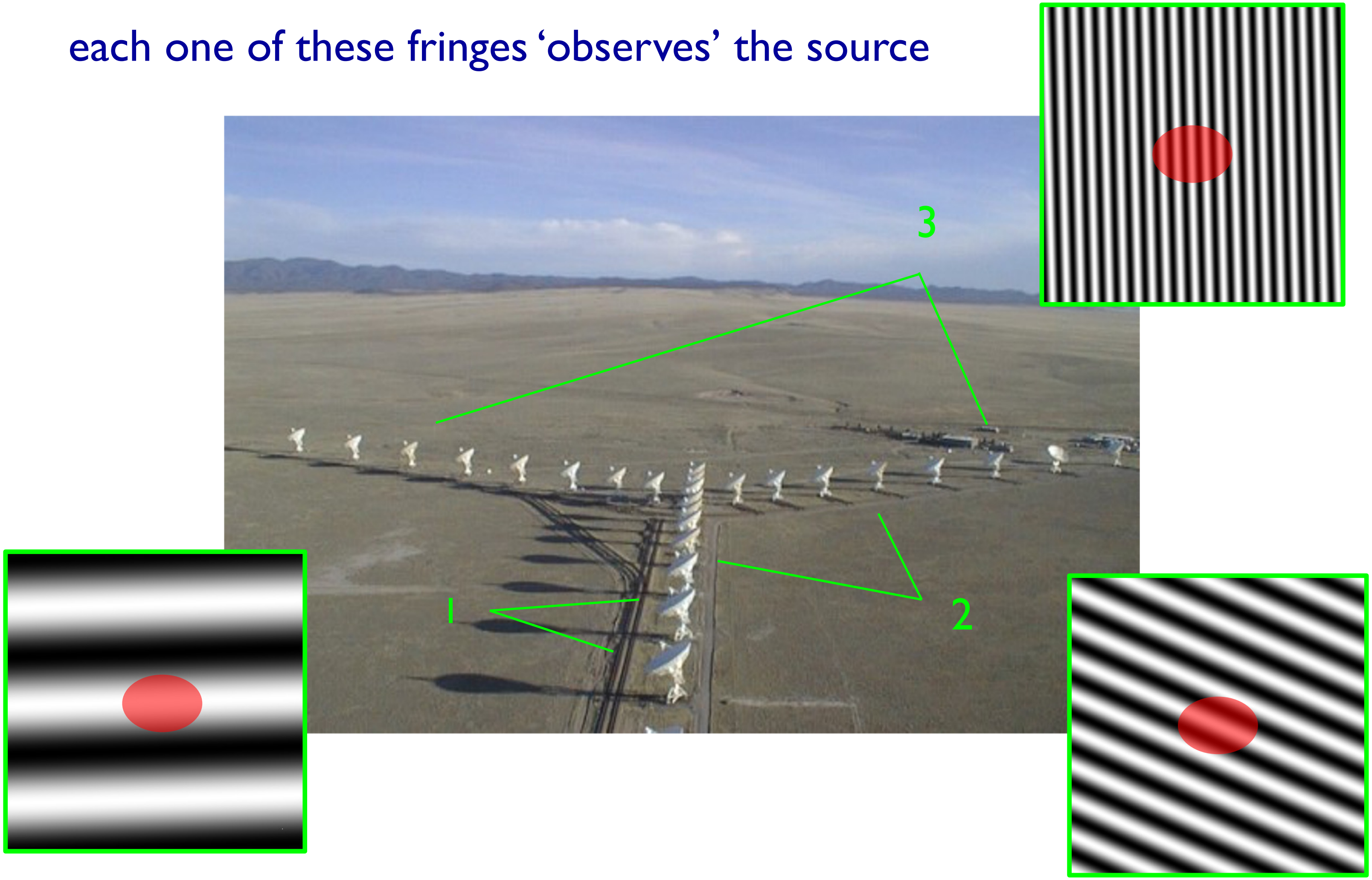


(Figure adapted from Condon & Ransom)

(Figure from wikipedia)

Visibility ←→ Sky relationship

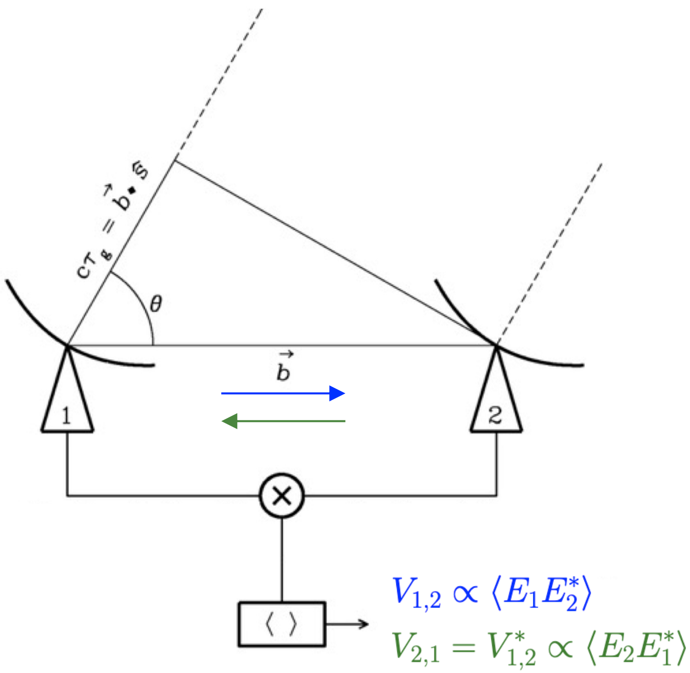
each one of these fringes ‘observes’ the source

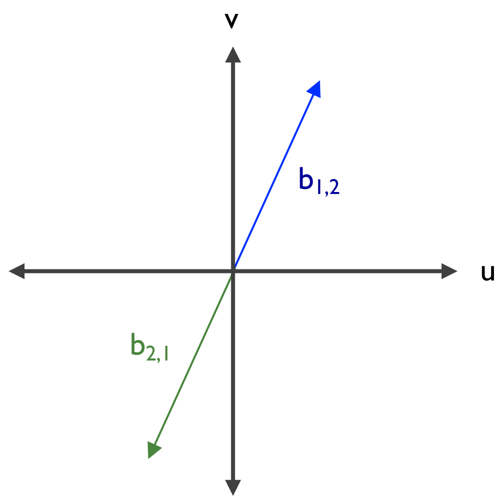


Visibility ←→ Sky relationship

If f(x) is real, F(s) is Hermitian: f(x) = f*(-x)

The sky is Real, so our Visibilities are Hermitian:
→ ***conjugate*** visibilities



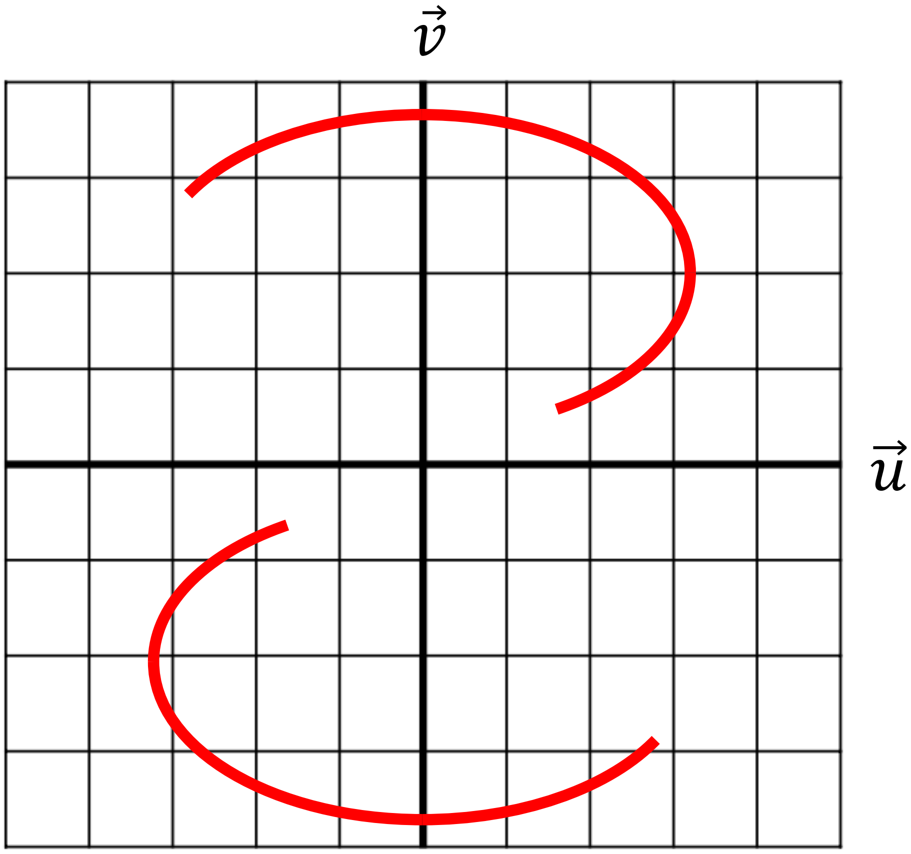


$b_{2,1}(u,v) = b_{1,2}(-u,-v)$

Visibility ←→ Sky relationship

Earth rotation synthesis: a baseline and its conjugate trace out elliptical arcs through the u-v plane over time, based on its Earth-centric coordinates X_λ, Y_λ, Z_λ, the source's declination δ, and the range of hour angles H observed

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \sin H & \cos H & 0 \\ -\sin\delta\cos H & \sin\delta\sin H & \cos\delta \\ \cos\delta\cos H & -\cos\delta\sin H & \sin\delta \end{bmatrix} \begin{bmatrix} X_\lambda \\ Y_\lambda \\ Z_\lambda \end{bmatrix}$$

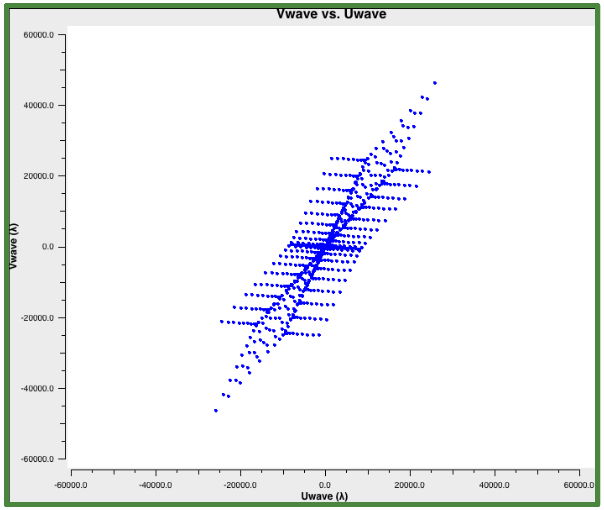


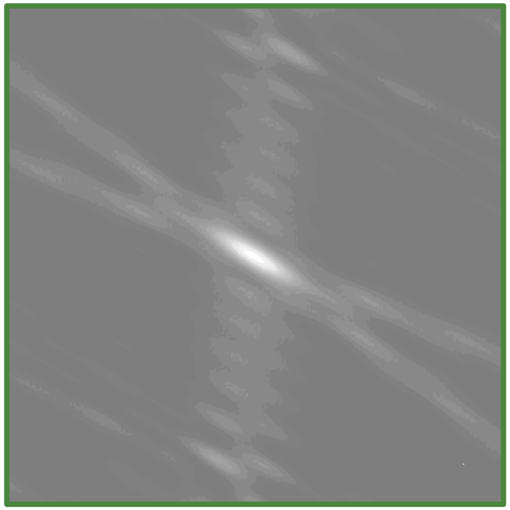
semi-major axis $\sqrt{X_\lambda^2 + Y_\lambda^2}$

semi-minor axis $\sin\delta_0\sqrt{X_\lambda^2 + Y_\lambda^2}$

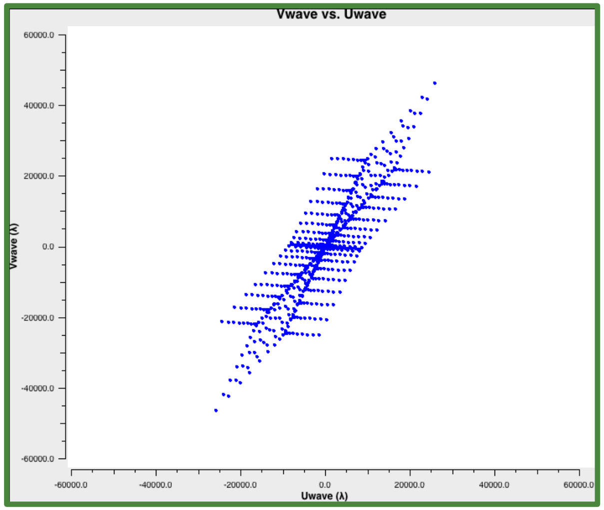
displacement along ***v*** $Z_\lambda \cos\delta_0$

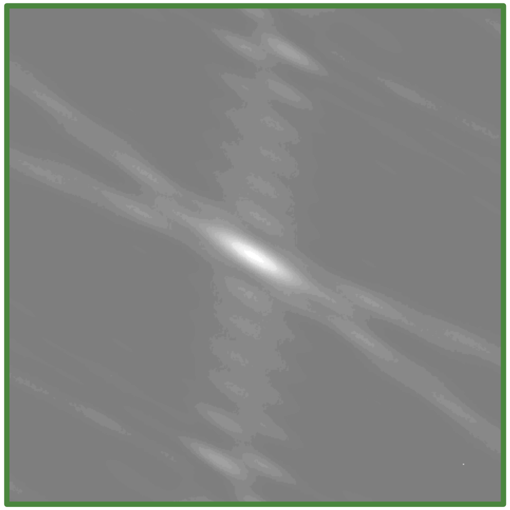
Visibility ←→ Sky relationship





Snapshot





Cumulative

Synthesized aperture

Synthesized beam

Table of visibilities

Essential columns:

- **b** in wavelengths or meters
- calibrated **V** in Janskys = 10^{-26} Wm^{-2}Hz^{-1}
- frequency, phase-tracking center

Other columns:

- data weight, flags
- timestamps, intervals
- antenna & receiver IDs

many, many rows

b (u, v, w)	V (x+iy)	direction	frequency	…

Ancillary tables: antenna pointing, weather, system calibration

Imaging with the Fourier transform

Performance considerations for M visibilities, NxN image size:

- Direct Fourier Transform: $O(M*N^2)$
- Discrete Fourier Transform: $O(N^4)$
- Fast Fourier Transform: $O(N^2 \log_2 N)$

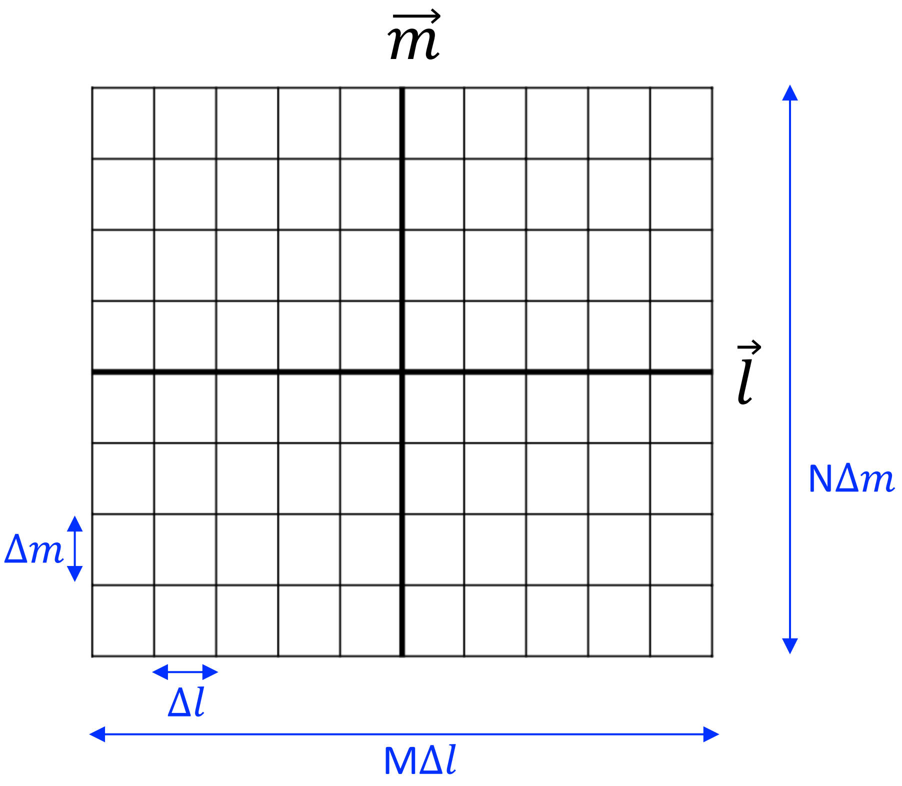
require regular grid

Also image plane gridding and other advanced solutions

Gridding

Practical to think first about the image grid:

- predict your angular resolution from diffraction: $\theta_{syn} \approx \frac{\lambda}{b_{max}}\ (radians)$
- choose an image cell size Δl that oversamples θ_{syn} by a factor of 4~5
- choose an image size MxN (pixels) based on your desired field of view (FOV)
- FOV will be M Δl x N Δm
- typically want Δl=Δm, M=N



Gridding

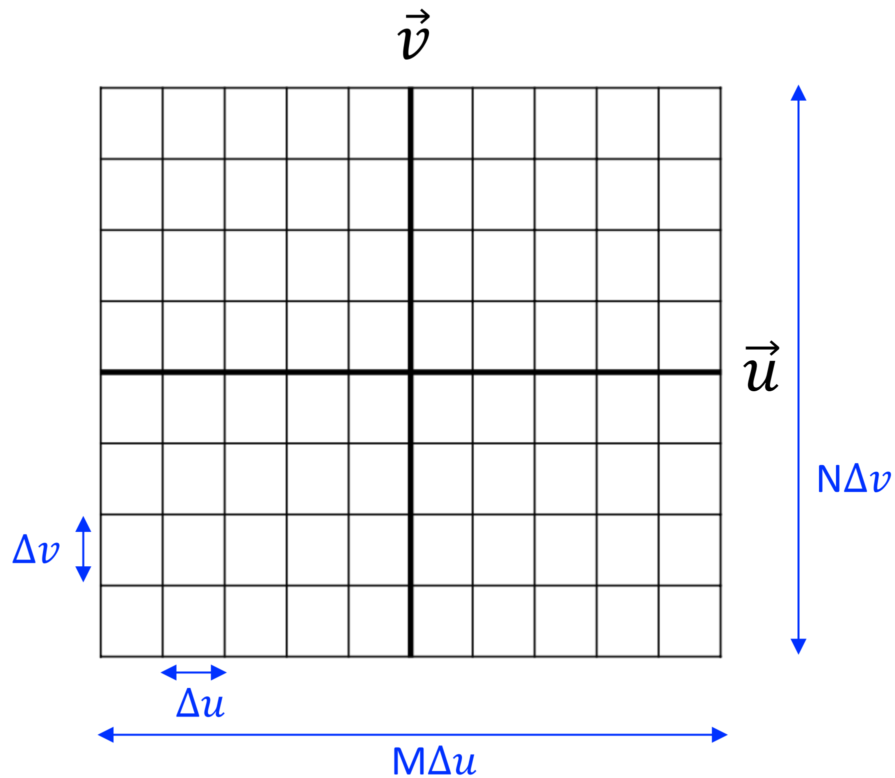
UV grid properties will be determined from the image grid

- the number of UV grid cells (pixels) will be the same as what was chosen for the image size, MxN

- the UV cell size will be inversely proportional to the image FOV:

  Δu=1/(M Δl)

  Δv =1/(N Δm)

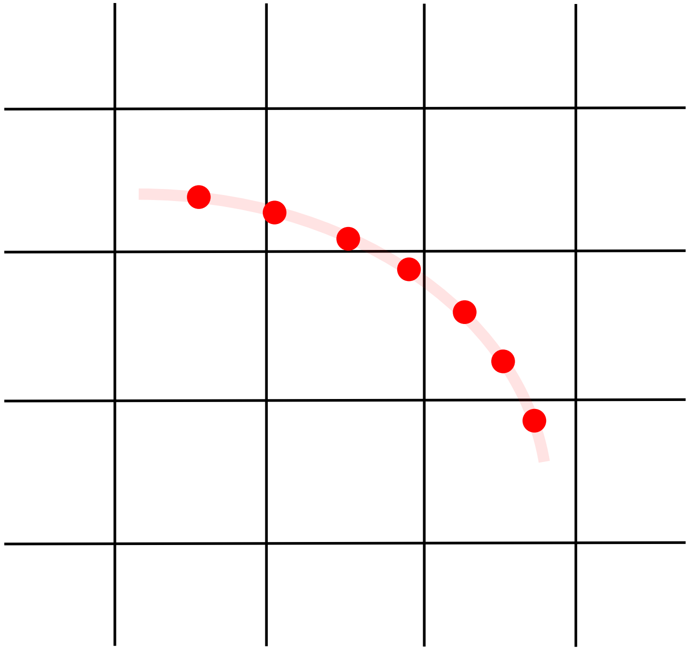


Gridding

Visibilities are averaged in the correlator and written out as discrete samples

The correlator integration times are often 1 -10 seconds

Measured visibilities will have values of **u** and **v** that do not align with grid cells, but the FFT requires a grid with regular spacing



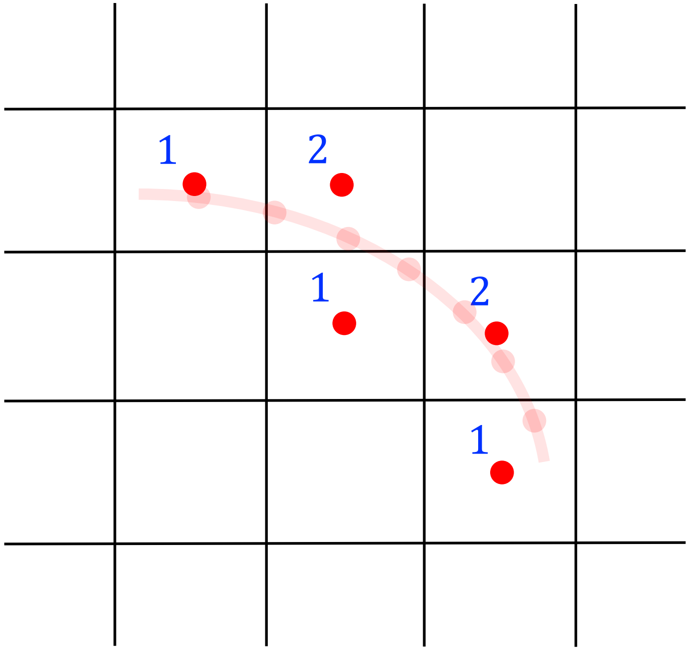
Gridding

The most basic gridding process is called ***nearest neighbor***

Move each visibility to the center of the closest UV cell

If there is more than one visibility per cell then take the weighted average, using the ***data weights***
→ ***cell averaging***

Keep track of the sum of weights in each UV cell using a separate grid
→ ***gridded weights***



Gridding

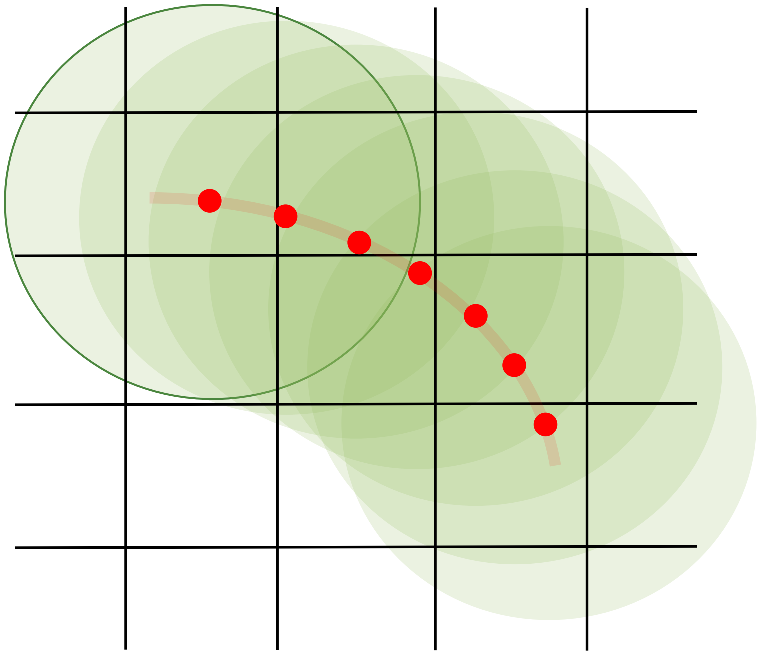
We can improve the gridding process through ***convolutional resampling***

Each visibility is convolved by a ***gridding kernel*** that distributes the visibility across multiple cells

We can oversample the kernel to reduce interpolation errors

Specific kernels can address other issues including:

- aliasing
- w-term
- mosaicking



The effect of the kernel can be divided out in the image plane

First Image of the Visibilities

The following equation requires that we know $V(u, v)$ at all values of u, v

$$I_\nu(l,m) = \iint V_\nu(u,v) e^{2\pi i(ul+vm)}\, du\, dv$$

but we know that our UV coverage is incomplete. So we add a new ***sampling function*** term $S(u, v)$

$$I_\nu^D(l,m) = \iint V_\nu(u,v) S(u,v) e^{2\pi i(ul+vm)}\, du\, dv$$

where

$$S(u,v) = \sum_{k=1}^{M} \delta(u - u_k, v - v_k)$$

M: # of visibility measurements

First Image of the Visibilities

from the multiplication property of the Fourier transform, we can equate

$$I_\nu^D(l,m) = \iint V_\nu(u,v)S(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

with

$$I_\nu^D(l,m) = \mathcal{F}^{-1}[V_\nu(u,v)] * \mathcal{F}^{-1}[S_\nu(u,v)]$$

$$= I_\nu(l,m) * B_\nu(l,m)$$

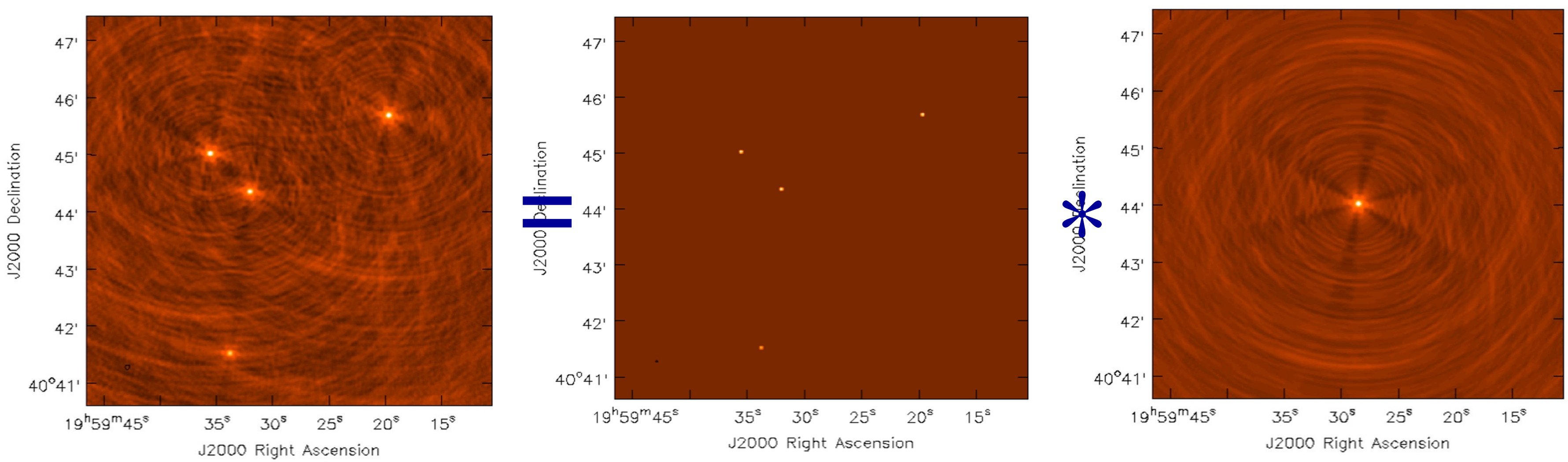
we call $I^D(l,m)$ the ***dirty image*** and $B(l,m)$ the ***dirty beam***

Reversing this convolution to recover $I(l,m)$ is called ***deconvolution***

First Image of the Visibilities

Example dirty image resulting from the convolution of the sky model with the dirty beam. The dirty beam is also called the ***point spread function.***

$$I_\nu^D(l,m) = I_\nu(l,m) * B(l,m)$$



(Figure by U. Rao)

The Point Spread Function (PSF)

The **PSF** is computed from the iFFT of the ***gridded weights.*** We define the '***clean beam***' as a Gaussian fit to the central peak of the PSF.

Properties of the PSF can be adjusted through manipulation of these weights, at the expense of image sensitivity:

natural weighting: no adjustments; optimal averaging

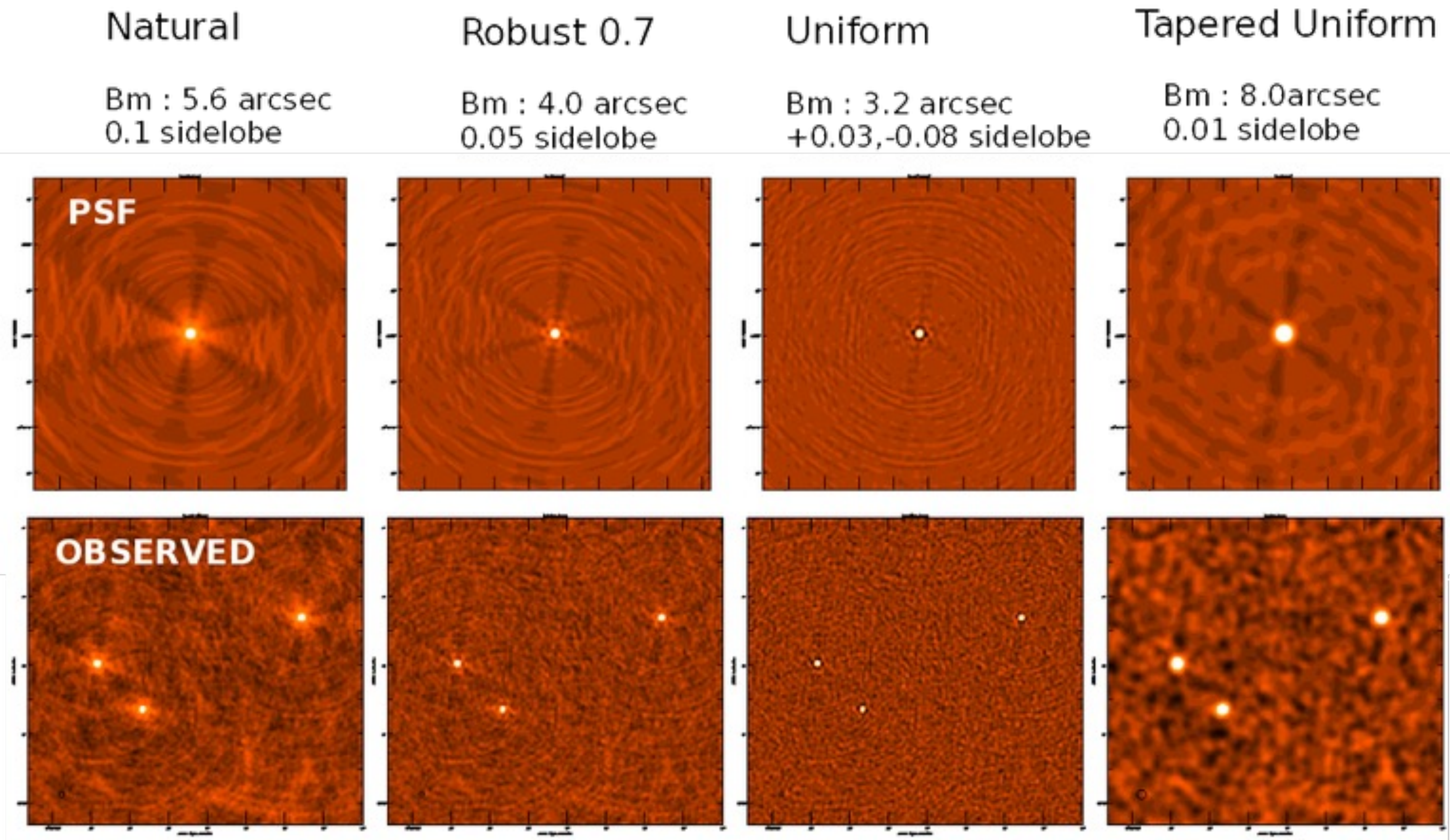
uniform weighting: constant weight per cell

super-uniform weighting: constant weight density

robust (Briggs) weighting: in between uniform and natural

(outer) UV taper: weight decreases with increasing UV-distance

The Point Spread Function (PSF)



(Figure by U. Rao)

Deconvolution

There is a formal, linear solution to the general problem of deconvolution, called the ***principal solution***

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

The principal solution is equivalent to the uniformly weighted dirty image

Linear deconvolution can not fill in the unmeasured UV cells

The FT of these zero-weight cells treats them as having zero value

Deconvolution

The FT of zero-weight cells treats them as having zero value, leading to spatial filtering

UV max due to the longest baseline

→ resolution limit, low-pass filter

UV min due to b_{min} = $D_{antenna}$

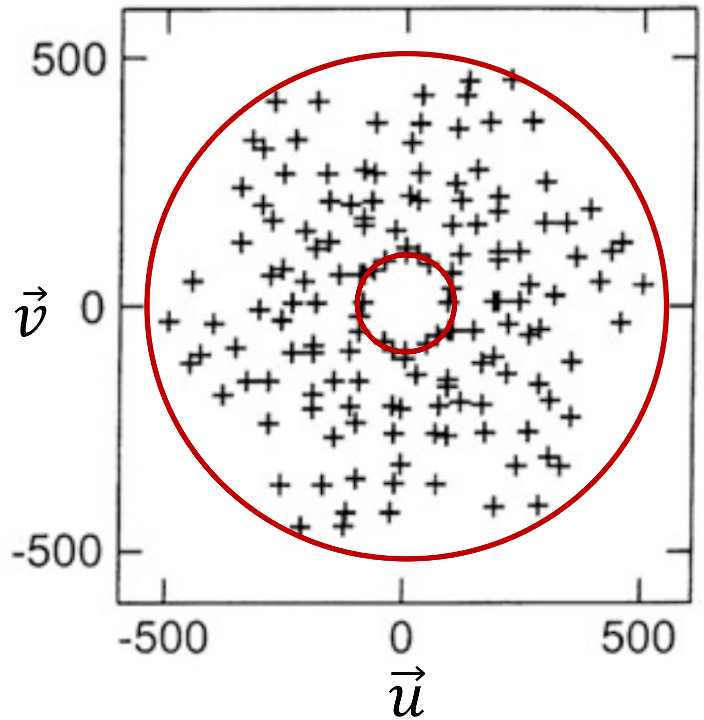
→ largest angular scale (LAS)

→ high-pass filter

Missing data in between samples:

spatial frequencies that are missing from the dirty image

→ PSF structure



(Figure adapted from Thompson, Moran, & Swenson)

Deconvolution

Deconvolution results can be improved with non-linear methods, e.g.

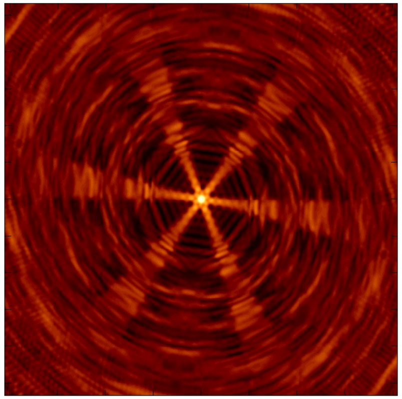
- Maximum Entropy
- Non-negative Least Squares
- Bayesian Inference
- CLEAN

These methods work by creating a model that is a good fit to the sampled visibilities. The FT of these models is not required to be zero at the location of unmeasured cells.

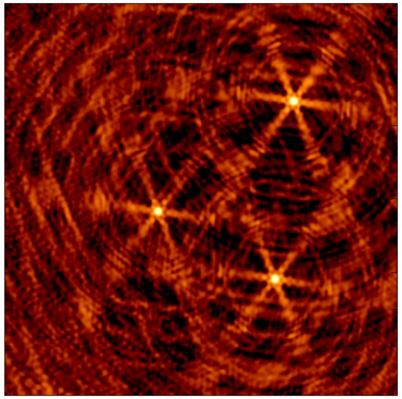
The final image is created from the ***model image*** without using any FTs. Therefore, only the residuals (data-model), which are hopefully noise-like with no remaining signal, are convolved by the dirty beam.

Deconvolution

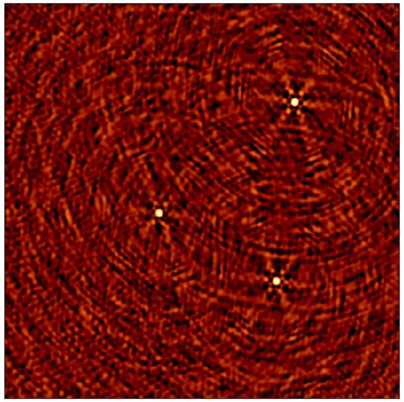
Simulated VLA example of 3 point sources



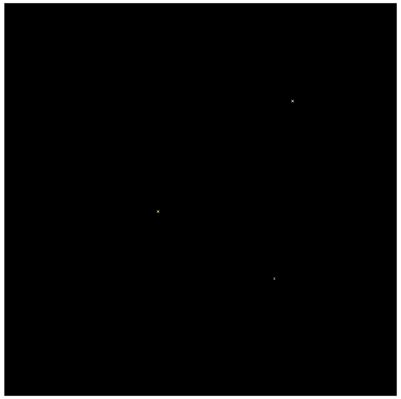
PSF



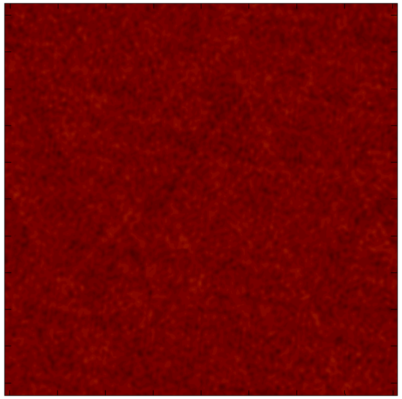
dirty image



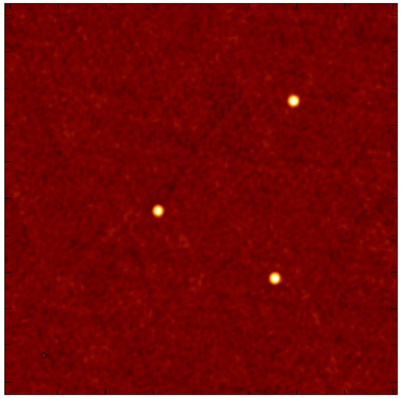
linear deconvolution



model



residual



restored

Deconvolution

The most basic CLEAN algorithm has the following phases:

Use the iFT to create the PSF and dirty image

Run the following loop until some stopping threshold is met:

- Find the brightest peak P in the dirty image
- Measure the position of the peak (x_p,y_p)
- Shift the PSF to (x_p,y_p), multiply the PSF by loop gain $\boldsymbol{\gamma}$ times P, and subtract this from the dirty image
- Add a point source with flux $\boldsymbol{\gamma}$P to the model at (x_p,y_p)

Smooth the model by a Gaussian ***clean beam*** having size ~ PSF

Add the ***residual image*** to the smoothed model

minor cycle

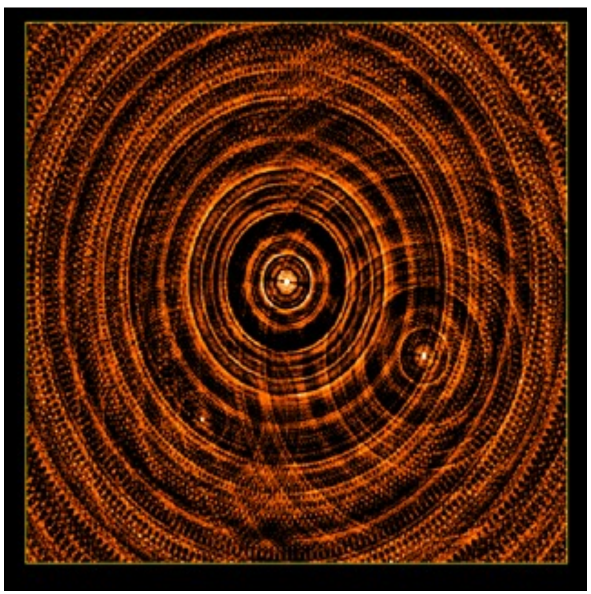
restoration

Deconvolution

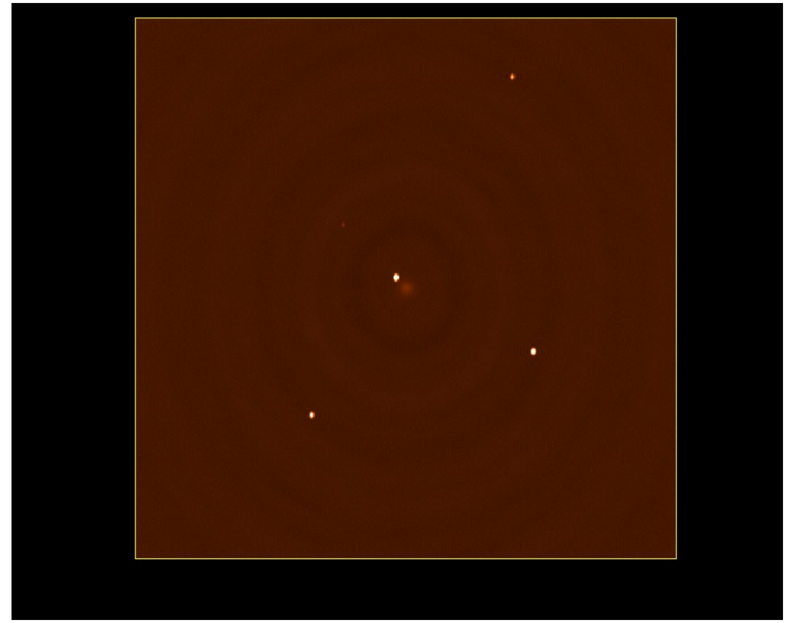
Simulated example from the Australian Telescope Compact Array

Model: 5 point sources and 1 Gaussian

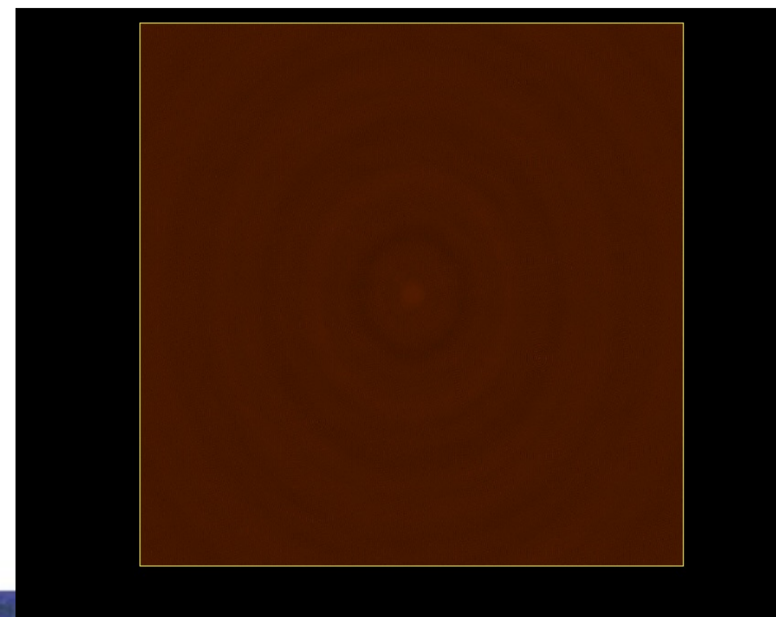
Clean iterations: 1,5,10,20,50,100,200,500, 1000



Dirty image



Restored



Residual

Deconvolution

Variations of the CLEAN algorithm:

Major cycles: gridding, FTs, UV-subtraction

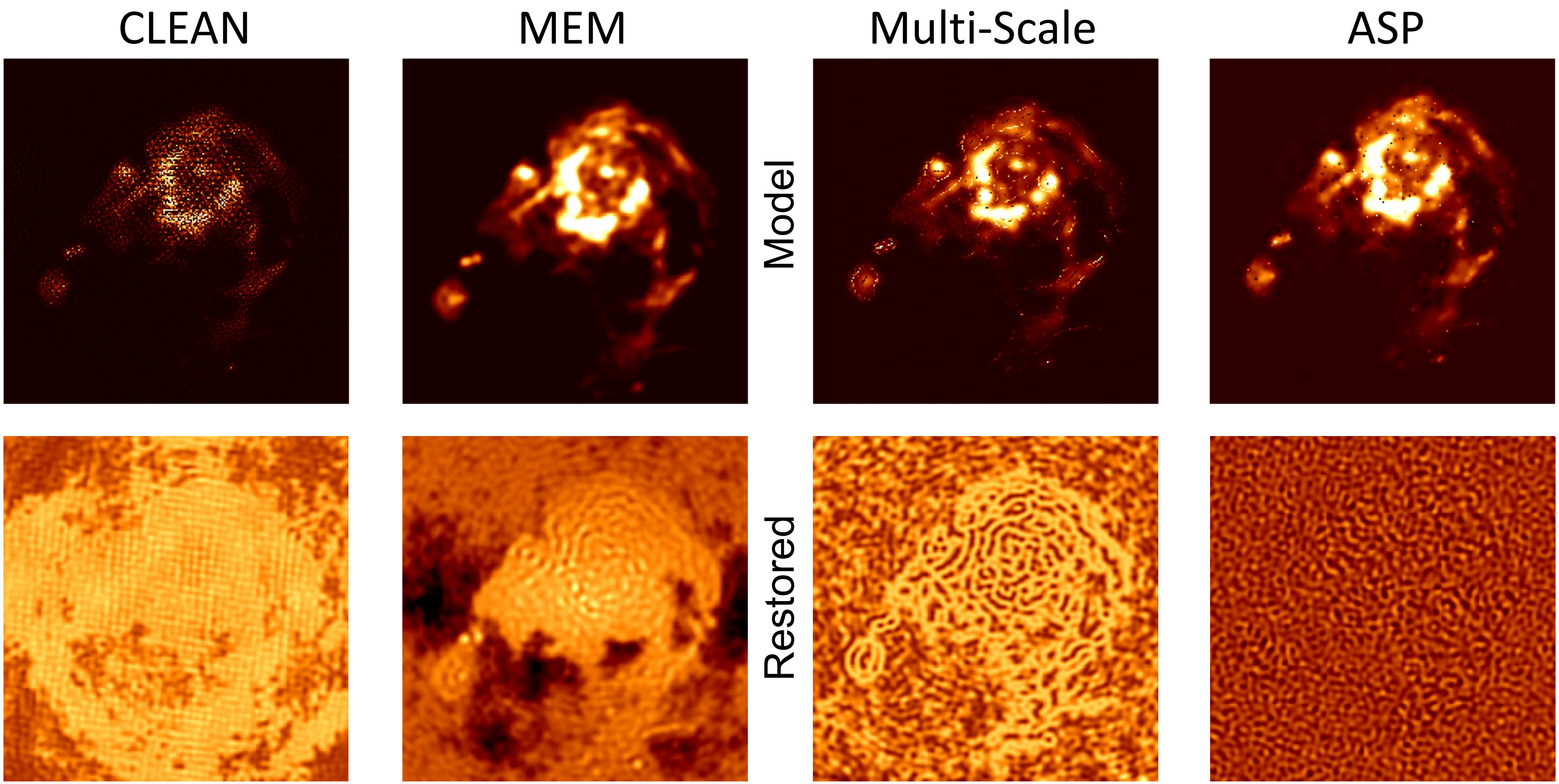
- Hogbom: only one cycle, large PSF
- Clark: many cycles, using the gridded data, PSF patch
- Cotton-Schwab: many cycles, using the measured data, PSF patch

Minor cycles: PSF subtraction, model building

- Hogbom: only point sources
- Multi-scale: point and a set of fixed paraboloids
- Adaptive Scale Pixel (ASP): point and a set of variable Gaussians

constrain the location of model components with ***masks***

Deconvolution



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

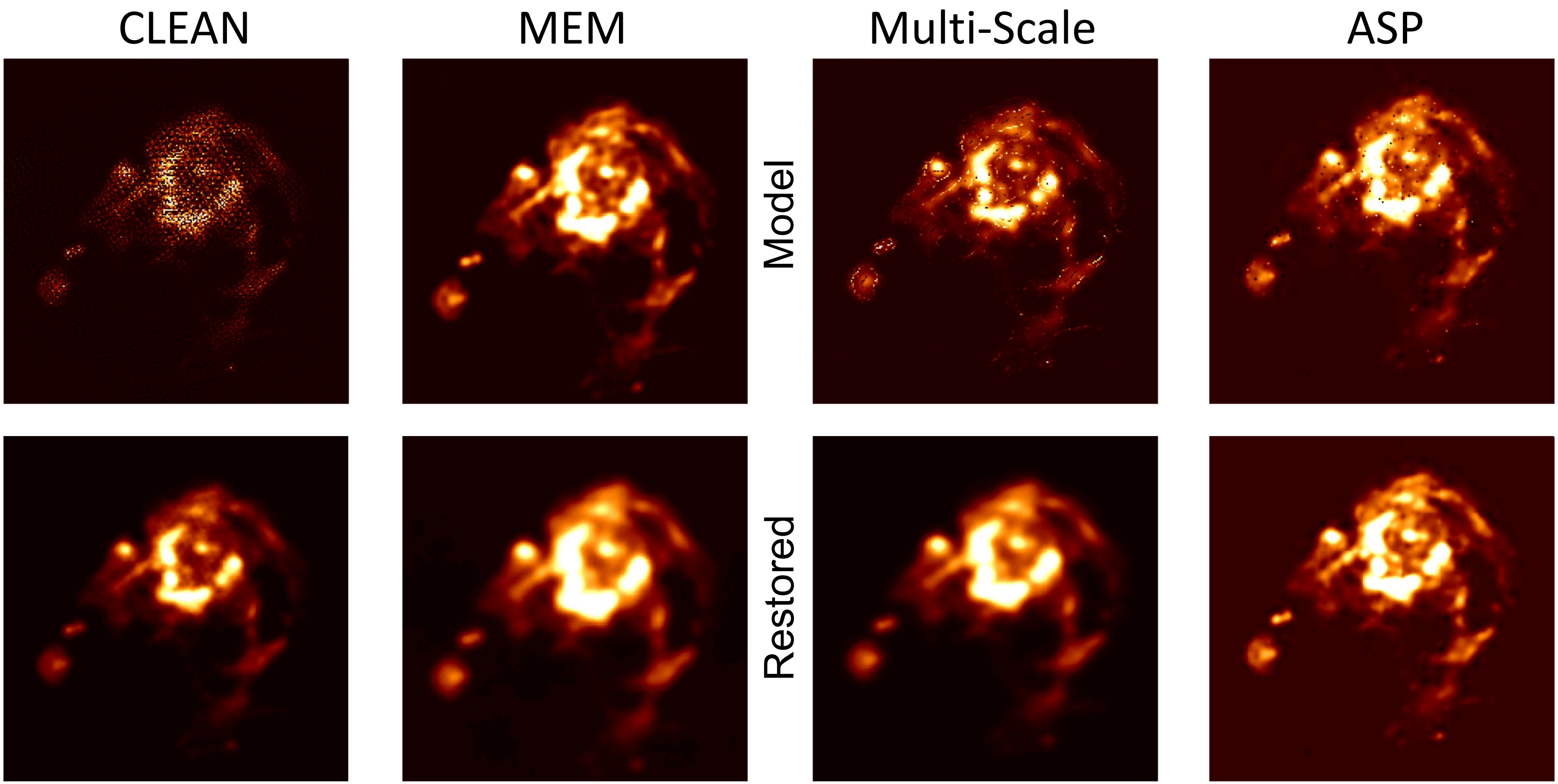
(Cornwell & Evans, 1985)

(Cornwell 2008)

(Bhatnagar & Cornwell 2004)

(figures from U. Rao)

Deconvolution



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

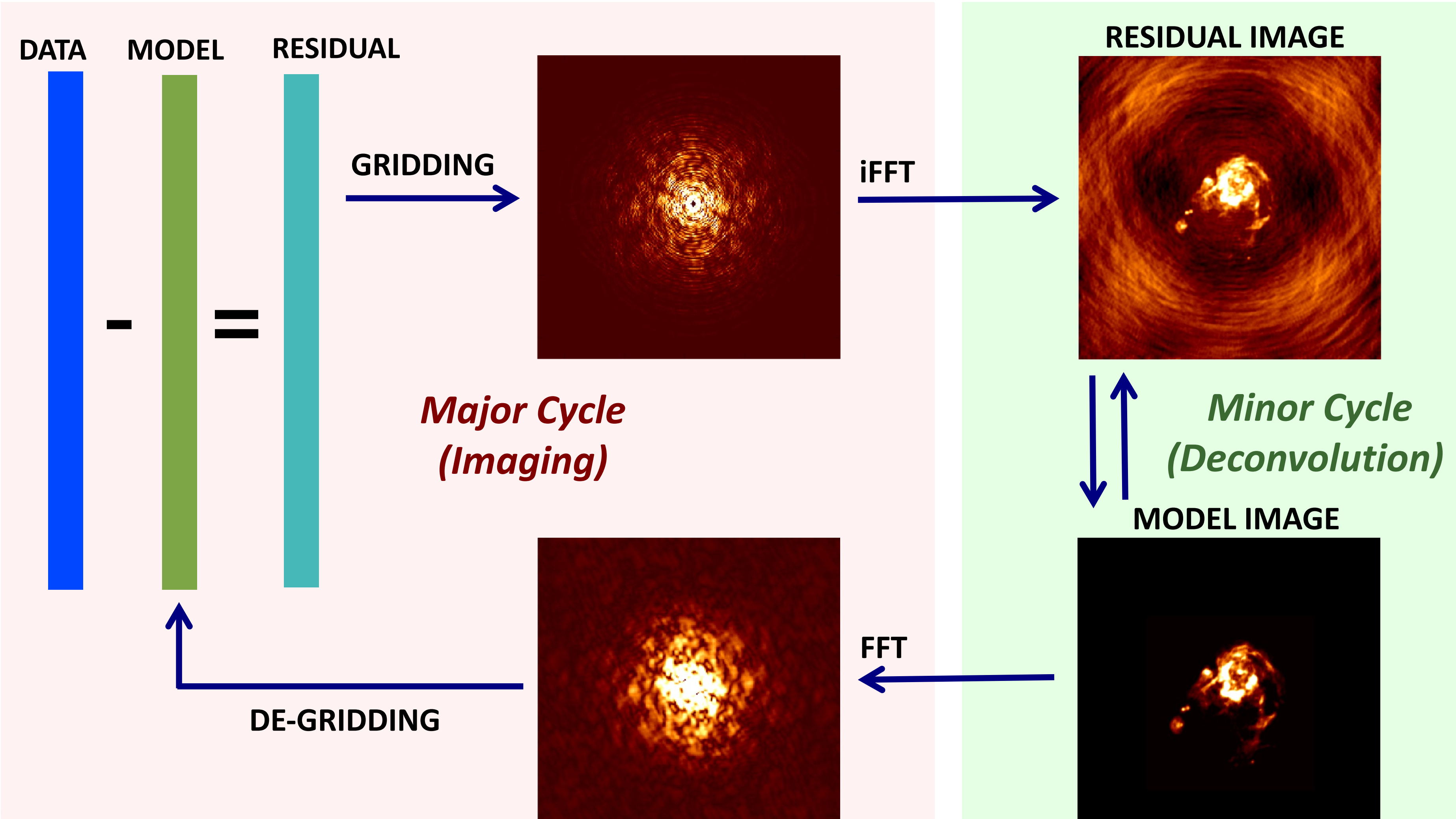
(Cornwell & Evans, 1985)

(Cornwell 2008)

(Bhatnagar & Cornwell 2004)

(figures from U. Rao)

Deconvolution



(figure adapted from U. Rao)

Summary

Visibility – a measurement of one Fourier component of the sky brightness distribution.

Dirty image – the inverse transform of the measured visibilities. Typically created by evaluating the visibilities on a regular grid, then using the Fast Fourier Transform.

Point spread function – what every point source looks like in the dirty image; also called the dirty beam. It is the FT of the gridded weights.

Deconvolution – replaces occurrences of the dirty beam in the dirty image by a Gaussian of equivalent size to make a '***cleaned***' image; this works by constructing a sky model in the image domain.

References & Credits

– Interferometry and Synthesis in Radio Astronomy
(Thompson, Moran, & Swenson)

– Synthesis Imaging in Radio Astronomy II
(Editors: Taylor, Carilli, & Perley)

– Essential Radio Astronomy
(Condon & Ransom)

– Tutorials, telescope observing guides, software documentation

– Imaging lectures from previous workshops

www.nrao.edu
science.nrao.edu
public.nrao.edu

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.