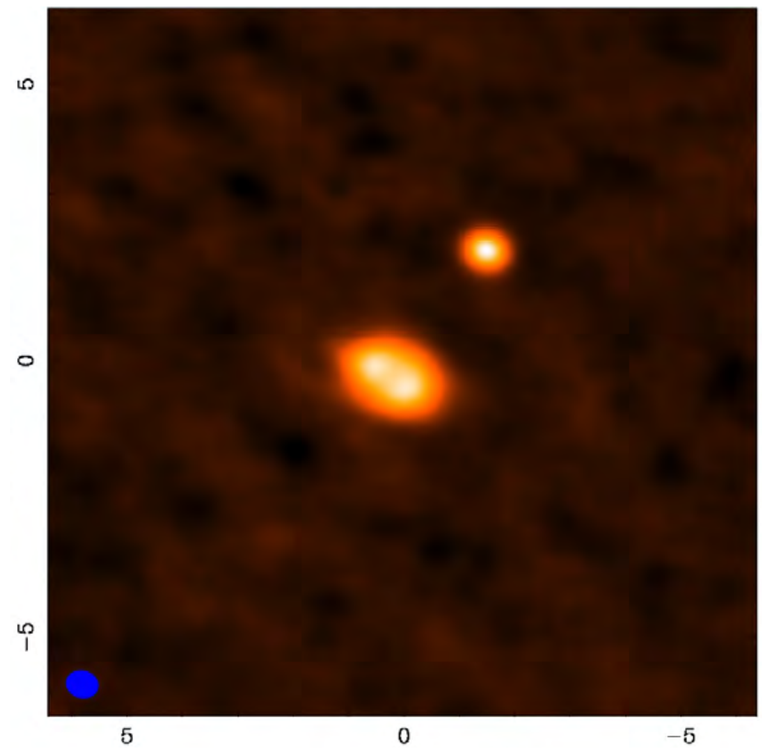
Imaging and Deconvolution

David J. Wilner

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN



Overview

goals

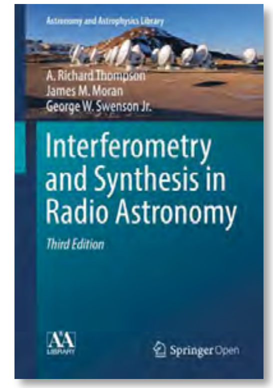
- gain intuition about interferometric imaging
- understand the need for deconvolution

topics

- Review of Fourier fundamentals
- Aperture synthesis and *(u,v)* plane sampling
- FFT and Gridding
- Visibility weighting schemes
- Deconvolution
- clean and maximum entropy algorithms
- Measures of image quality
- Spectral Line Considerations

References

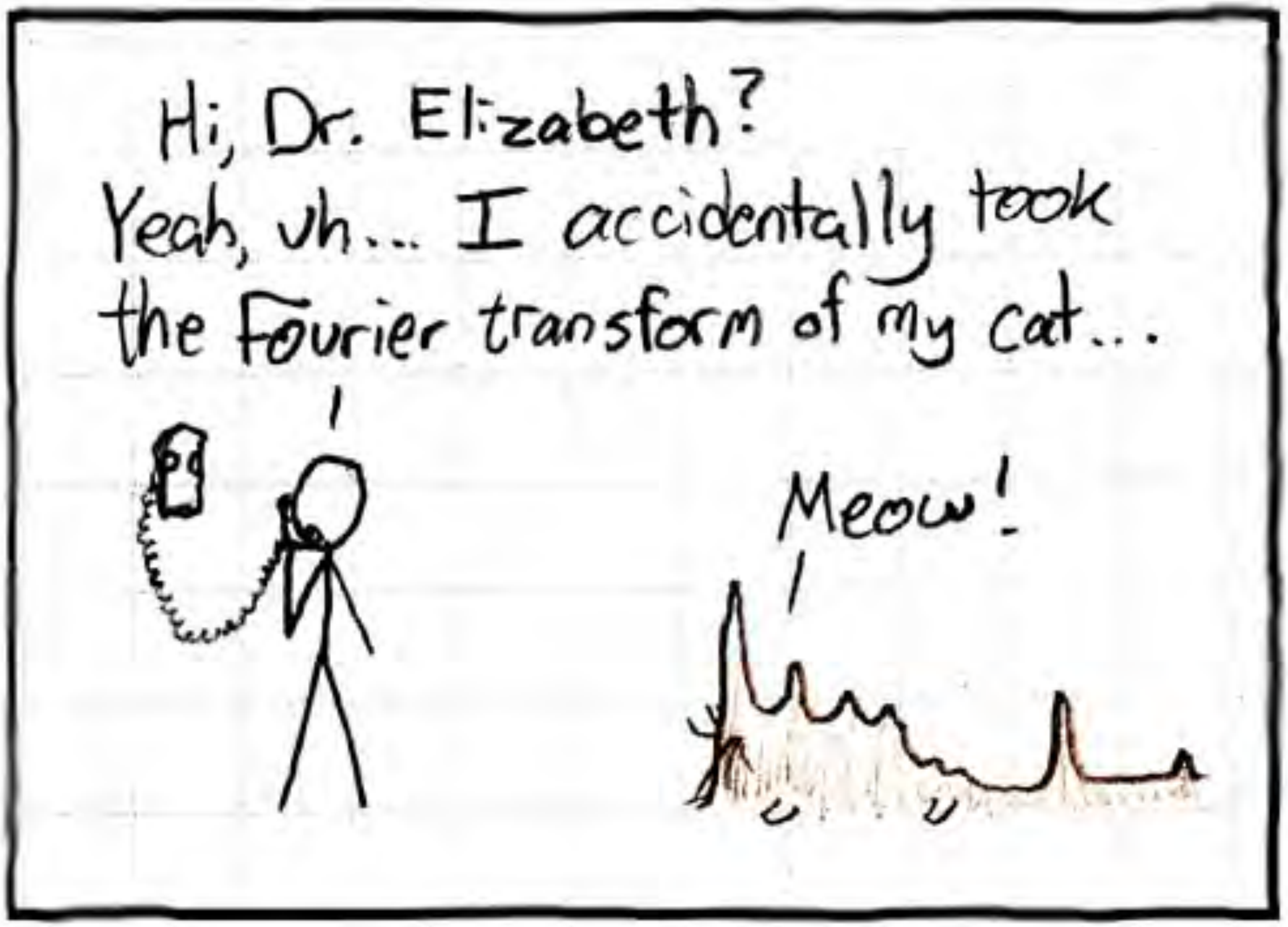
- Thompson, A.R., Moran, J.M. & Swensen, G.W. “Interferometry and Synthesis Imaging in Radio Astronomy” 3rd edition, 2017, Springer Open



- NRAO Synthesis Workshop proceedings
  - Synthesis Imaging in Radio Astronomy II,, ASP Conference Series, Vol. 180, 1999, eds. Taylor, G.B., Carilli, C.L., Perley, R.A.
  - lecture slides: www.aoc.nrao.edu/events/synthesis

- IRAM 2000 Interferometry School proceedings
  - www.iram.fr/IRAMFR/IS/IS2008/archive.html
- Condon, J.J. & Ransom, S.M. 2016 “Essential Radio Astronomy”
  - science.nrao.edu/opportunities/courses/era
- Many useful pedagogical presentations available on-line
  - ALMA primer, casadocs, CASS radio school, ERIS lectures, …
- EHT M87 imaging paper, 2019, ApJ Letters, 875, L4

xkcd.com/26/

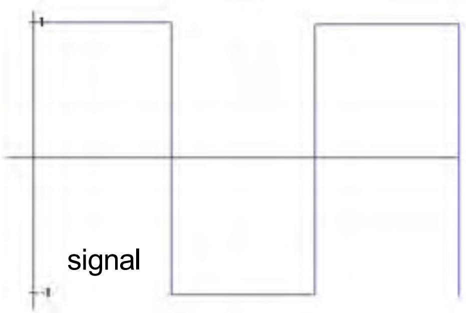


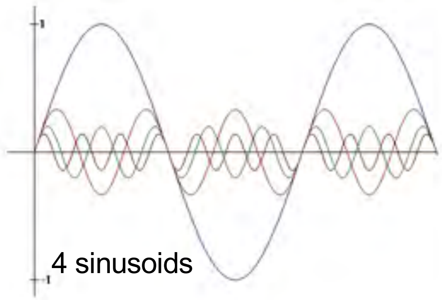
CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

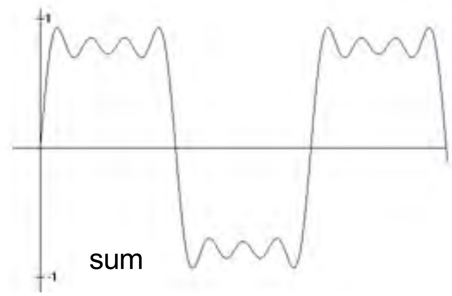
Fourier Fundamentals

Jean Baptiste Joseph Fourier 1768-1830

- Fourier theory states that any well behaved signal can be represented as the sum of sinusoids







$$x(t) = \frac{4}{\pi}\left(\sin(2\pi f t) + \frac{1}{3}\sin(6\pi f t) + \frac{1}{5}\sin(10\pi f t) + \cdots\right)$$

- the Fourier transform is the mathematical tool that decomposes a signal into its sinusoidal components
- the Fourier transform contains *all* information of the original signal

The Fourier Domain

- acquire comfort with the Fourier domain…
  - in older texts, functions and their Fourier transforms occupy *upper* and *lower* domains, as if "functions circulated at ground level and their transforms in the underworld" (Bracewell 1965)

- a few properties of the Fourier transform $g(x) \xrightarrow{\mathcal{F}} G(s)$

  An addition in one domain is an addition in the other

$$g(x) + h(x) = G(s) + H(s)$$

  A multiplication in one domain is a convolution in the other

$$g(x) = h(x) * k(x) \qquad G(s) = H(s)K(s)$$

  Scaling: large in one domain is small in the other

$$g(\alpha x) = \alpha^{-1} G(s/\alpha)$$

  An offset in one domain is a phase shift in the other

$$g(x - x_0) = G(s)e^{i2\pi x_0 s}$$

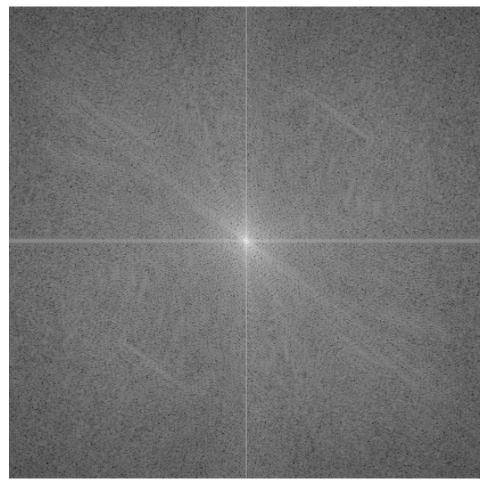
Visibilities

- each *V(u,v)* is a complex quantity
  - expressed as *(real, imaginary)* or *(amplitude, phase)*

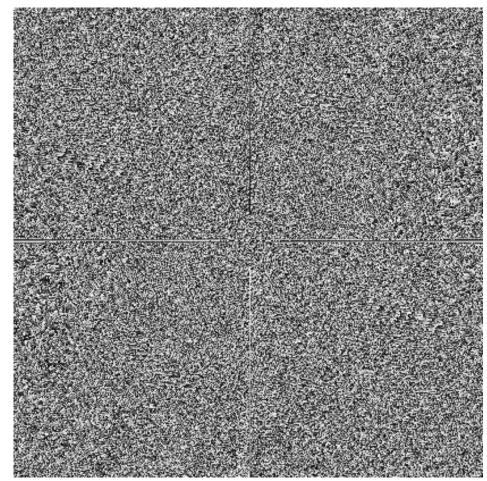
T(l,m)

$\xrightarrow{\mathcal{F}}$

V(u,v) amplitude



V(u,v) phase

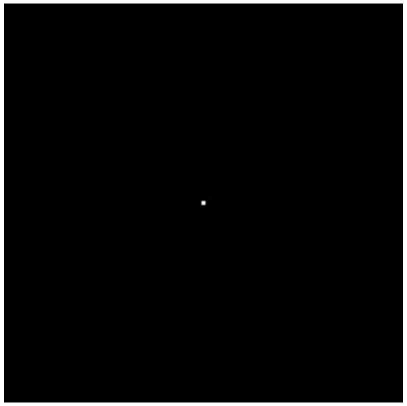


each V(u,v) contains information everywhere in the image T(l,m)

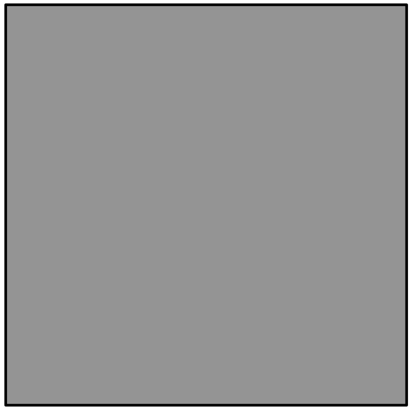
Some 2D Fourier Transform Pairs

T(l,m) | *V(u,v) amplitude*

δ function

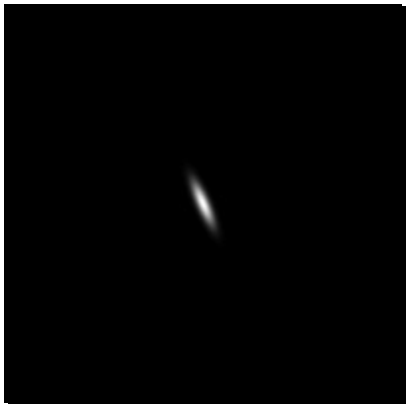


$\xrightarrow{\mathcal{F}}$

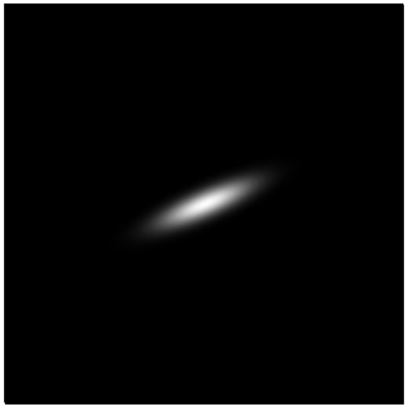


constant

elliptical Gaussian



$\xrightarrow{\mathcal{F}}$



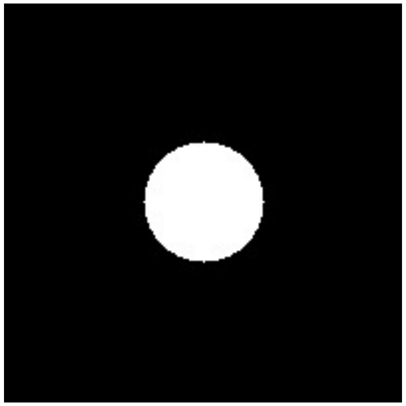
elliptical Gaussian

narrow features transform into wide features (and vice-versa)

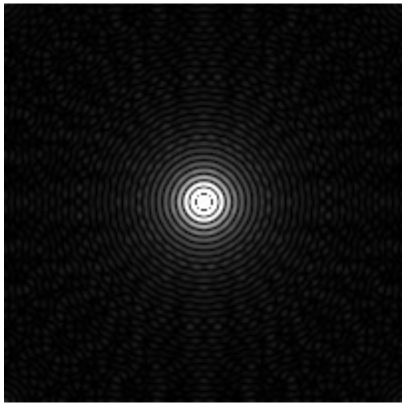
Some 2D Fourier Transform Pairs

T(l,m) *V(u,v) amplitude*

uniform disk



$\xrightarrow{\mathcal{F}}$



Bessel function

anything sharp in one domain generally oscillates ("rings") in the other

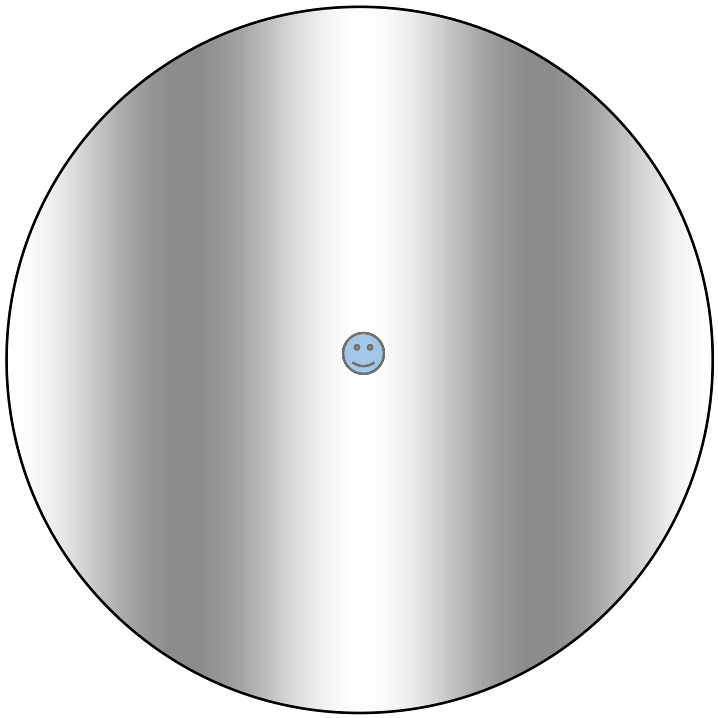
Response of an Interferometer

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

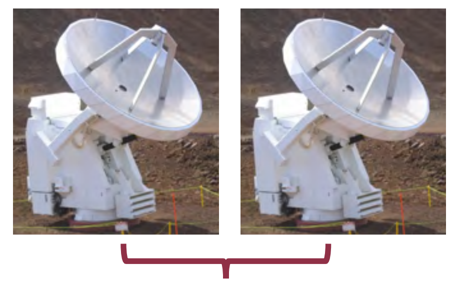
V(u,v), the complex visibility function, is the 2D Fourier transform of T(l,m) the sky brightness distribution (for an incoherent source, small field of view, far field, quasi-monochromatic, etc.)

- short baseline: wide fringe pattern, low resolution
  - flux from extended sources adds up
- long baseline: narrow fringe pattern, high resolution
  - flux from extended source averages out → low *V(u,v)*
- *V(u=0,v=0)* is the integral of *T(l,m)dldm* = total flux density
- *T(l,m)* is real: *V(-u,-v) = V*(u,v)*   *where * = complex conjugate*
  - get two visibilities for one measurement

$$V(u,v) = \iint T(l,m)e^{-i2\pi(ul+vm)}\mathbf{dldm}$$



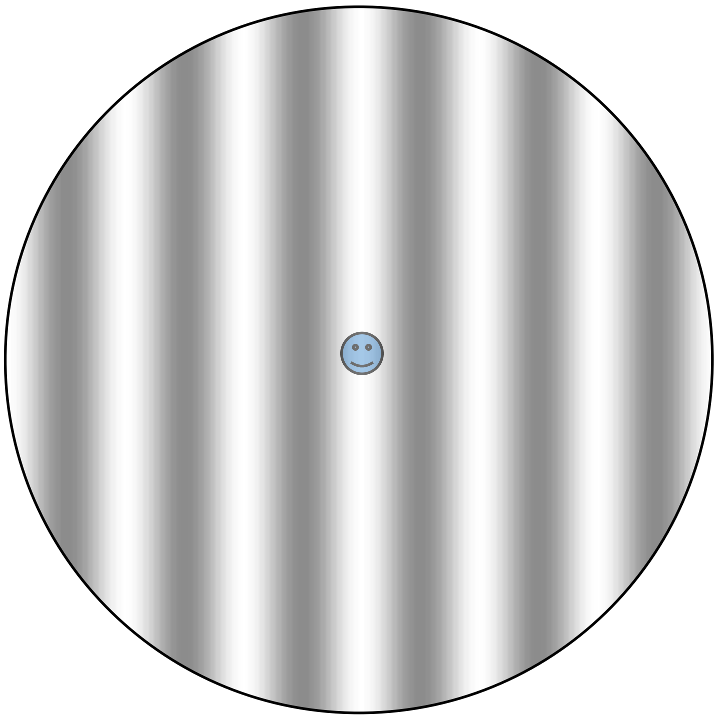
compact source



short baseline

wide fringe pattern

$$V(u,v) = \iint T(l,m)e^{-i2\pi(ul+vm)}\mathbf{dldm}$$

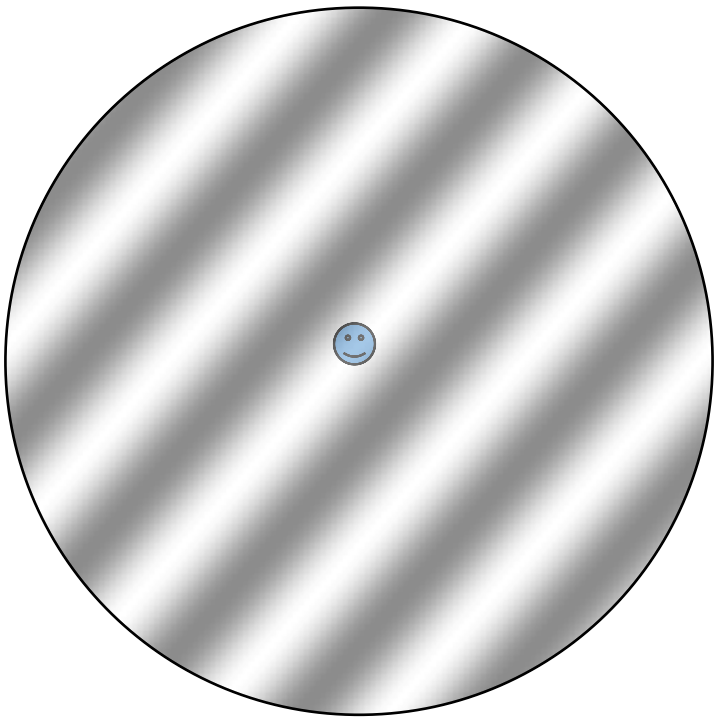


compact source

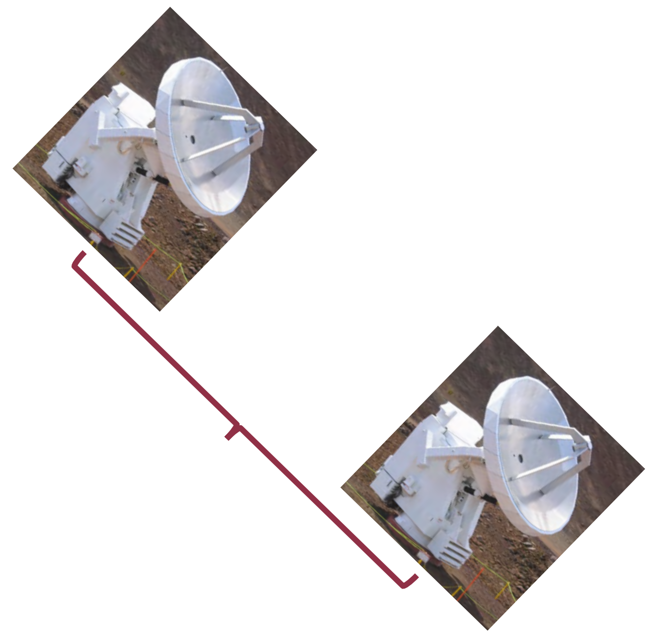
long baseline

narrow fringe pattern

$$V(u,v) = \iint T(l,m)e^{-i2\pi(ul+vm)}\mathrm{dldm}$$



compact source

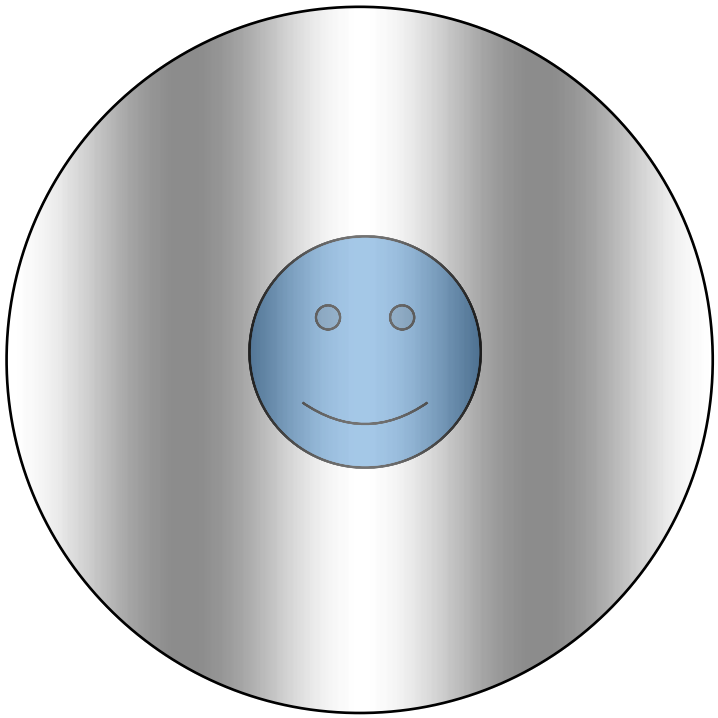


long baseline

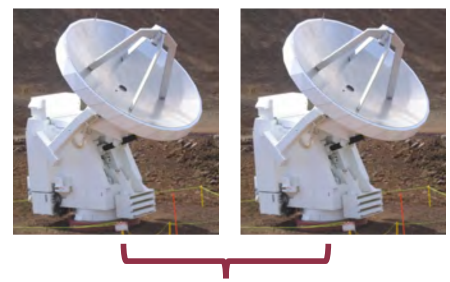
narrow fringe pattern

different orientation

$$V(u,v) = \iint T(l,m) e^{-i2\pi(ul+vm)} \mathrm{d}l\mathrm{d}m$$



extended source



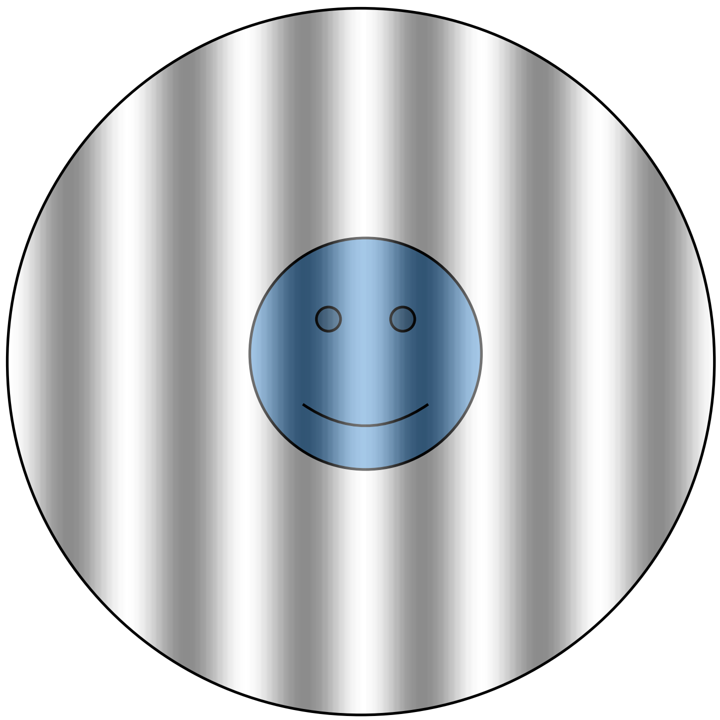
short baseline

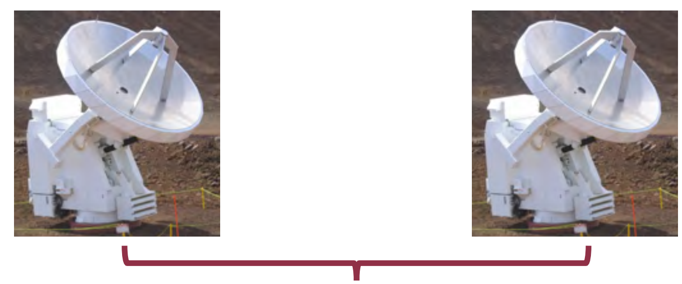
wide fringe pattern

CENTER FOR ASTROPHYSICS

HARVARD & SMITHSONIAN

$$V(u,v) = \iint T(l,m)e^{-i2\pi(ul+vm)}\mathbf{dldm}$$





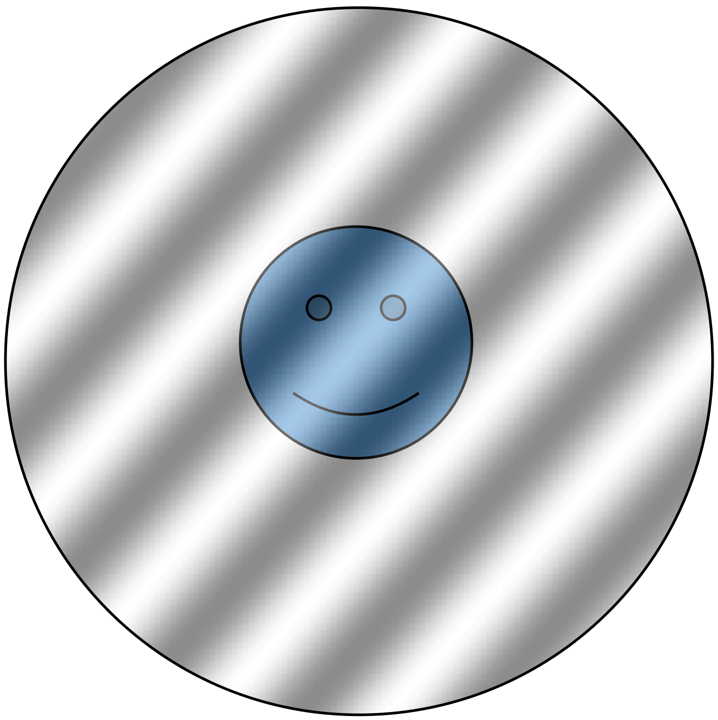
long baseline

narrow fringe pattern

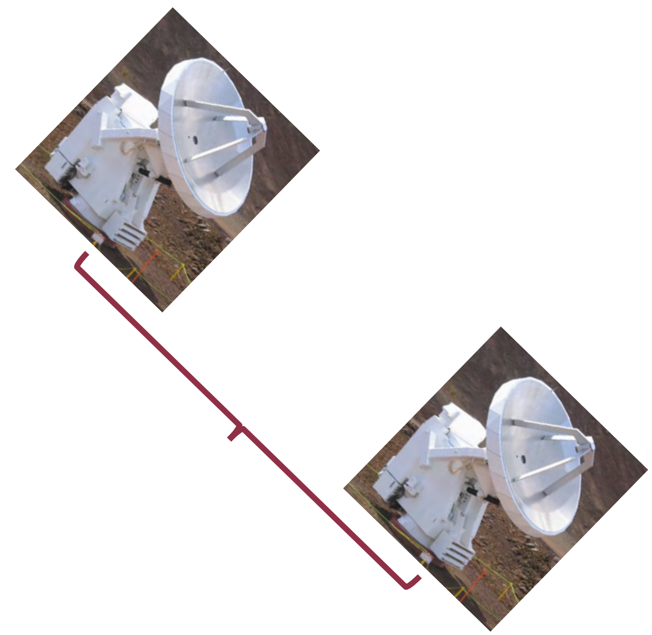
extended source

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

$$V(u,v) = \iint T(l,m)e^{-i2\pi(ul+vm)}\,\mathbf{dldm}$$



extended source



long baseline

narrow fringe pattern

different orientation

Aperture Synthesis

basic idea: sample $V(u,v)$ at enough (u,v) points using distributed small apertures to synthesize a large aperture of size (u_{max},v_{max})

use **more antennas** for more samples

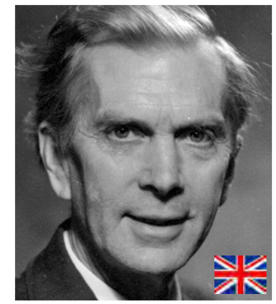
- one pair of antennas = one baseline
  = two (u,v) samples at one time
- N antennas = N(N-1) samples at one time
- reconfigure physical layout of N antennas for more (as long as source structure doesn't change with time)

use **Earth rotation** for more samples

- baseline length/orientation relative to sky change with time

use **more wavelengths** for more samples

- baselines stretch with increasing wavelength
- "multi-frequency synthesis" for continuum imaging: determine structure at some fiducial wavelength and the change with wavelength, e.g. Taylor expansion



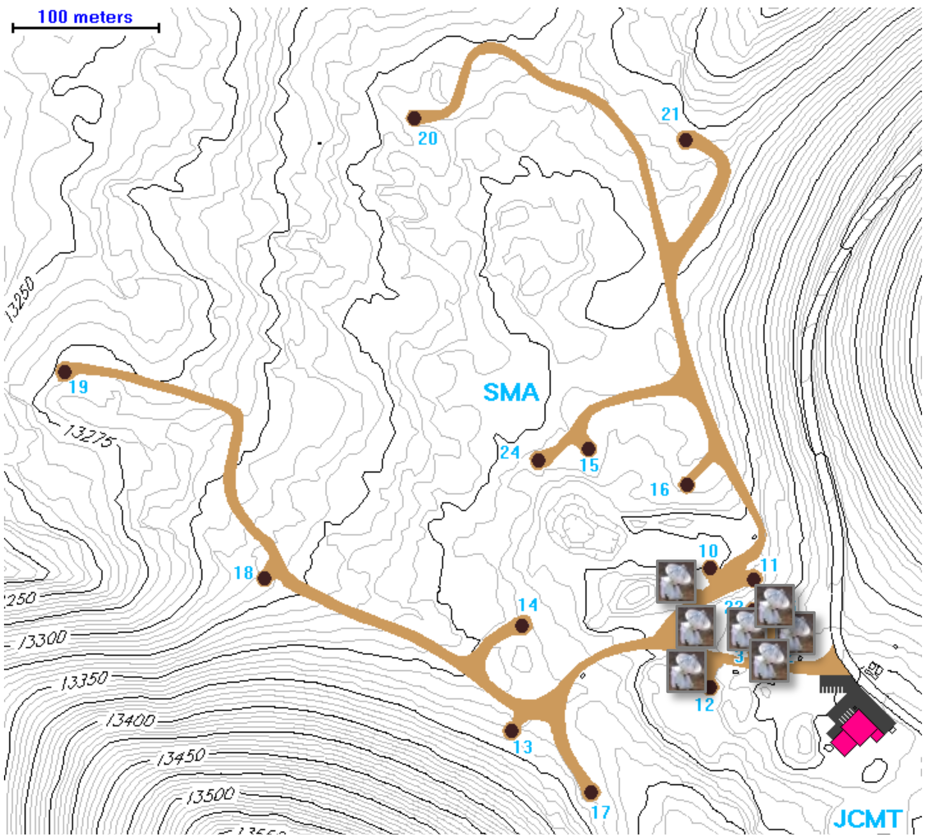
Sir Martin Ryle
1918-1984

1974 Nobel Prize in Physics

Aperture Synthesis Telescopes for Wavelengths less than 3 mm

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling



SMA antenna locations on October 23, 2009

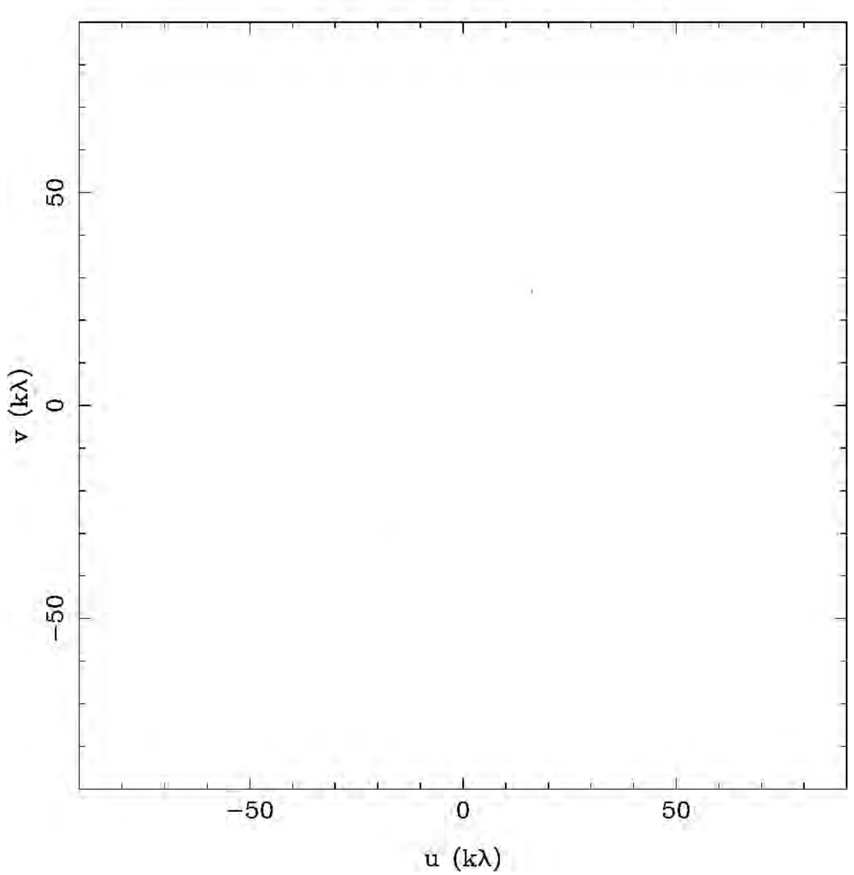
ν = 345 GHz

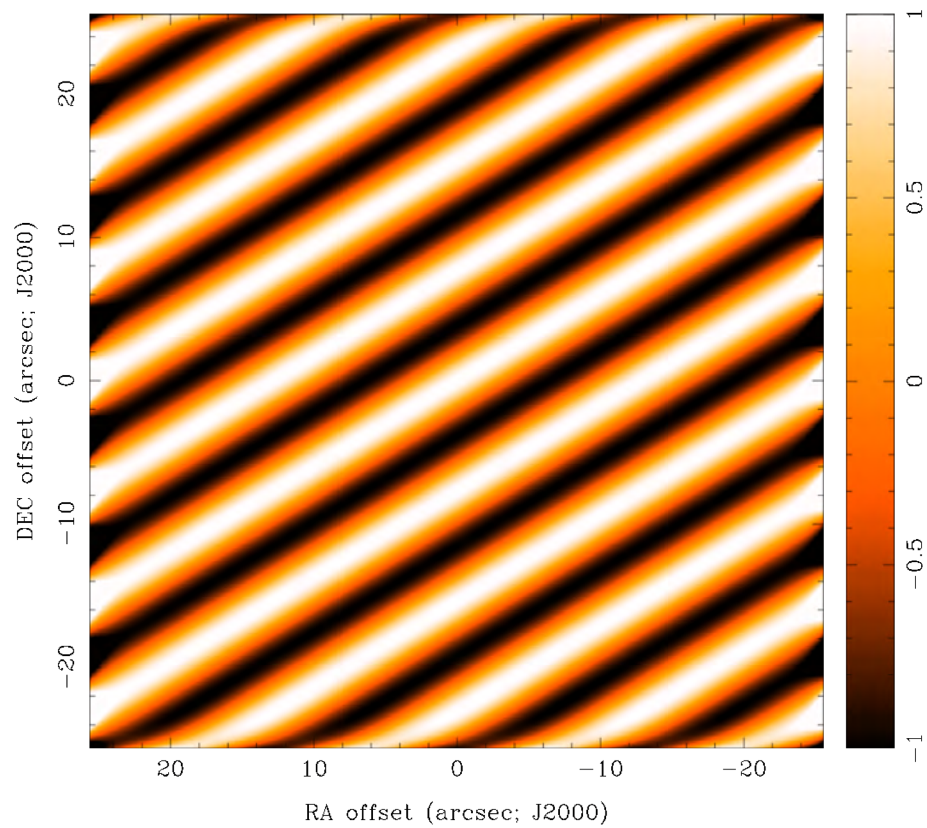
source at dec = + 22 deg

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

2 antennas, 1 x 30s sample

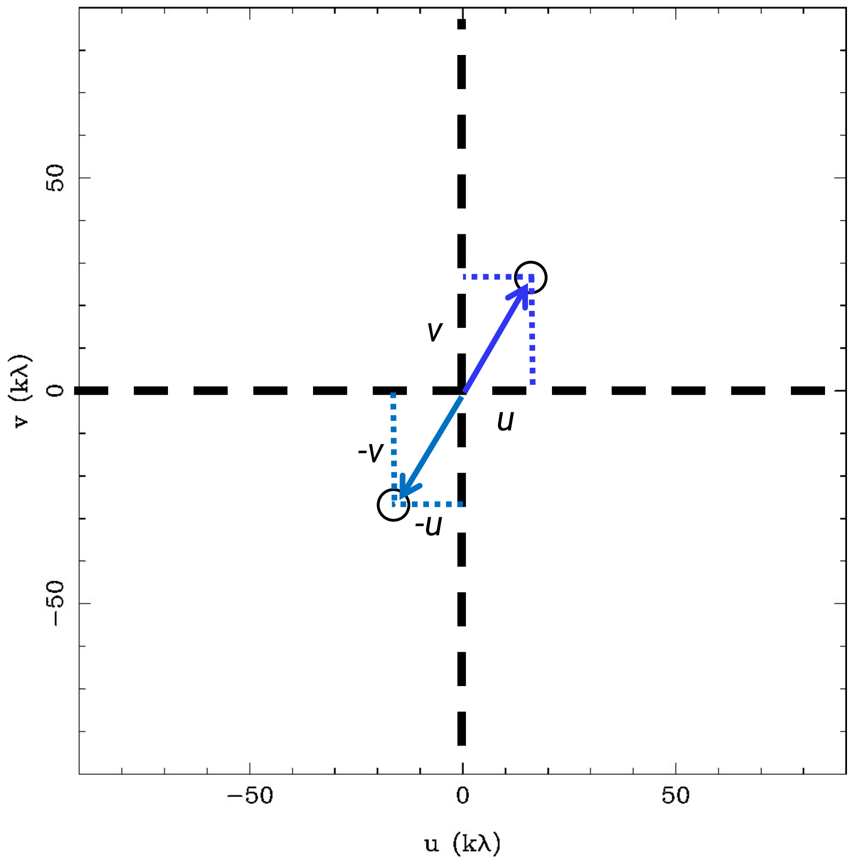


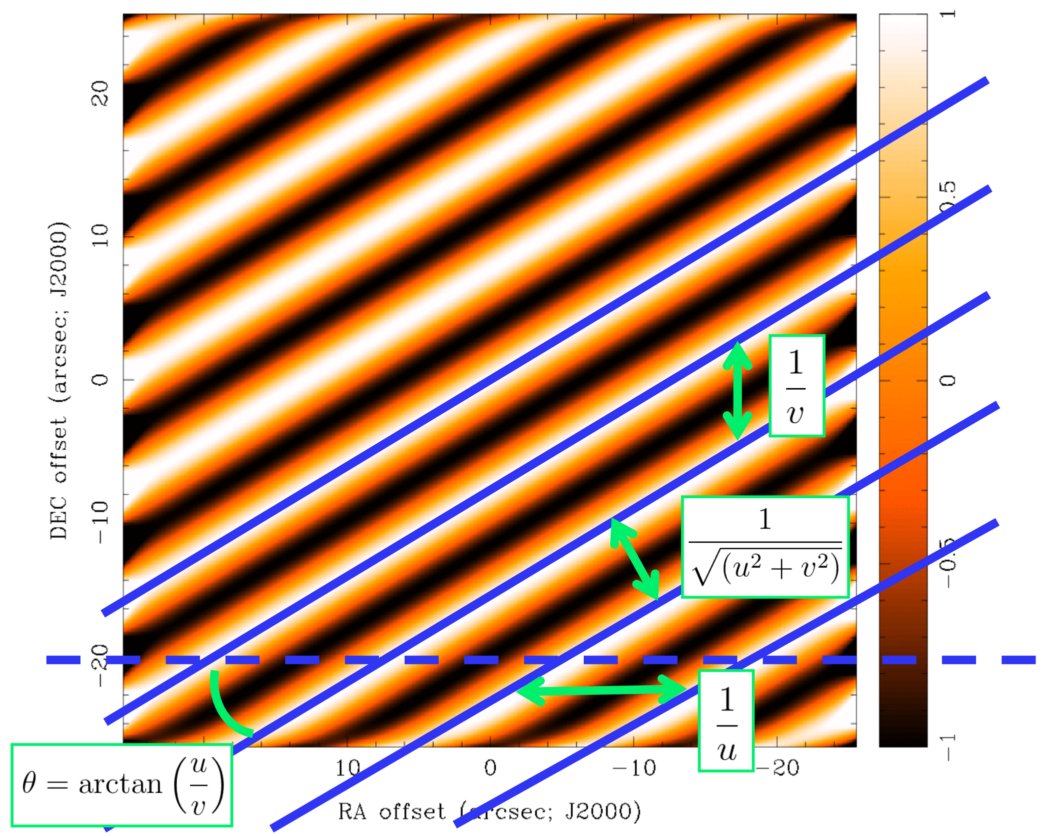


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

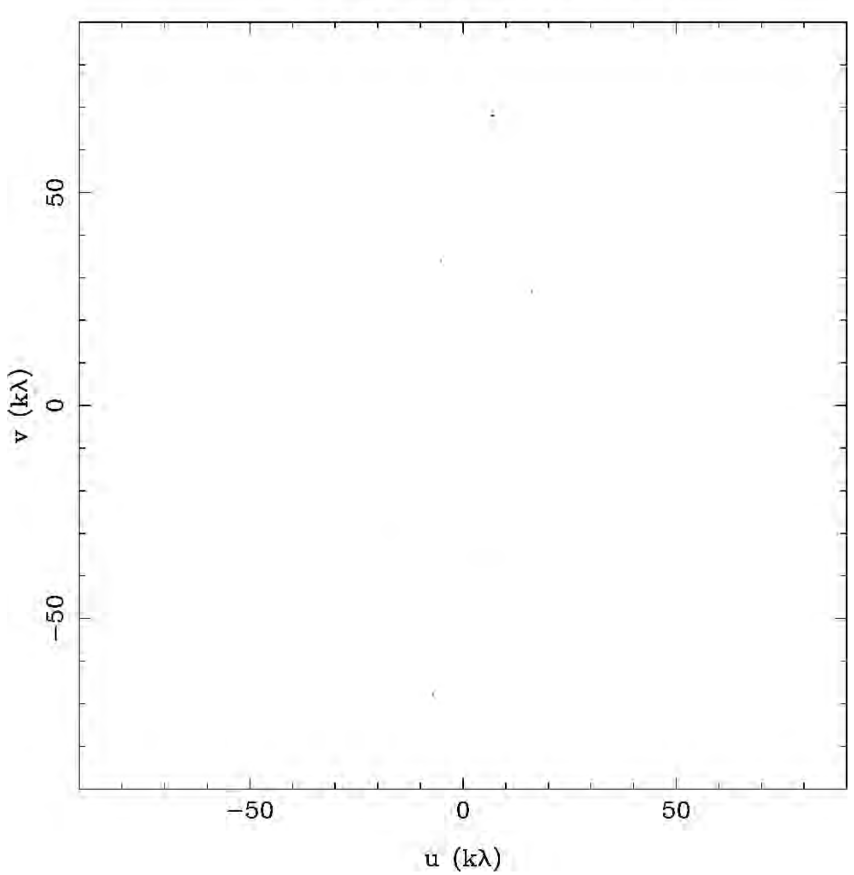
2 antennas, 1 x 30s sample

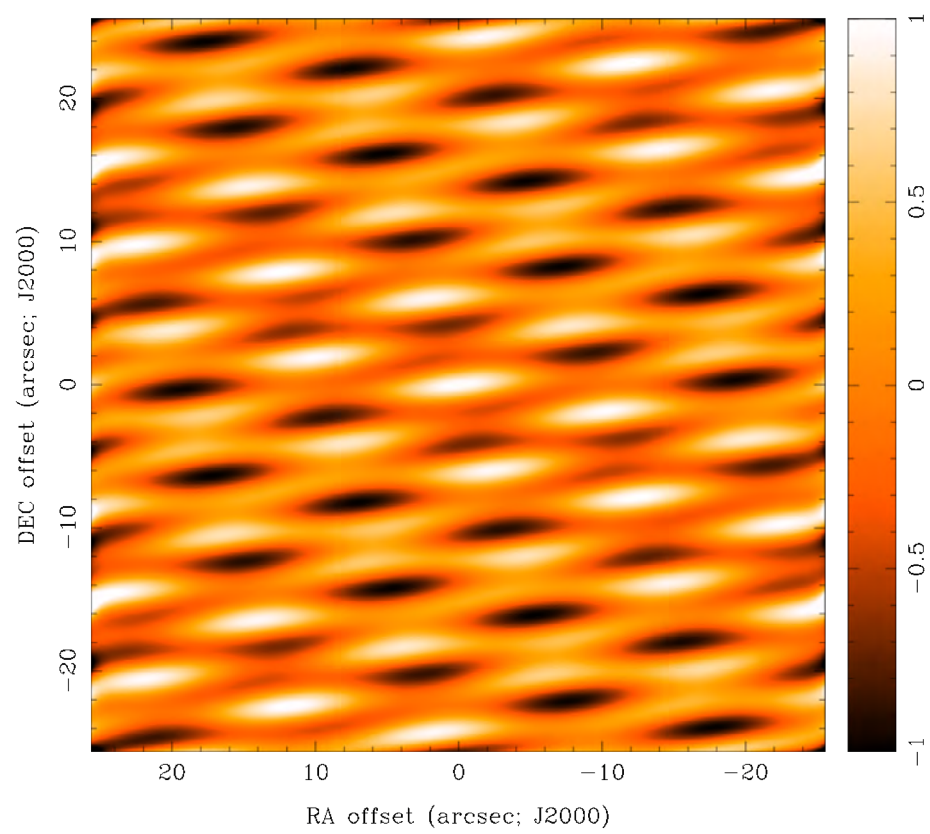




Example SMA *(u,v)* Plane Sampling

3 antennas, 1 x 30s sample

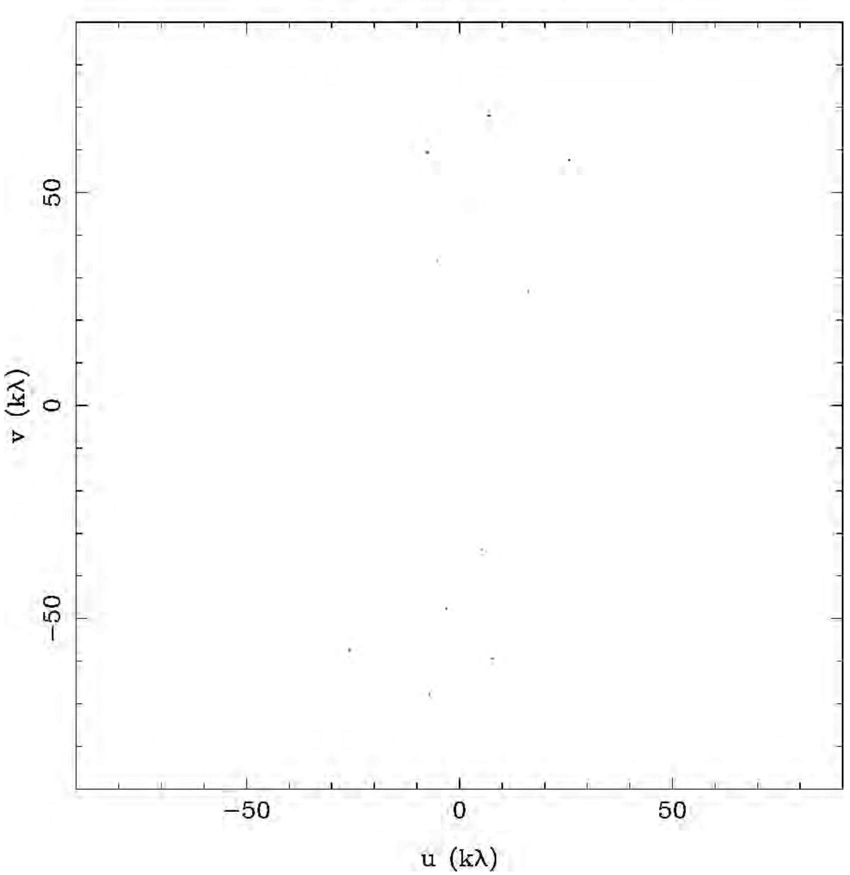


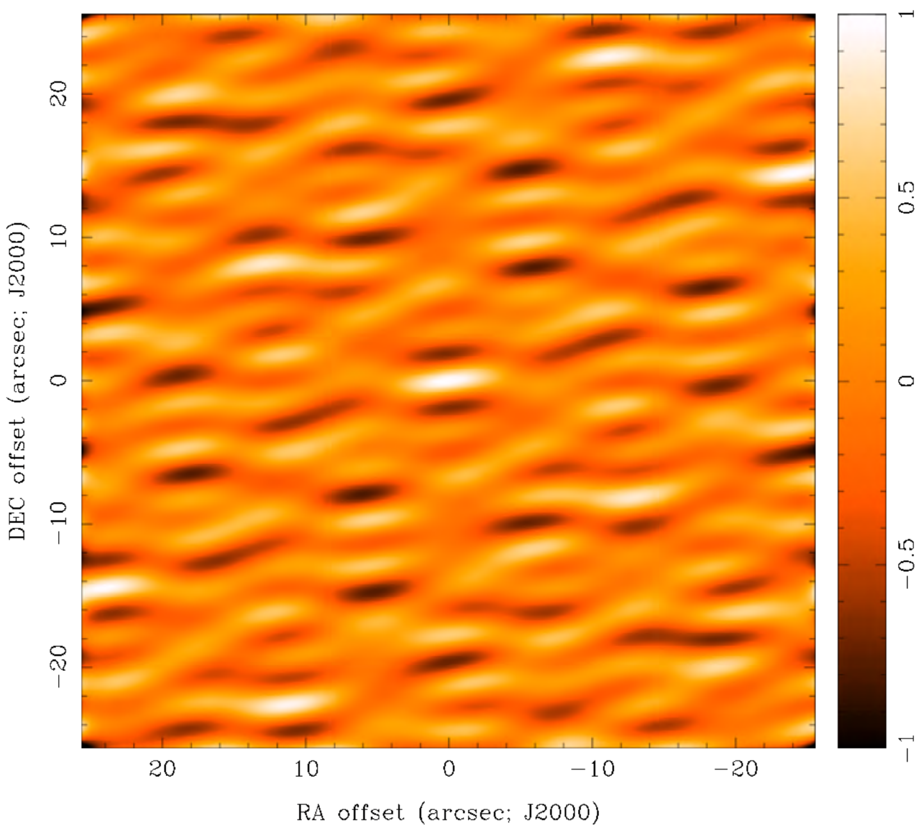


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

4 antennas, 1 x 30s sample

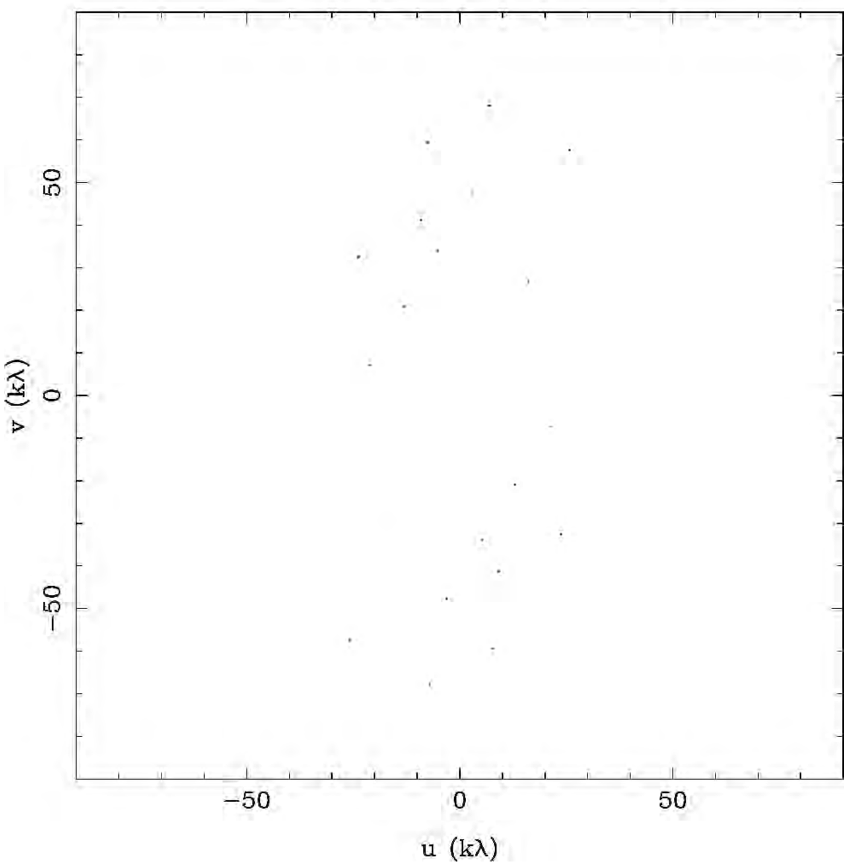


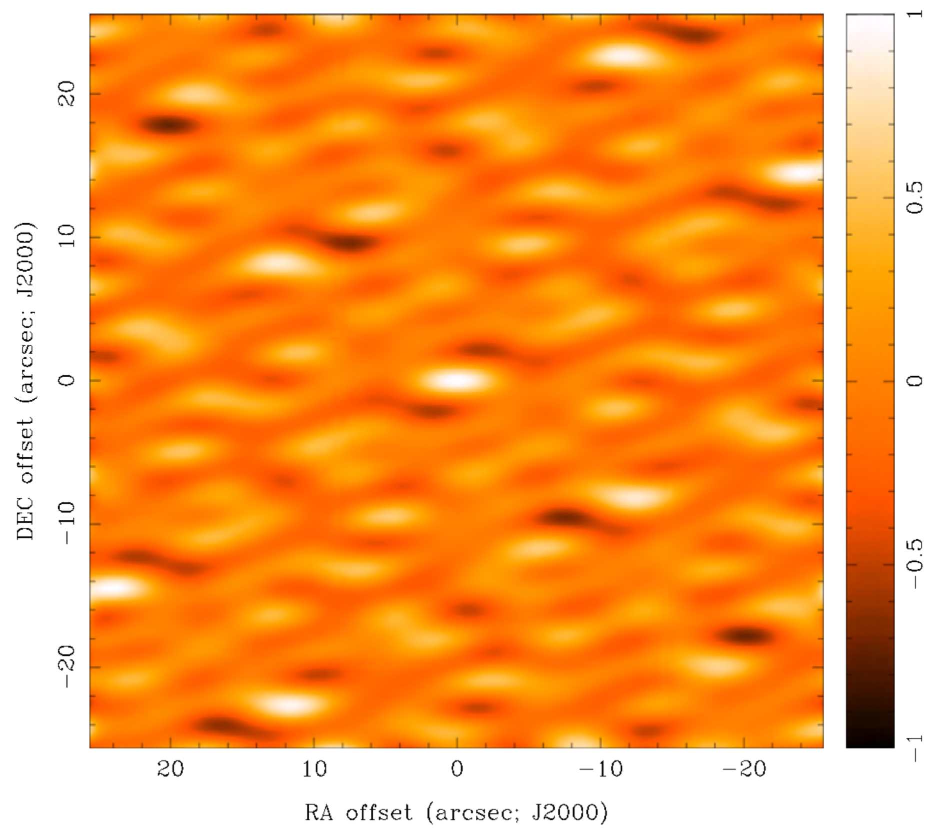


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

5 antennas, 1 x 30s sample

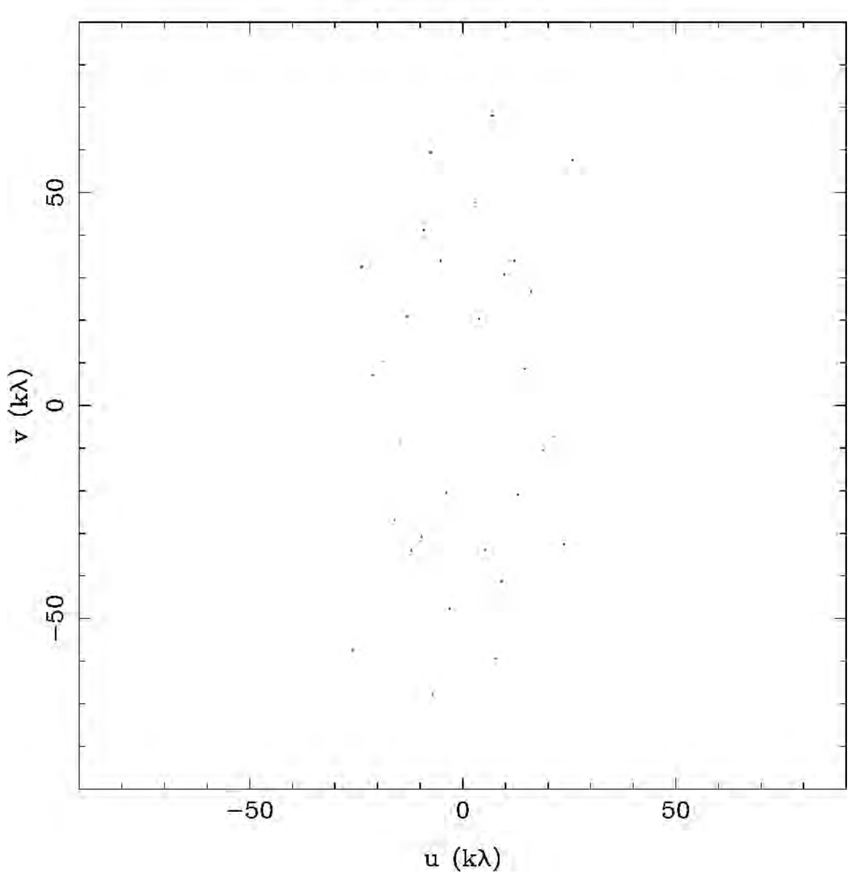


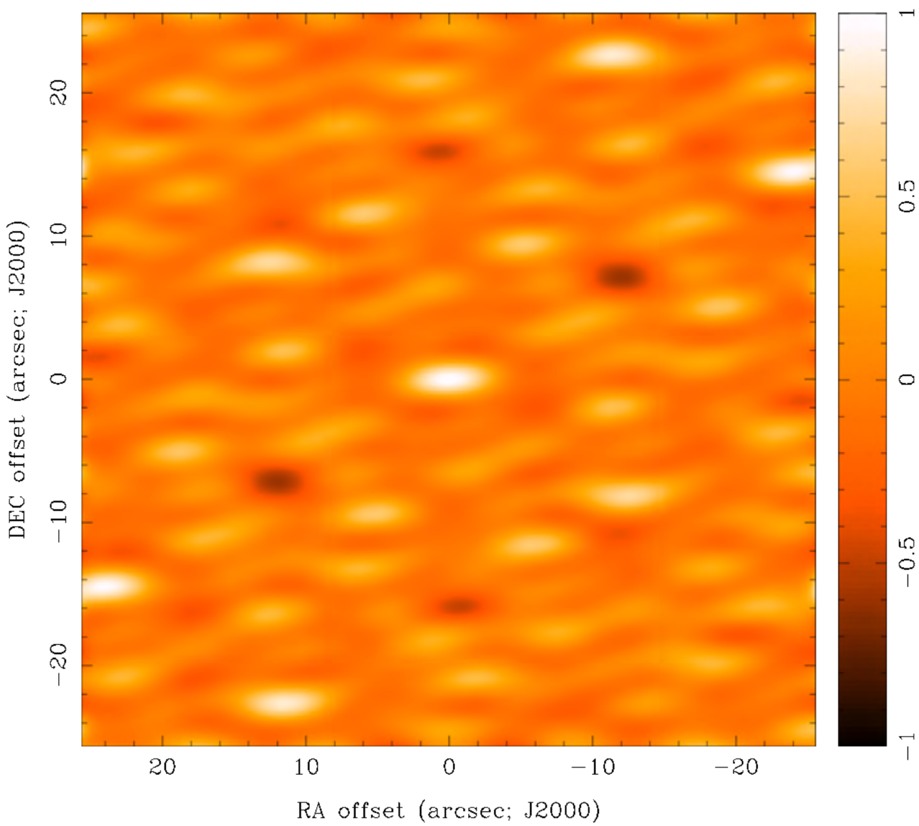


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

6 antennas, 1 x 30s sample

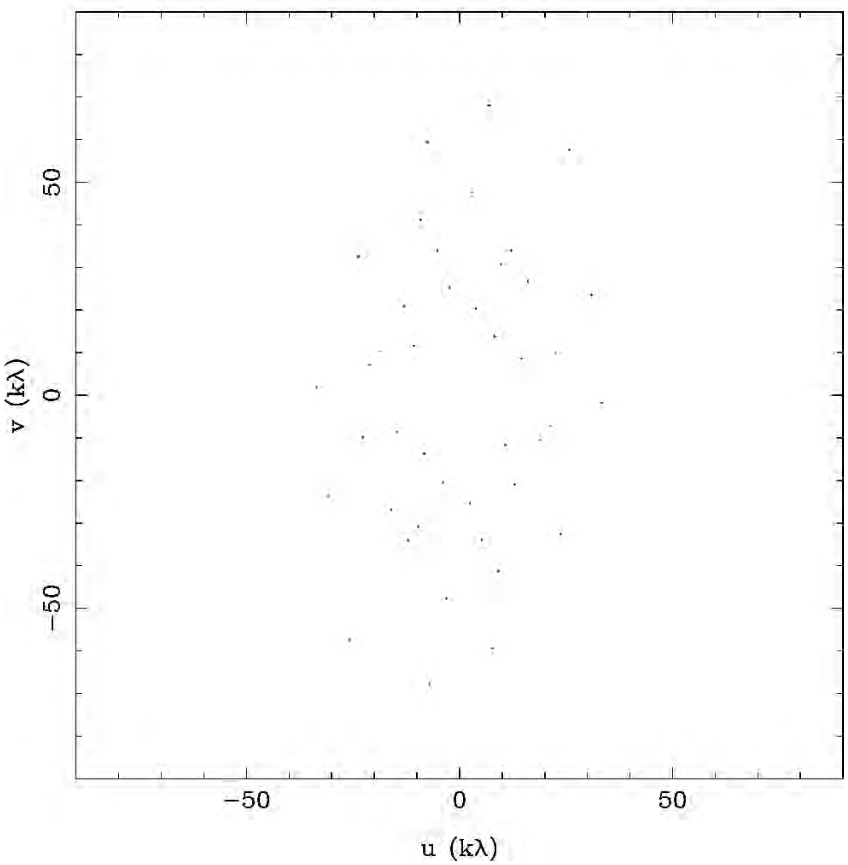


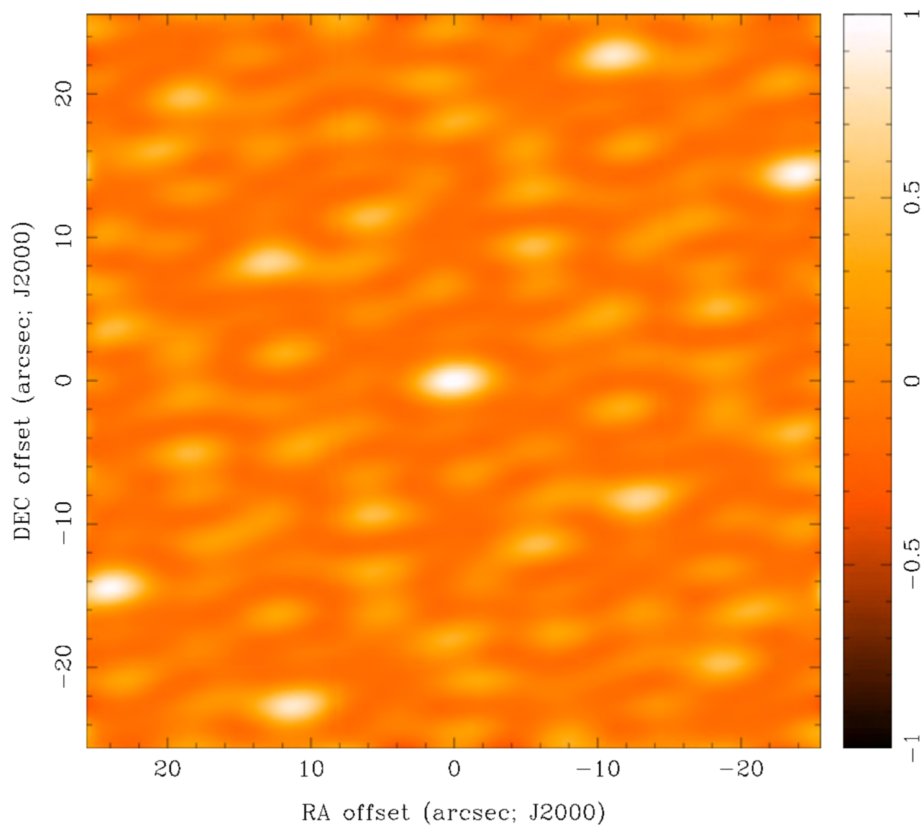


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

7 antennas, 1 x 30s sample

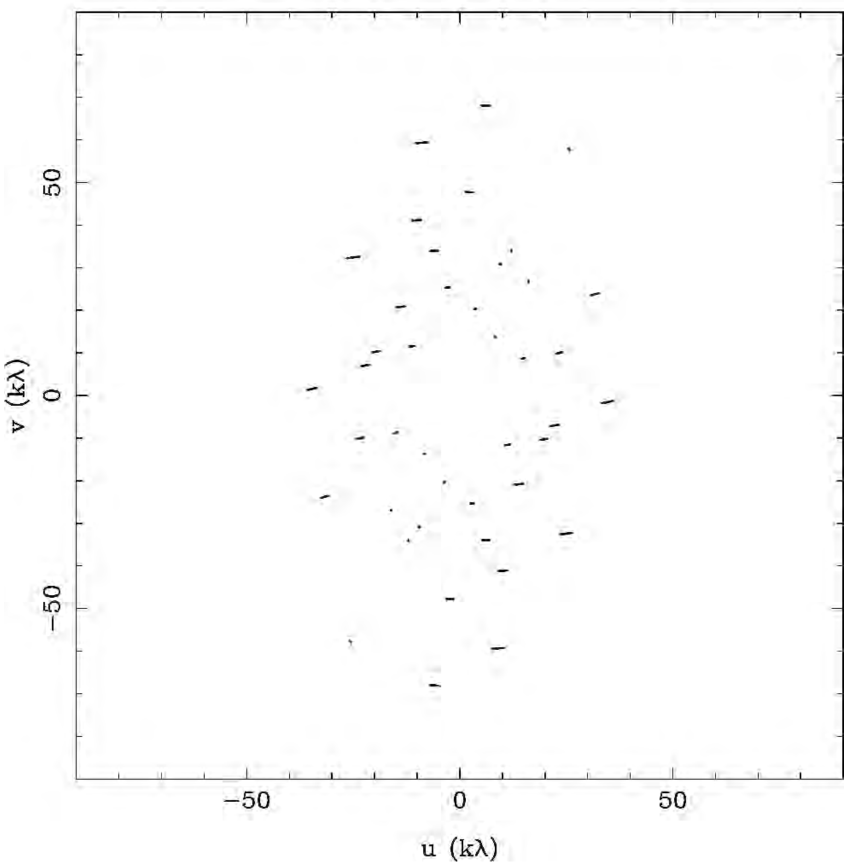


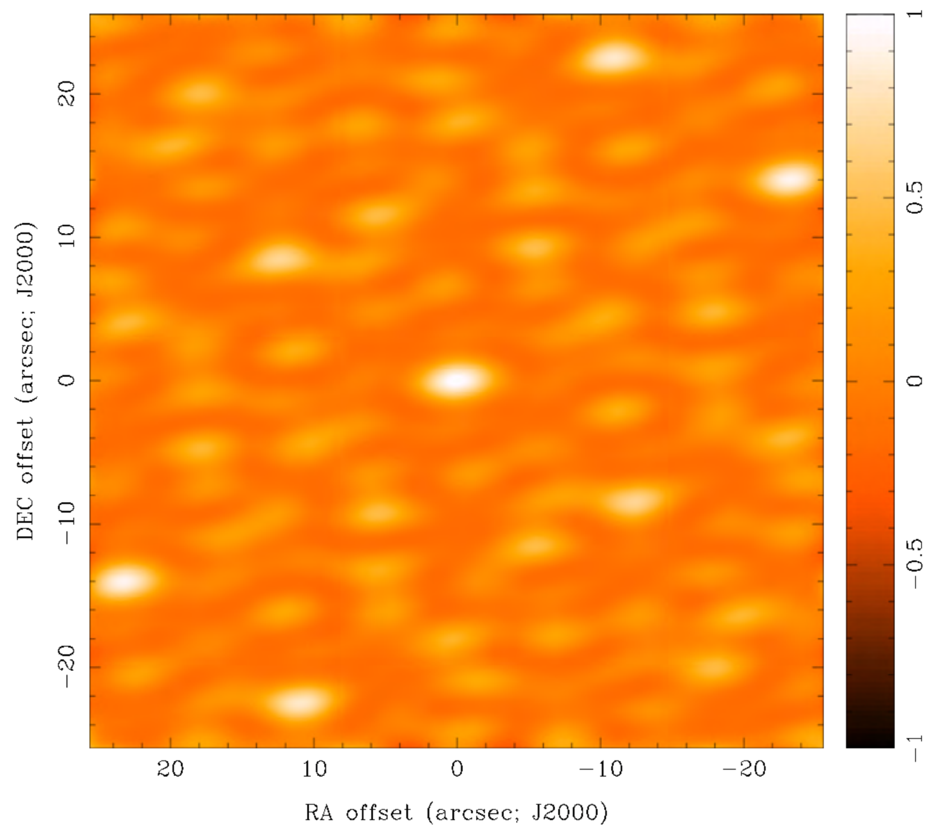


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

7 antennas, 10 x 30s samples

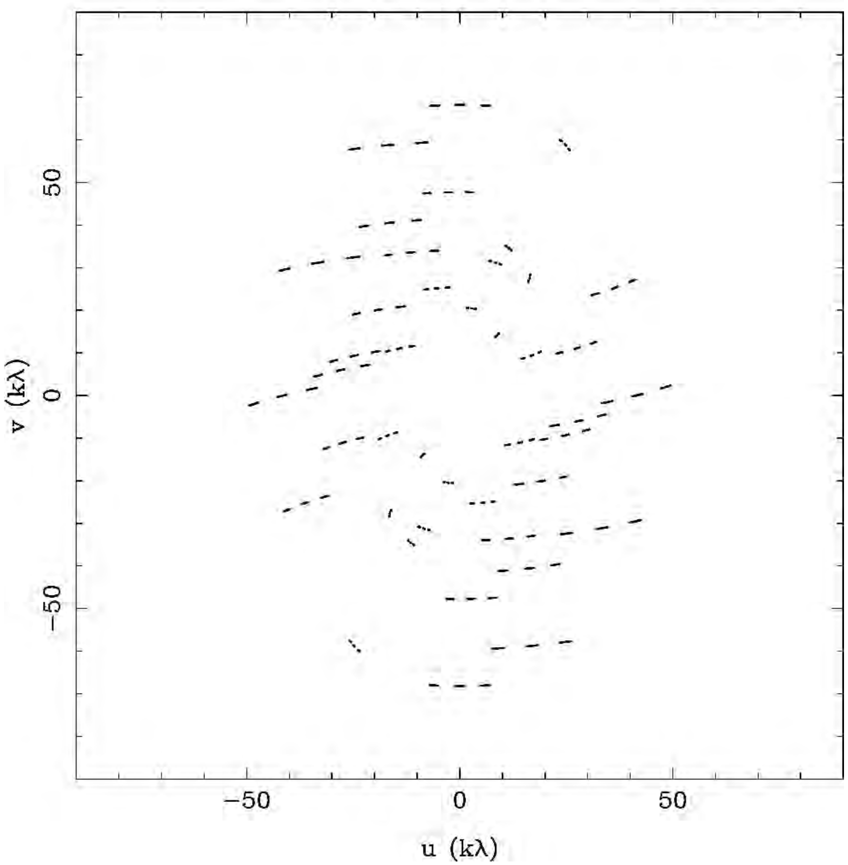


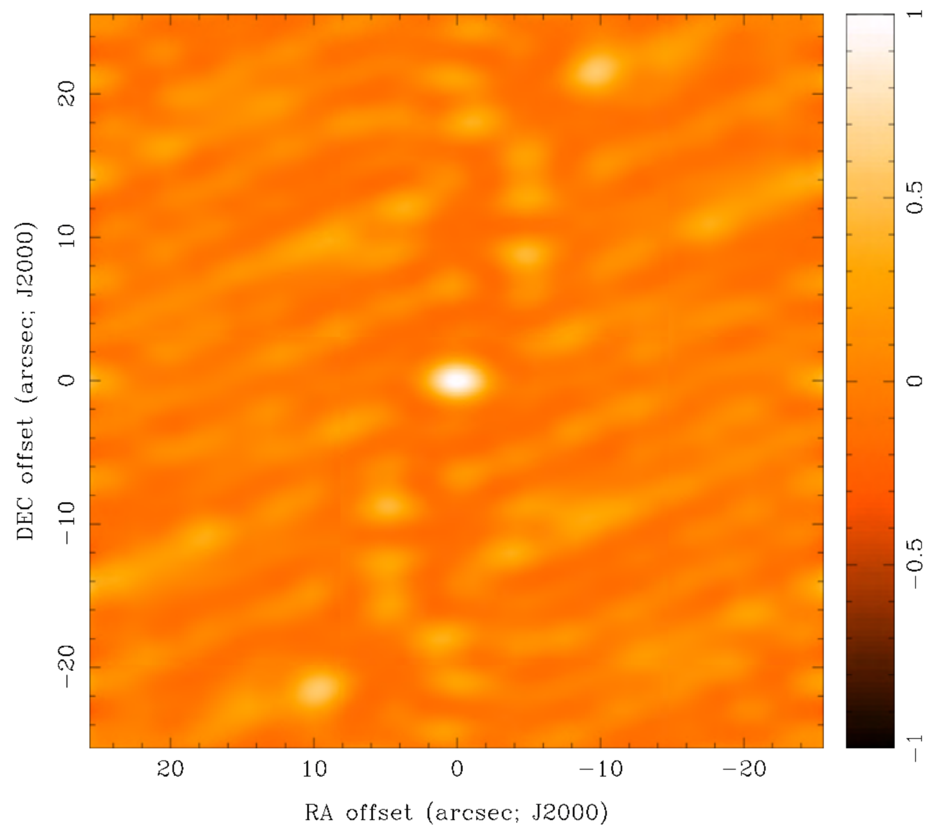


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

7 antennas, 1 hour

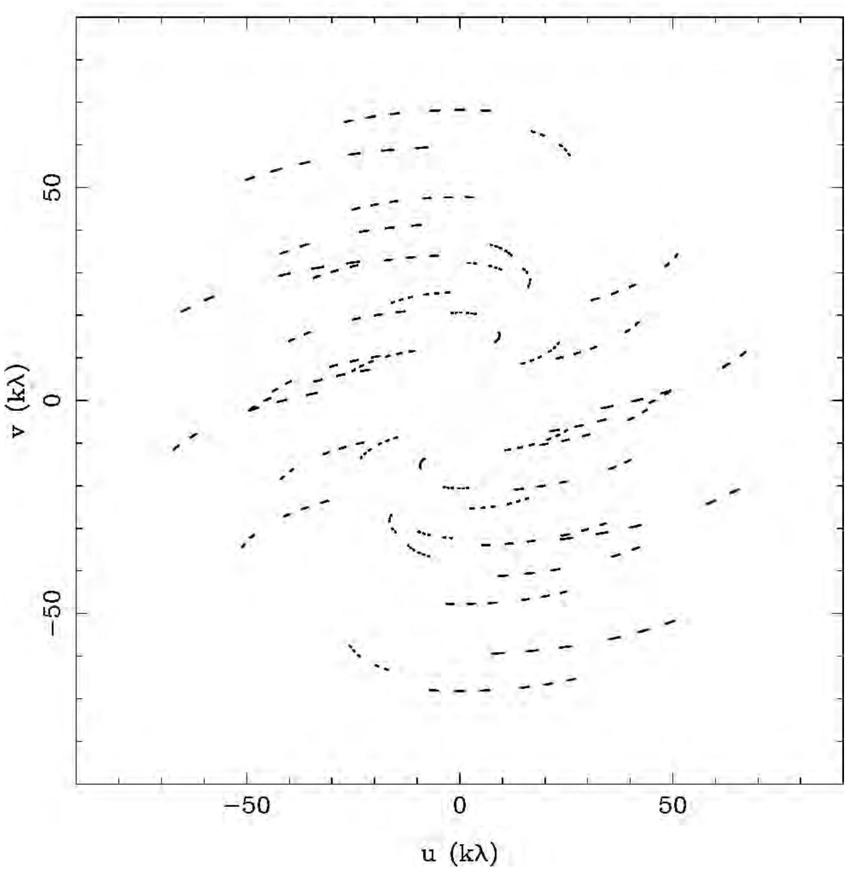


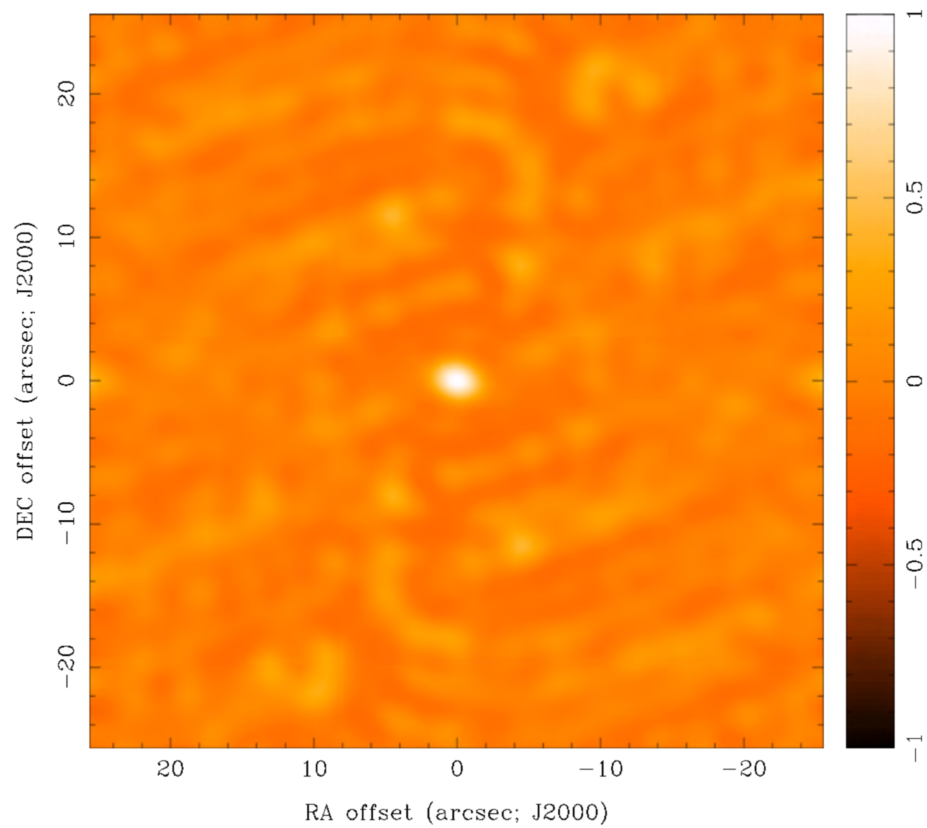


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

7 antennas, 3 hours

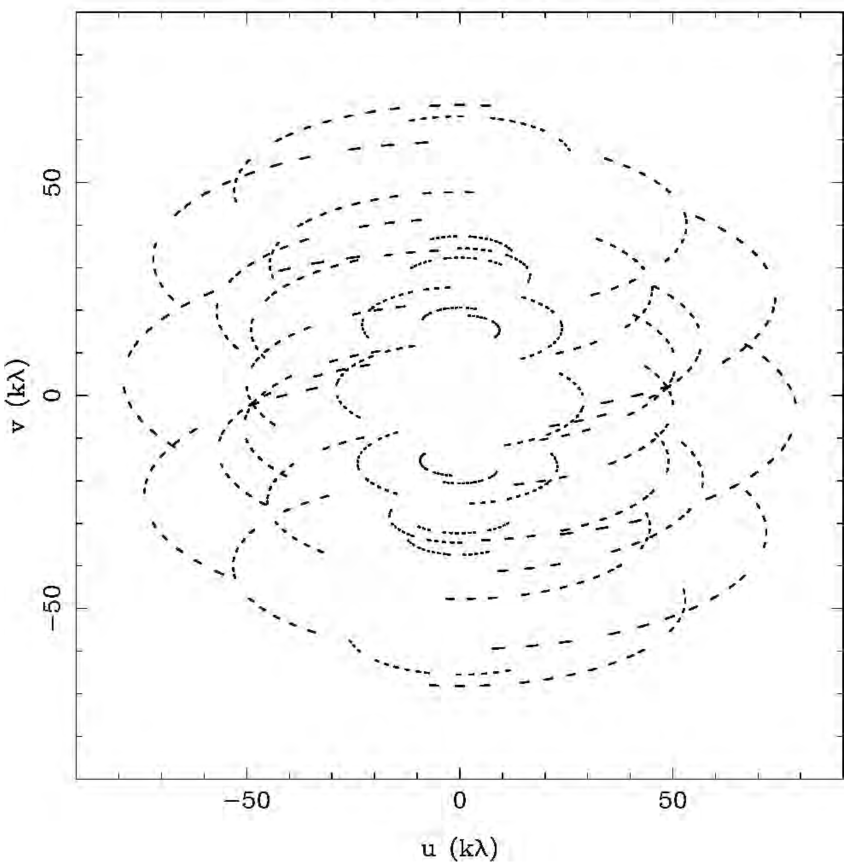


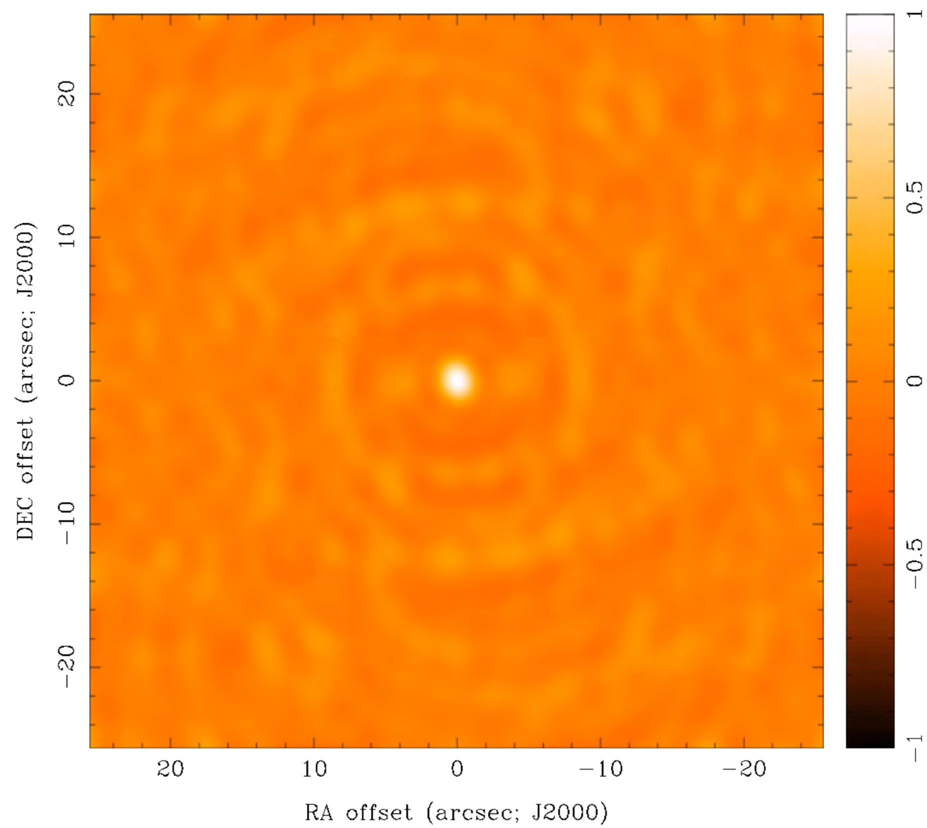


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

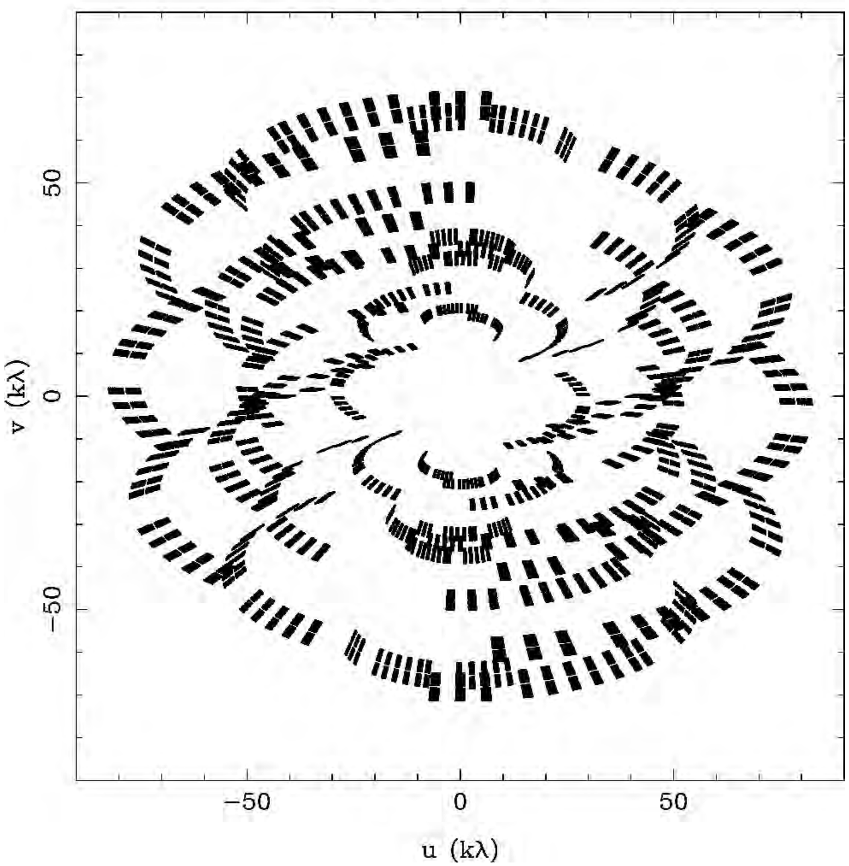
7 antennas, 7 hours





CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

SMA Multi-frequency Synthesis



for continuum

“multi-frequency synthesis”

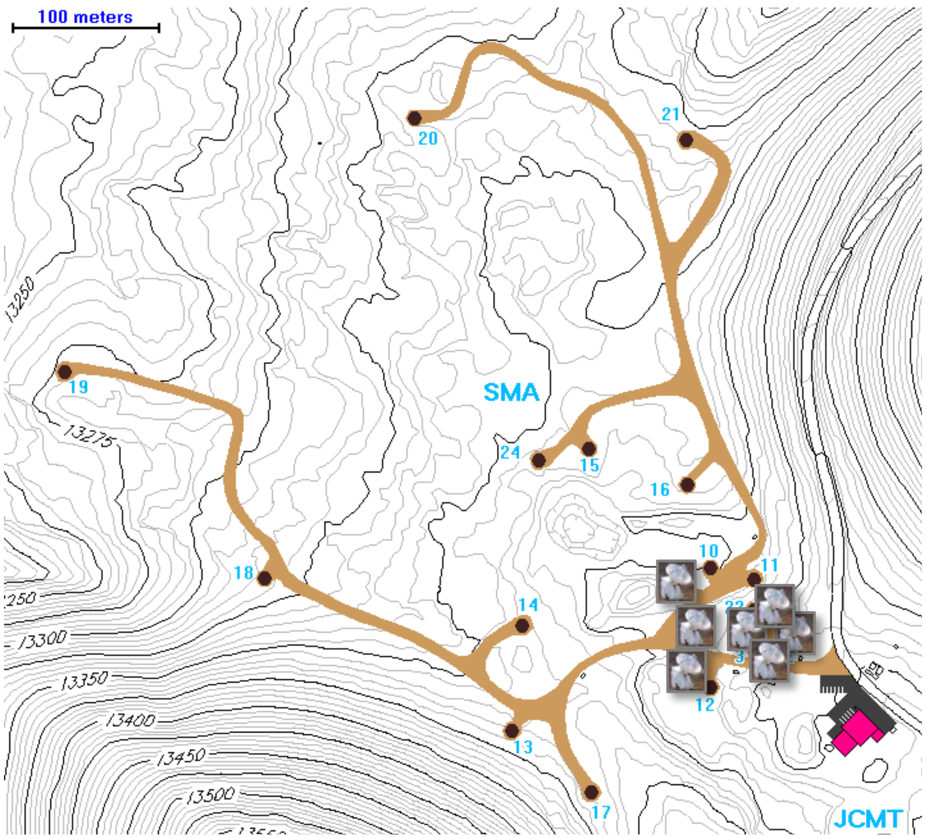
e.g. SWARM 32 GHz

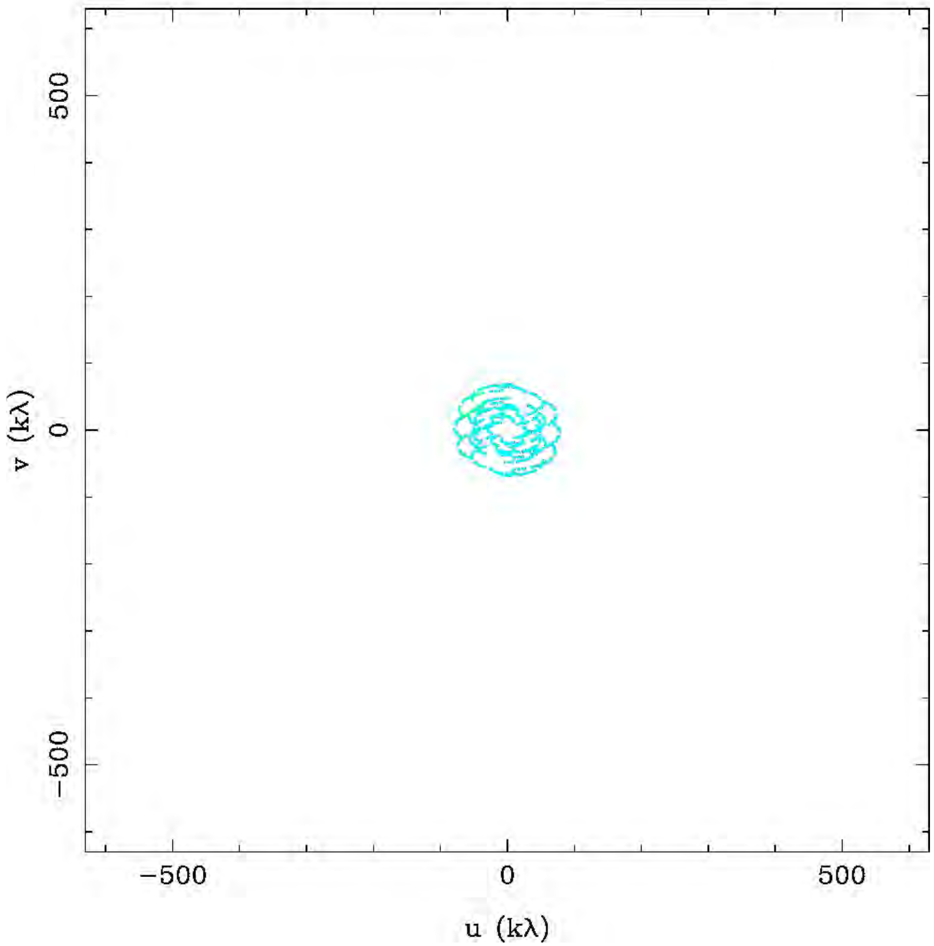
(8 GHz x 2 SB x 2 Pol)

→ *(u,v)* samples spread radially

Example SMA *(u,v)* Plane Sampling

COM configuration of 7 SMA antennas, ν = 345 GHz, dec = + 22 deg

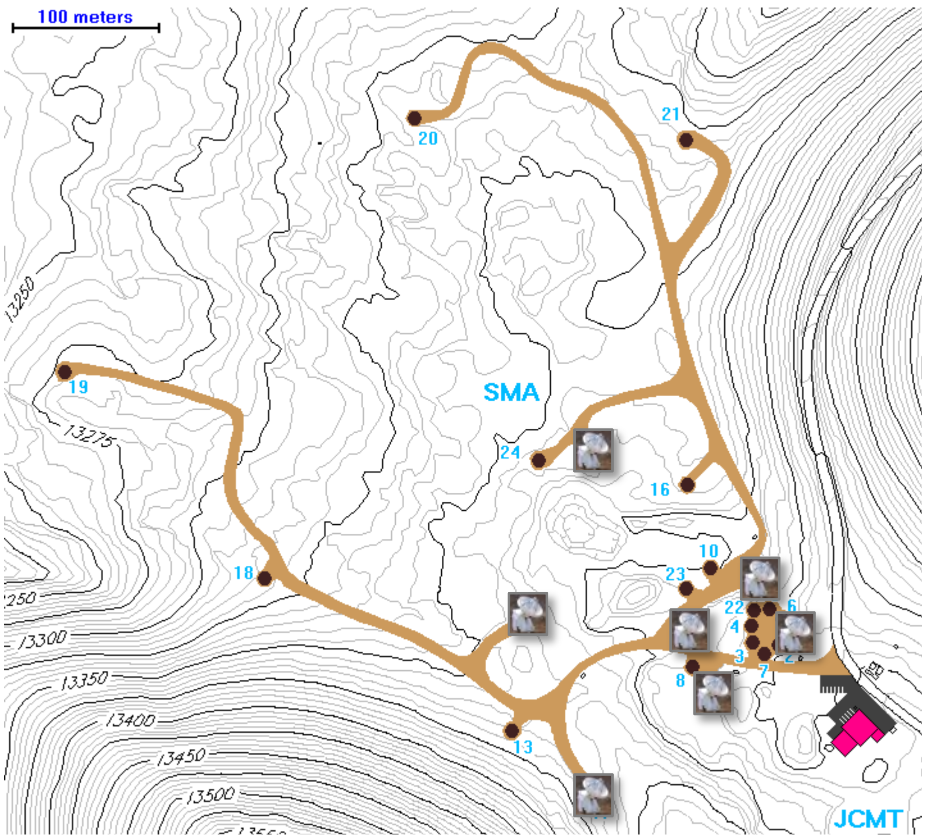


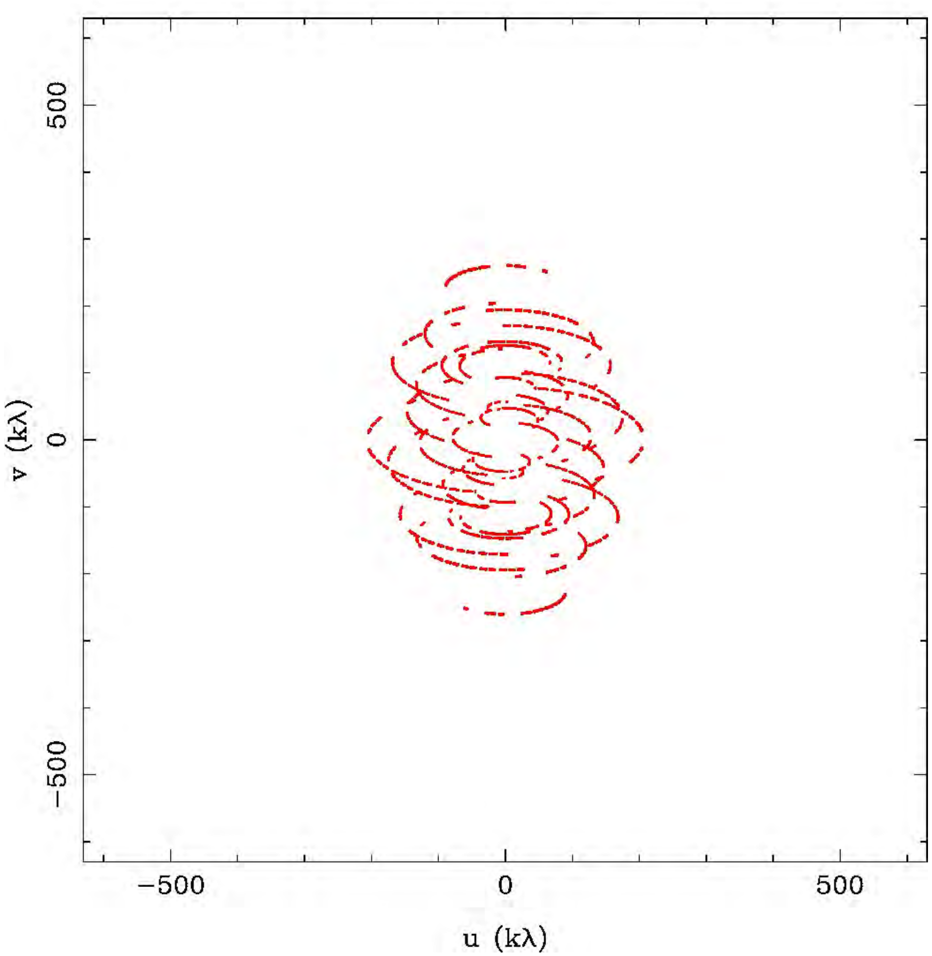


CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Example SMA *(u,v)* Plane Sampling

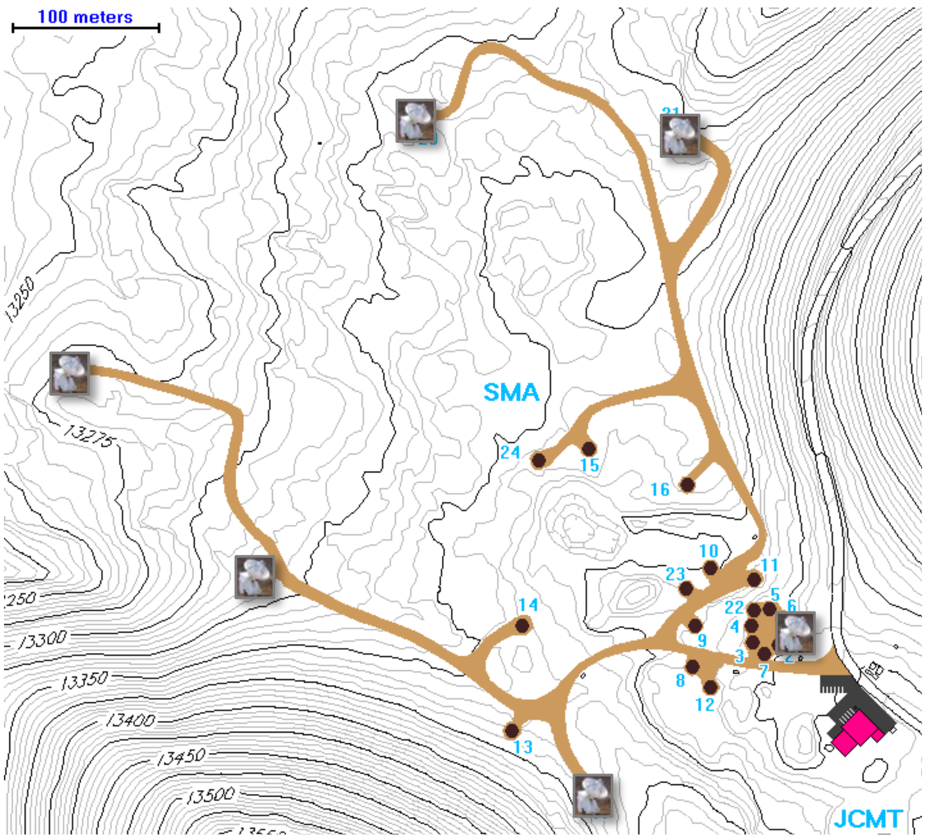
EXT configuration of 7 SMA antennas, ν = 345 GHz, dec = + 22 deg

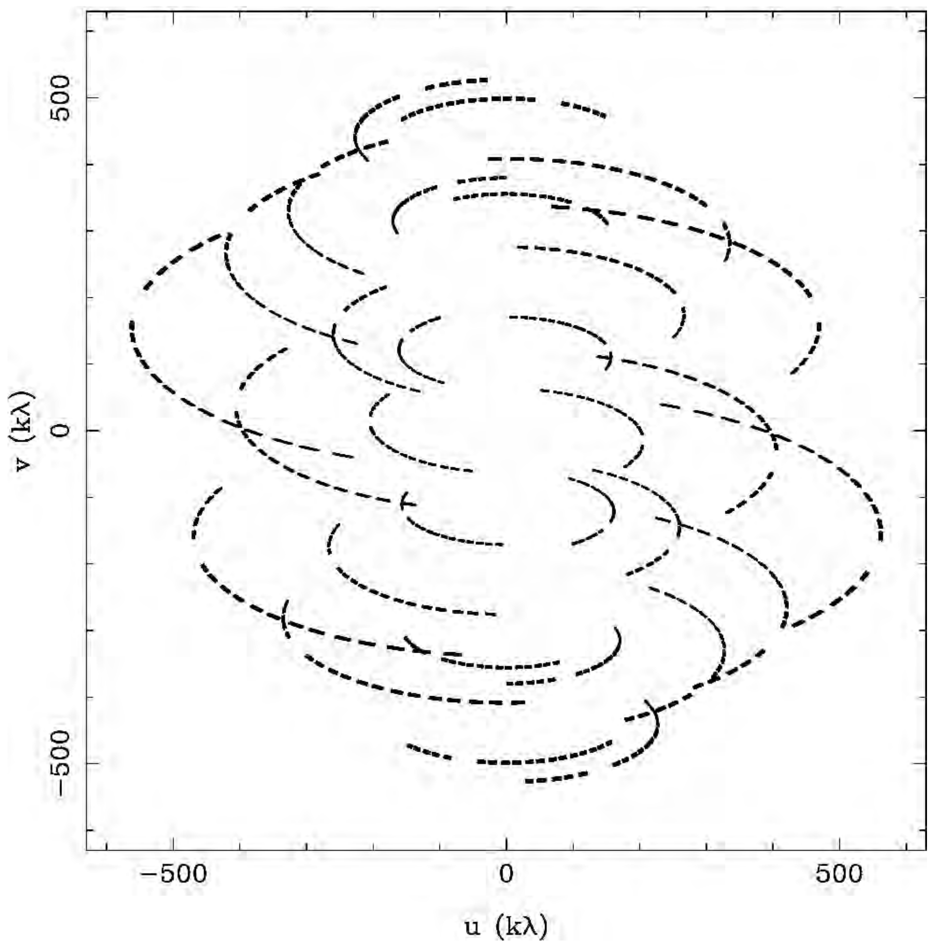




Example SMA *(u,v)* Plane Sampling

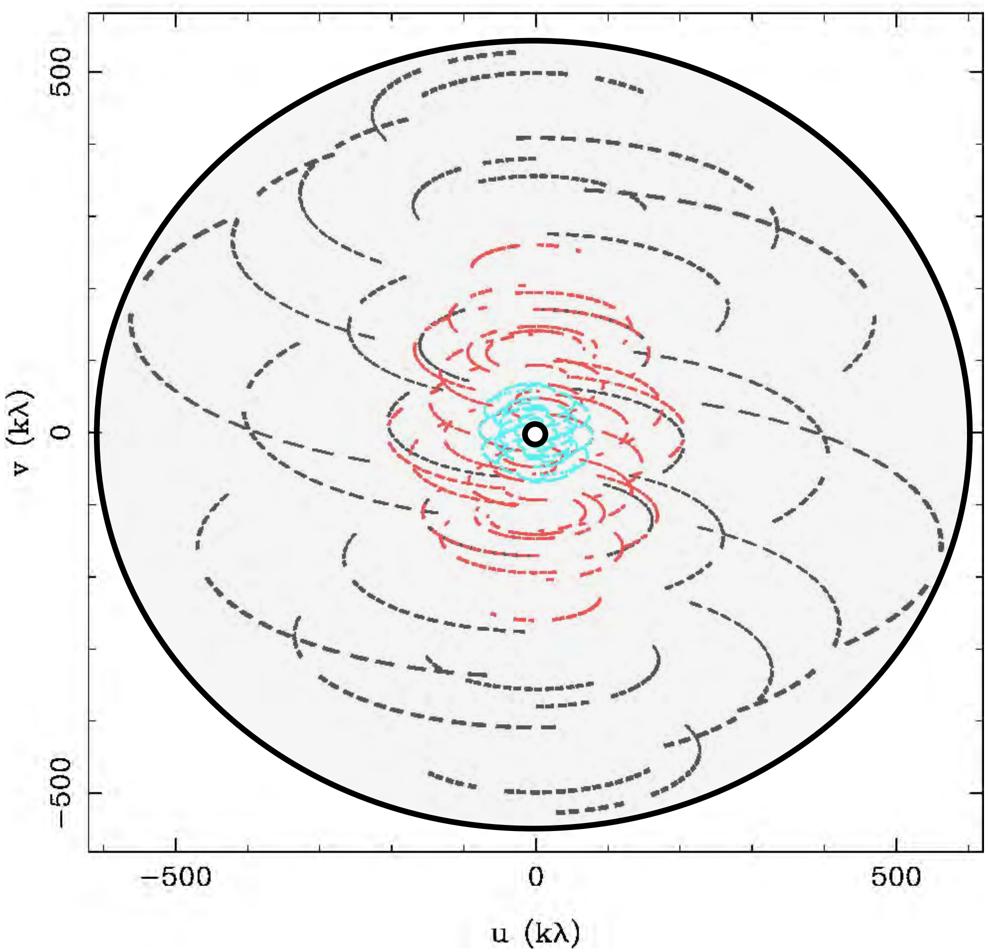
VEX configuration of 6 SMA antennas, ν = 345 GHz, dec = + 22 deg





Implications of *(u,v)* Plane Sampling

samples of *V(u,v)* are limited by array and Earth-sky geometry



outer boundary

- no info on smaller scales
- resolution limit

inner boundary

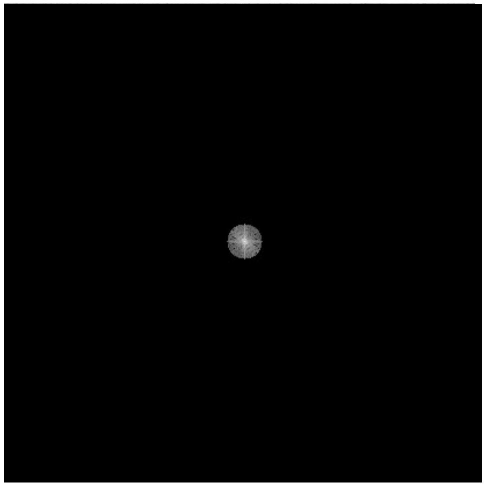
- no info on larger scales
- extended sources invisible

irregular coverage in between

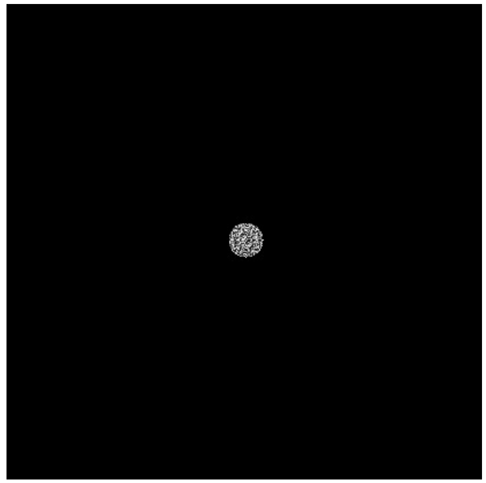
- sampling theorem violated
- information missing

Inner and Outer *(u,v)* Boundaries

V(u,v) amplitude



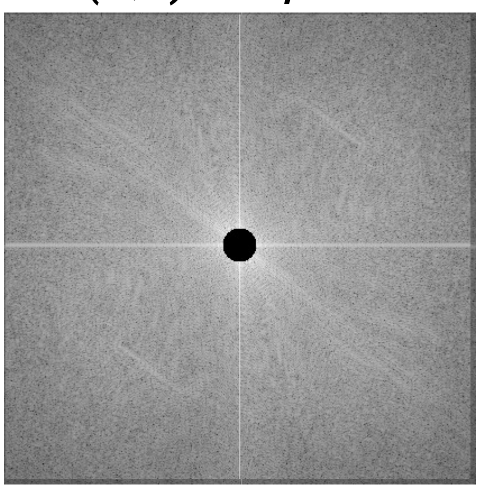
V(u,v) phase



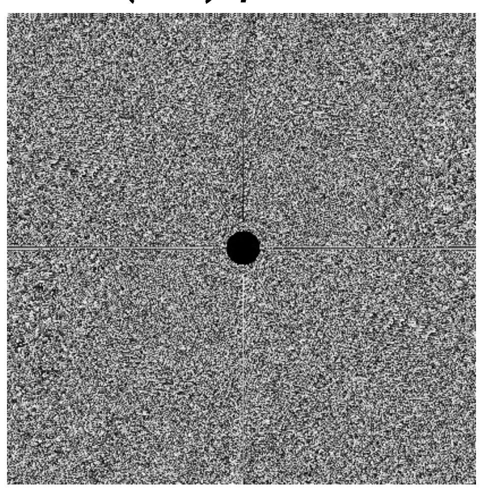
$\xrightarrow{\mathcal{F}}$

T(l,m)

V(u,v) amplitude

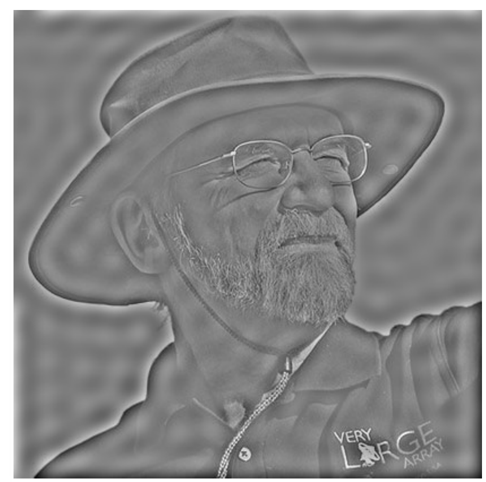


V(u,v) phase



$\xrightarrow{\mathcal{F}}$

T(l,m)



Calibrated Visibilities… Now what?

analyze *V(u,v)* samples directly by model fitting

- good for simple structures, e.g. point sources, symmetric rings
- for a purely statistical description of sky brightness, e.g. fluctuations
- visibilities have well defined noise properties

use incomplete, noisy samples of *V(u,v)* to make an image

- Fourier transform to create a distorted version of *T(l,m)*
- account for incomplete sampling to build up a model of *T(l,m)*
- work with the model of *T(l,m)* to do science

some software packages and key tasks for imaging and deconvolution

- CASA `tclean`
- miriad `invert,clean,restore`
- AIPS `imagr`
- GILDAS `uvmap,clean`

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Formal Description of Imaging

- sample Fourier domain at discrete points

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

- Fourier transform sampled visibility function

$$V(u,v)S(u,v) \xrightarrow{\mathcal{F}} T^D(l,m)$$

- apply the convolution theorem

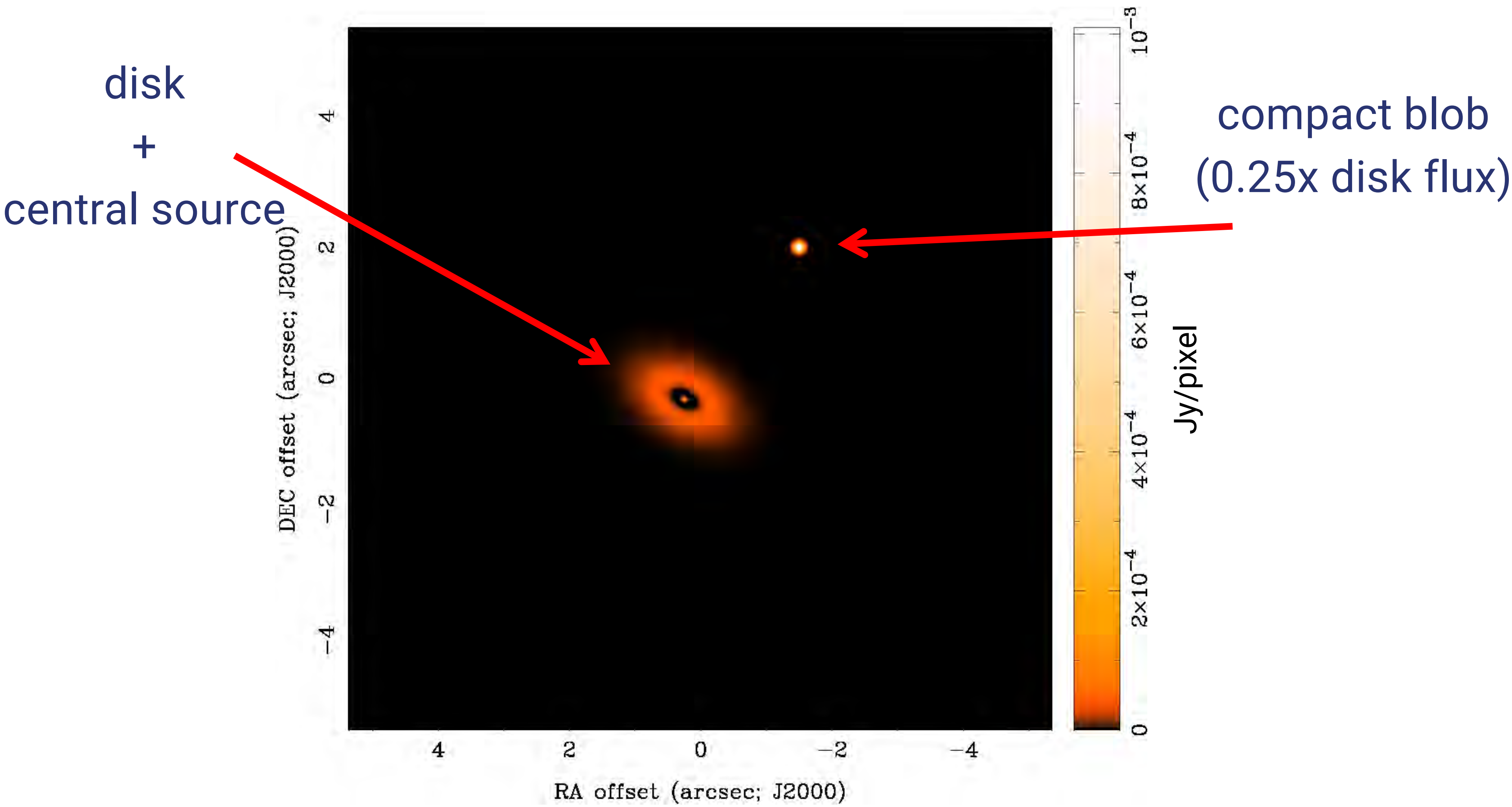
$$T(l,m) * s(l,m) = T^D(l,m)$$

where the Fourier transform of the sampling pattern $s(l,m) \xrightarrow{\mathcal{F}} S(u,v)$

is the “point spread function” or “synthesized beam”

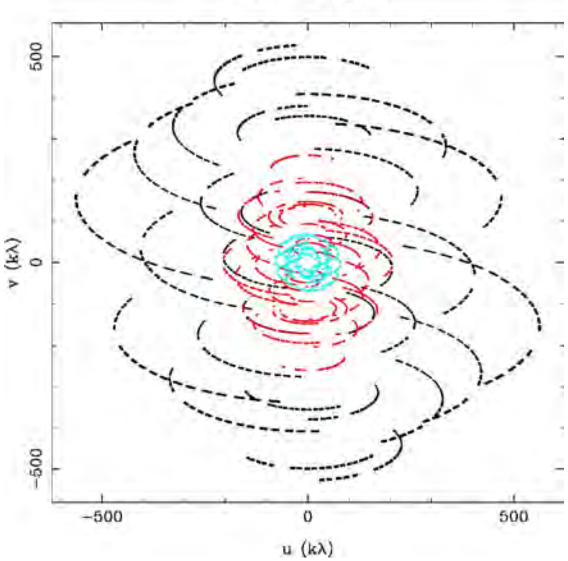
- radio astronomy jargon:
  the “dirty image” is the true image convolved with the “dirty beam”

Example model sky brightness $T(l,m)$

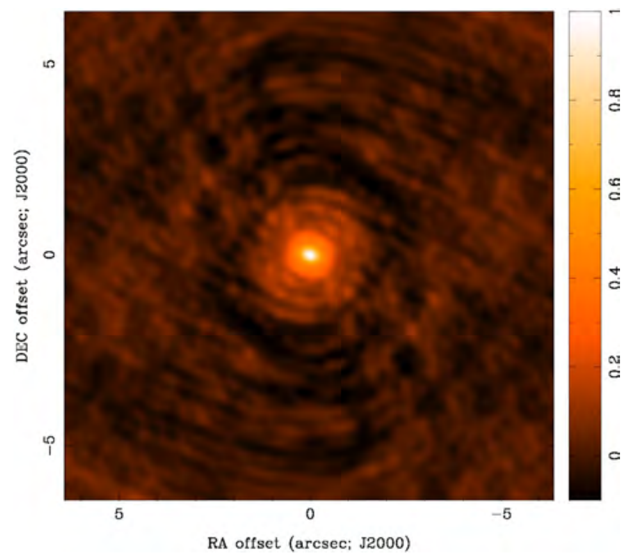


Dirty Beam and Dirty Image

S(u,v)



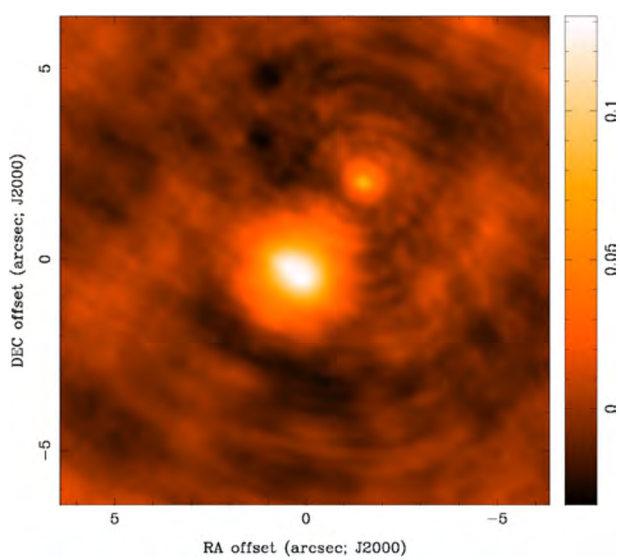
$\xrightarrow{\mathcal{F}}$



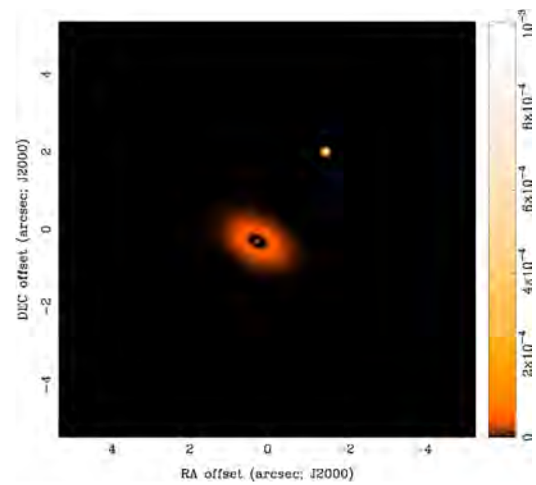
s(l,m)
"dirty beam"

$*$

T^D(l,m)
"dirty image"



=



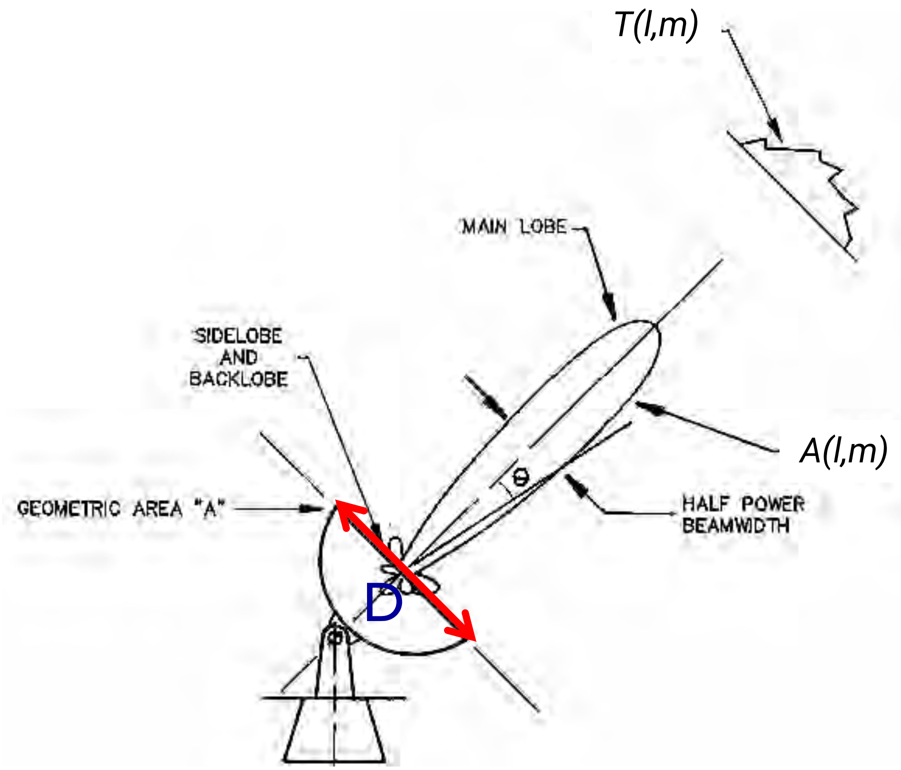
T(l,m)

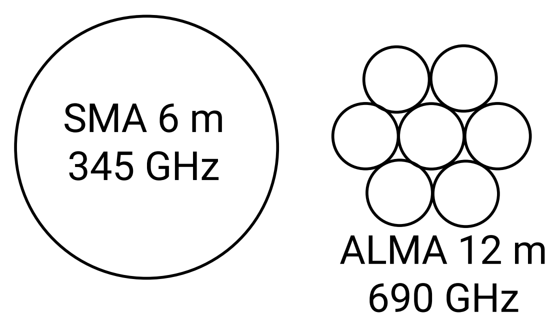
CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Field of View

- **antenna response *A(l,m)* is not uniform across the sky**
  - "primary beam" fwhm ~ λ/D
  - response beyond primary beam usually not important for SMA observations

- **antenna response *A(l,m)* modifies the sky brightness distribution**
  - $T(l,m) \rightarrow T(l,m)A(l,m)$
  - can correct with division by $A(l,m)$ in the image plane
  - larger fields of view require multiple pointings of the array antennas = mosaicking





Fast Fourier Transform

Fast Fourier Transform (FFT) is used to compute the Fourier integral

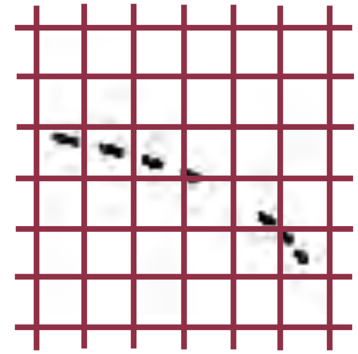
- Direct computation by simple summation is slow
  - must compute sin and cos functions directly for prescribed combinations of visibilities: $O(N^4)$ for N^2 image cells
  - *can* be managed computationally for modest values of N
  - but generally not practical for most modern imaging applications
- FFT algorithm
  - much faster than simple summation: $O(N^2 \log N)$
  - but FFT speed gain does not come for free
- FFT requires data on a regularly spaced grid… and aperture synthesis does not provide $V(u,v)$ samples on a regularly spaced grid
- also must pay attention to aliasing effects due to periodic form

Gridding

Gridding is used to resample *V(u,v)* onto a regular *(u,v)* grid to use FFT

- conventional approach is to use convolution
- *(u,v)* cell size ≈ 0.5 x D, where D = antenna diameter

$$V^G(u,v) = V(u,v)S(u,v) * G(u,v) \xrightarrow{F} T^D(l,m)g(l,m)$$



- prolate spheroidal functions are popular "gridding convolution functions"
  - compact in *(u,v)* plane: minimize smoothing, allow efficient gridding
  - drops to near zero at image edges, suppresses aliasing
- other gridding steps may include

functions that apply primary beam weighting and offsets ("mosaicking")

functions that apply wide-field phase shifts ("W projection")

functions that correct for primary beam differences ("A projection")

Image Size and Pixel Size

image size

- natural choice is often full primary beam *A(l,m)*
- e.g. SMA at 870 μm, 6 m antennas → image size 2 x 35 arcsec
- if there are strong sources in *A(l,m)* sidelobes, then the FFT will alias them into the image → make larger image (or image outlier fields)

pixel size

- satisfy Nyquist-Shannon sampling theorem for longest baselines

$$\Delta l < \frac{1}{2u_{max}} \qquad \Delta m < \frac{1}{2v_{max}}$$

- in practice, use 3 to 5 pixels across dirty beam main lobe to aid deconvolution process
- e.g. SMA at 870 μm, baselines to 500 m → pixel size < 0.1 arcsec

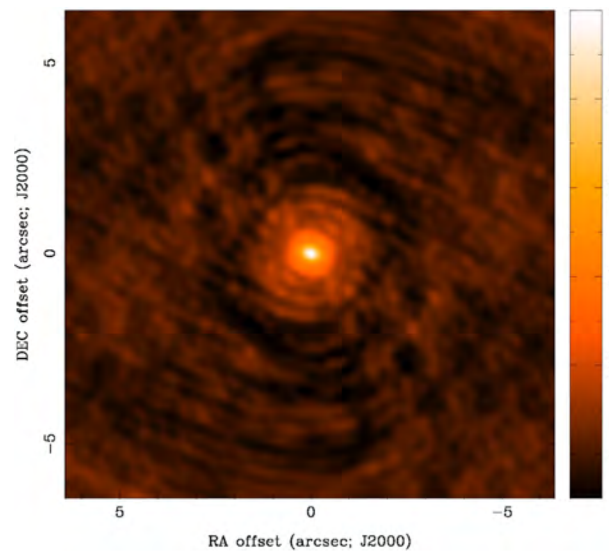
Visibility Weighting Schemes

- We can change the angular response of the interferometer by adding additional samples of *V(u,v);* this requires additional observing time
- Another way is to introduce a weighting function, *W(u,v),* into the visibility gridding process
  - *W(u,v)* modifies the sampling function: $S(u,v) \rightarrow S(u,v)W(u,v)$
  - changes the dirty beam shape
- *W(u,v)* can be used to bring out features on different angular scales from the same samples of *V(u,v)*

Natural Weighting

$W(u,v) = 1/\sigma^2$ in occupied cells (where σ is noise)

- advantages
  - maximize point source sensitivity
  - lowest rms in image
- disadvantages
  - usually many short baselines
    → lower angular resolution
  - many sample density variations
    → more structure in dirty beam



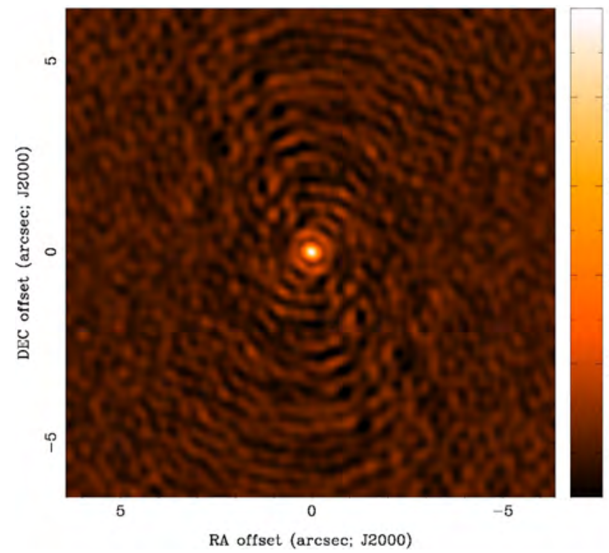
Gaussian fit to central core

0.59x0.50 arcsec

Uniform Weighting

W(u,v) inversely proportional to local density of samples; weight for occupied cell = constant

- advantages
  - fills *(u,v)* plane more uniformly → less structure in dirty beam
  - more weight to long baselines → higher angular resolution
- disadvantages
  - down weights some data → higher rms noise
  - can be trouble with sparse *(u,v)* coverage since cells with few samples have same weight as cells with many



Gaussian fit to central core

0.35x0.30 arcsec

Robust ("Briggs") Weighting

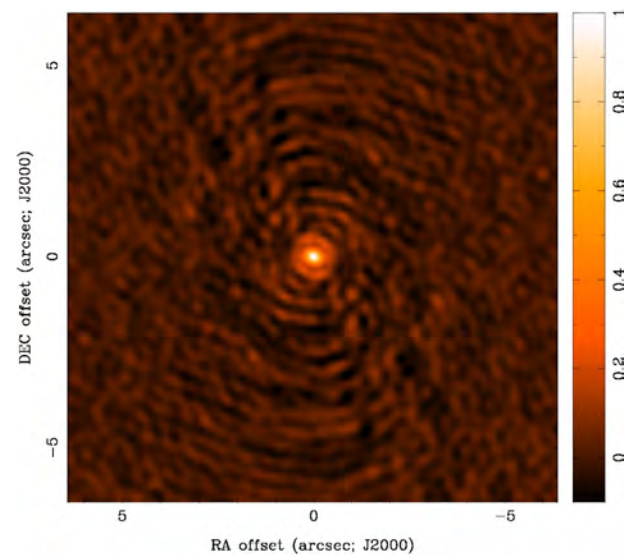
variant on uniform weighting that avoids giving too much weight to cells with low natural weight

- software implementations differ
- e.g. $$W(u,v) = \frac{1}{\sqrt{1 + S_N^2 / S_{thresh}^2}}$$

  S_N is cell natural weight

  S_{thresh} is a threshold parameter

- advantages
  - parameter for continuous variation between best point source sensitivity and highest angular resolution
  - usually can obtain most of natural weight sensitivity *at the same time* as most of uniform weight resolution (!)



Gaussian fit to central core

0.40x0.34 arcsec

Tapering

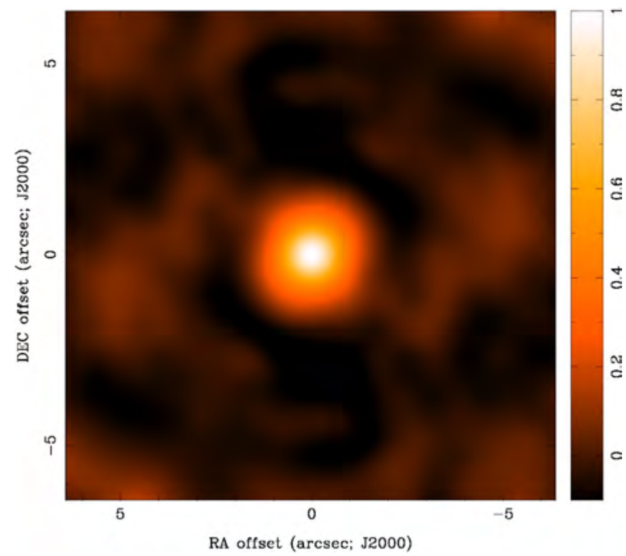
apodize *(u,v)* sampling by a Gaussian function

$$W(u,v) = \exp\left(-\frac{(u^2+v^2)}{t^2}\right)$$

t is an adjustable tapering parameter

- like convolving image by a Gaussian
- advantages
  - less weight to long baselines

  $\rightarrow$ lower angular resolution that can improve sensitivity to extended structure
- disadvantages
  - higher noise per beam
  - limits to usefulness as more and more data are down weighted



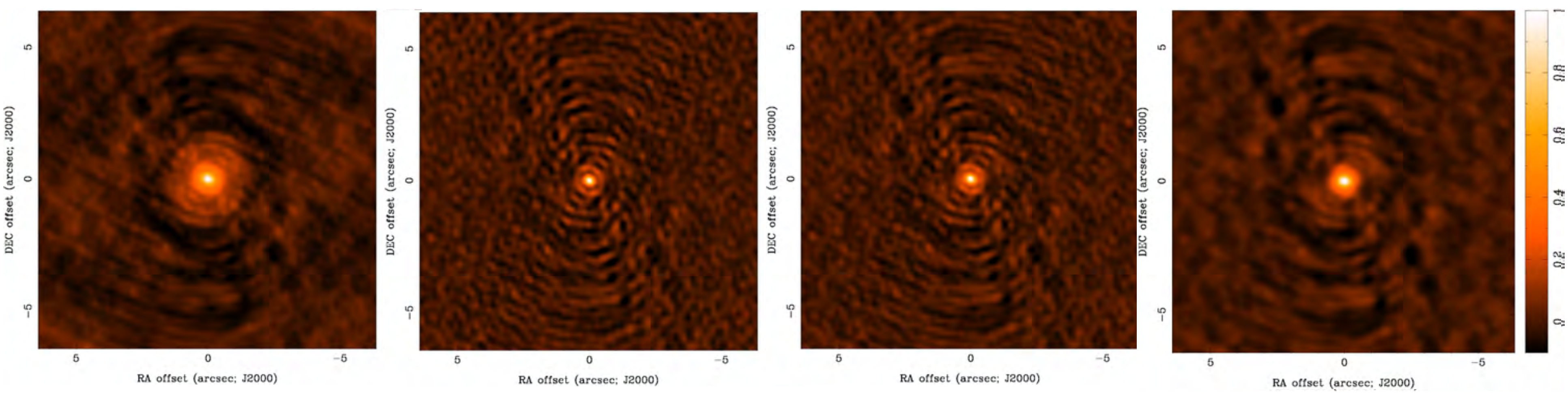
Gaussian fit to central core

1.5x1.5 arcsec

(natural weight + taper)

Weighting Schemes and Noise

- natural: equal weight to all visibilities → lowest noise
- uniform: equal weight for filled *(u,v)* cells → higher noise
- robust: continuous variation between natural and uniform
- taper: lower resolution but improved brightness temperature sensitivity



natural	uniform	robust=0	robust=0 + taper
0.59x0.50	0.35x0.30	0.40x0.34	0.59x0.50
ΔS = 1.0 mJy	ΔS = 2.1 mJy	ΔS = 1.3 mJy	ΔS = 1.2 mJy

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Visibility Weighting Scheme Summary

- imaging parameters provide a lot of freedom
- appropriate choices depend on science goals

natural	robust	uniform
lower angular resolution		highest angular resolution
best point source sensitivity		worse point source sensitivity
worse sidelobe structure		better sidelobe structure

+ tapering to
improve sensitivity to extended emission,
reduce dirty beam sidelobe structure,
match resolution from different observations,
… and other applications

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Beyond the Dirty Image

- to keep you awake at night…
  - ∃ an infinite number of $T(l,m)$ compatible with sampled $V(u,v)$
  - also noise → undetected and corrupted structure in $T(l.m)$

Deconvolution

use non-linear techniques to interpolate/extrapolate $V(u,v)$ samples into unsampled regions of (u,v) plane to find a *plausible* model of $T(l,m)$

- there is no unique prescription to extract an optimum model of $T(l,m)$
- requires *a priori* assumptions about $T(l,m)$ to pick a plausible “invisible” distributions that fill the unsampled parts of (u,v) plane
- key assumption: real sky does not look like dirty beam
  - not plausible: symmetric spokes, rings, negative regions, …

Deconvolution in Radio Astronomy

Two most common deconvolution algorithms

- clean (Högbom, J.A 1974, A&AS, 15, 417)
  - a priori assumption: *T(l,m)* can be represented by point sources
  - variants to improve computational efficiency, performance on extended structure
- maximum entropy (Gull, S.F. & Daniell, G.J 1978, Nature, 272, 686)
  - special case of forward modeling that minimizes an objective function that includes the data and a regularization term ("regularized maximum likelihood" methods)
  - a priori assumption for max entropy: *T(l,m)* is smooth and positive
  - vast literature about the deep meaning of entropy (Bayes theorem)

A very active research area! (see, e.g. EHT M87 imaging paper)

Clean Algorithm

- original version by Högbom is purely image based
  0. make dirty image and dirty beam (2x larger than the dirty image)
  1. find peak of dirty image
  2. subtract dirty beam centered at peak location, scaled by peak value and a loop gain factor (typically ~ 0.1) to make residual dirty image
  3. add subtracted peak to sky model (clean component list or image)
  4. if residual dirty image peak > stopping criterion, then goto step 1.

- Clark clean
  - use small patch of beam to improve speed, subtract clean components from gridded visibilities at once using FFT
- Cotton-Schwab clean
  - like Clark clean, but subtract clean components from ungridded visibilities and repeat entire image process to create residuals
  - avoids worst pixelization effects

Clean Algorithm Parameters

- stopping criterion
  - peak < threshold = multiple of theoretical rms noise
  - peak < threshold = fraction of dirty map maximum (useful if strong sources prevent a sensible noise threshold from being reached)
  - maximum number of clean components reached (no justification)
- loop gain parameter
  - values ~ 0.1 - 0.3 typically give good results
  - lower values can help with recovering smoother structures
- finite support
  - easy to include *a priori* information about where to search for clean components in the dirty map (image "masks" or clean "boxes")
  - useful, often essential for best results, but potentially dangerous
  - can be an arduous manual process; automatic algorithms OK

Clean: The Restored Image

- last step is to create a final "restored" image
  - make model image with all point source clean components
  - convolve point source model image with a "clean beam", an elliptical Gaussian fit to the main lobe of the dirty beam
  - add back residual dirty image with noise and structure below threshold
- the "restored" image is an estimate of the true sky brightness *T(l,m)*
- units of the "restored" image are (mostly) Jy per clean beam area

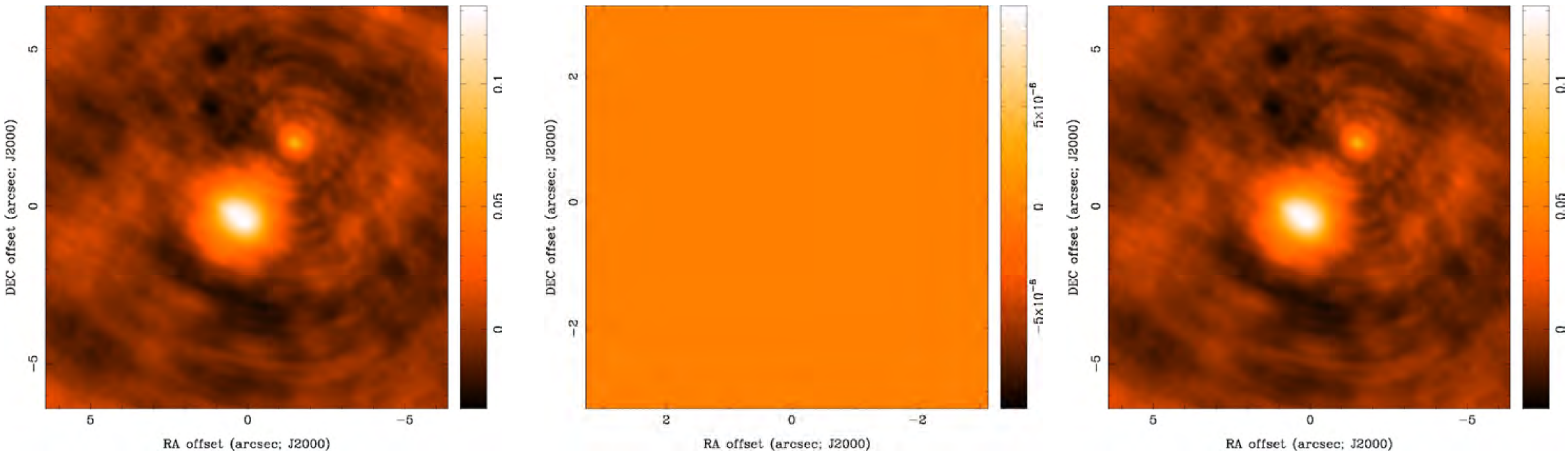
  = intensity, or brightness temperature

- Schwarz, U.J. 1978, A&A, 65, 345 proves that clean is statistically equivalent to a least squares fit of sine functions in the case of no noise

Clean Example

initialize

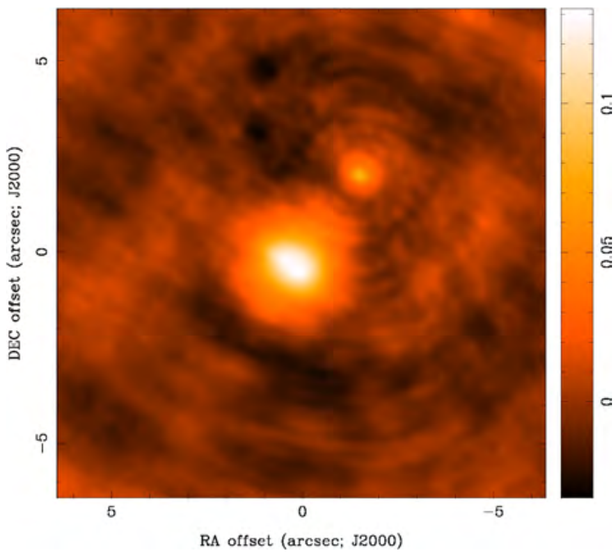
$T^D(l,m)$ | *0 clean components* | *residual map*



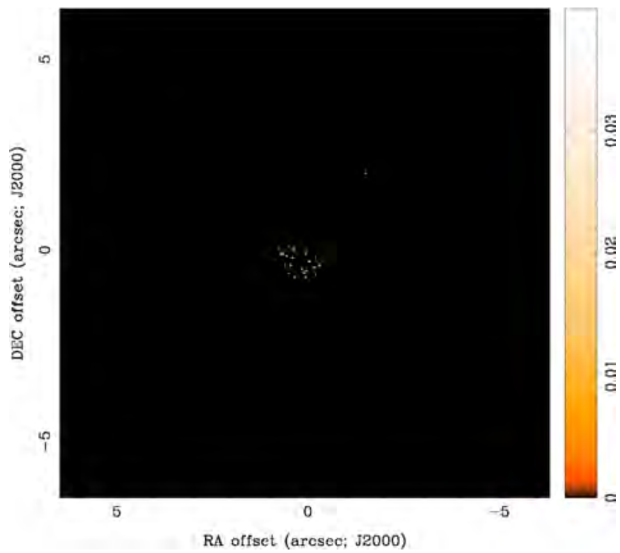
CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Clean Example

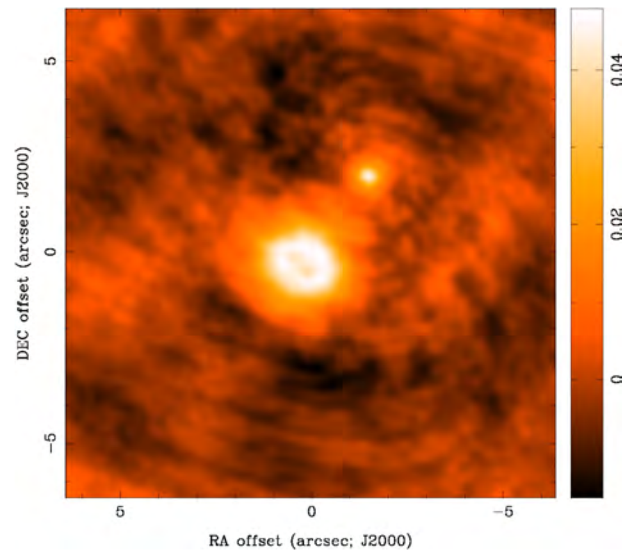
$T^D(l,m)$



30 clean components



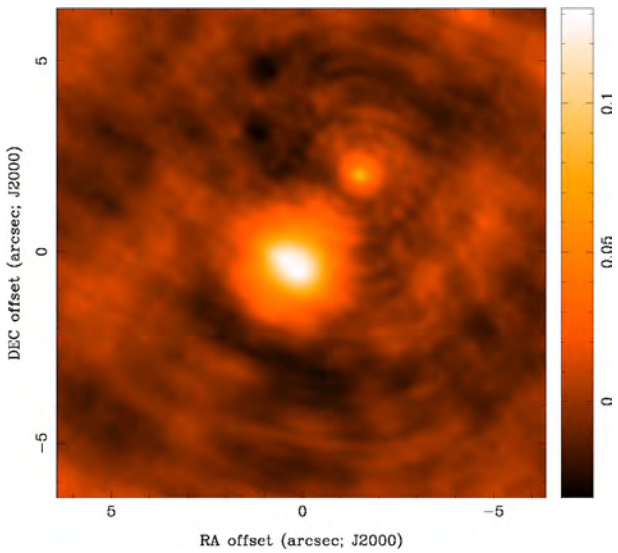
residual map



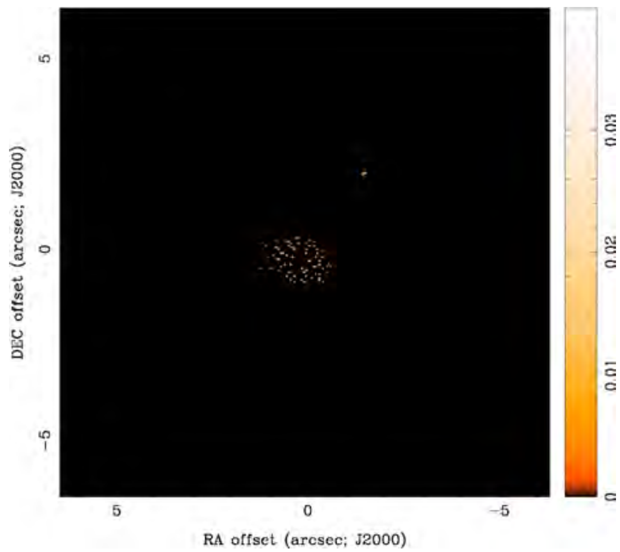
CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Clean Example

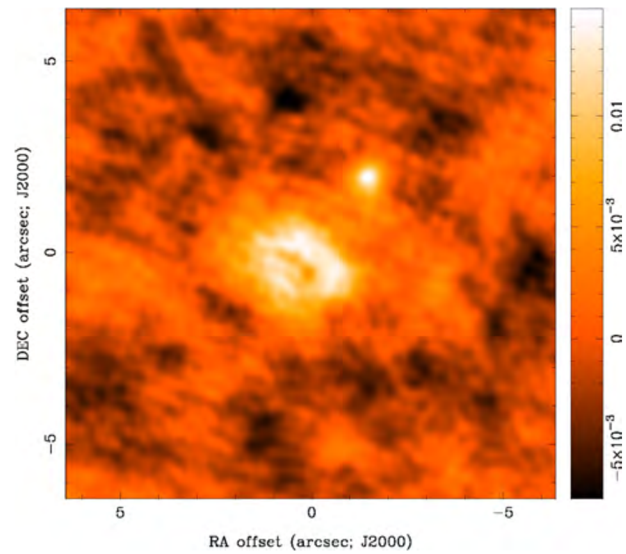
$T^D(l,m)$



100 clean components

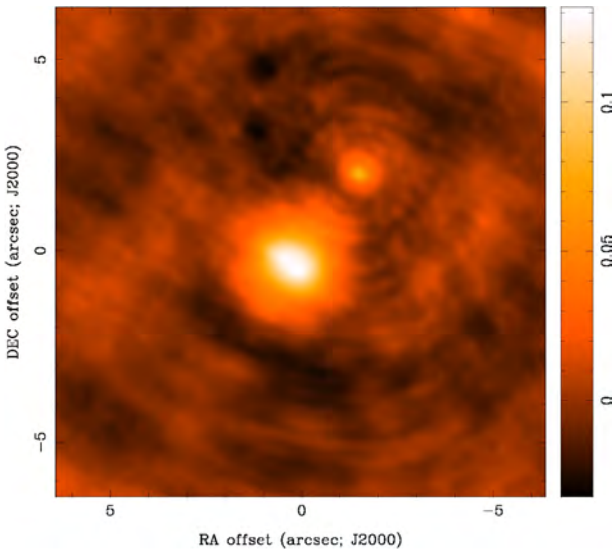


residual map

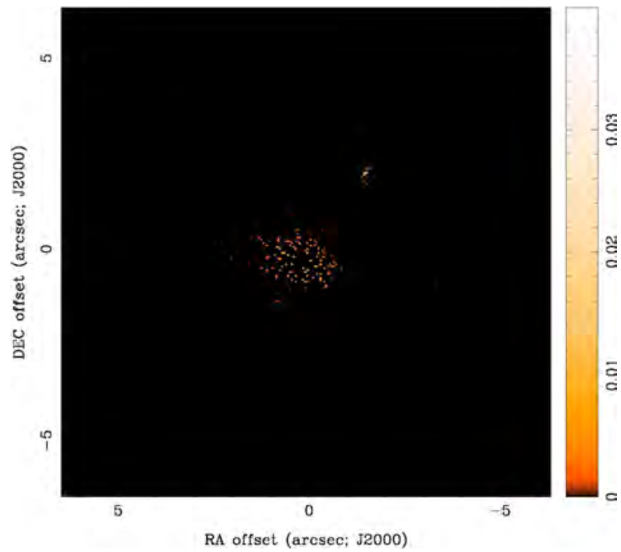


Clean Example

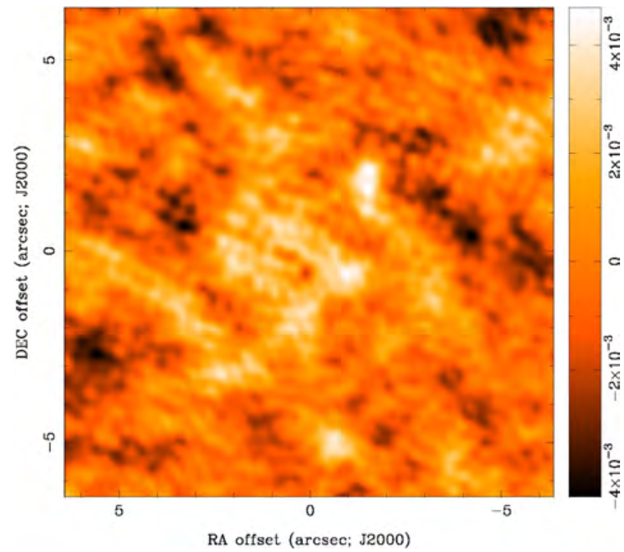
$T^D(l,m)$



300 clean components

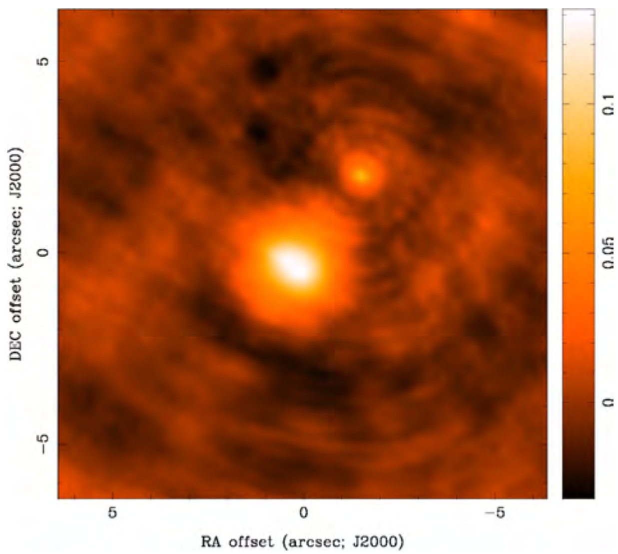


residual map

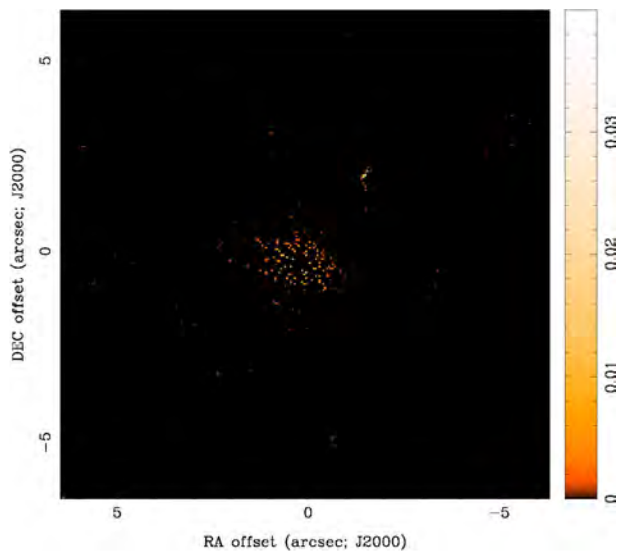


Clean Example

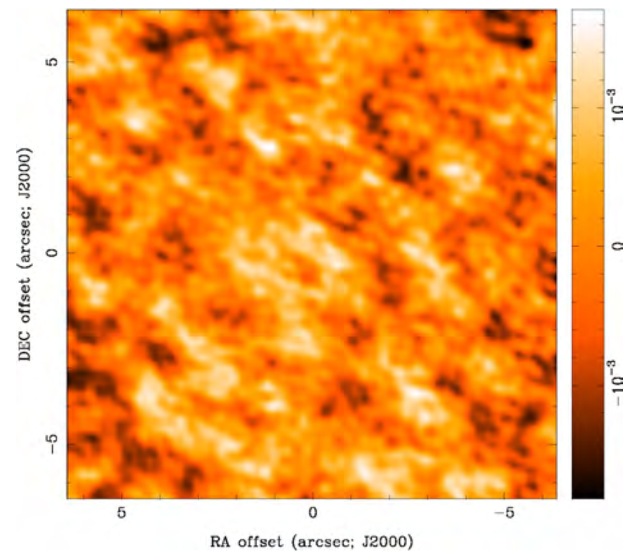
$T^D(l,m)$



585 clean components

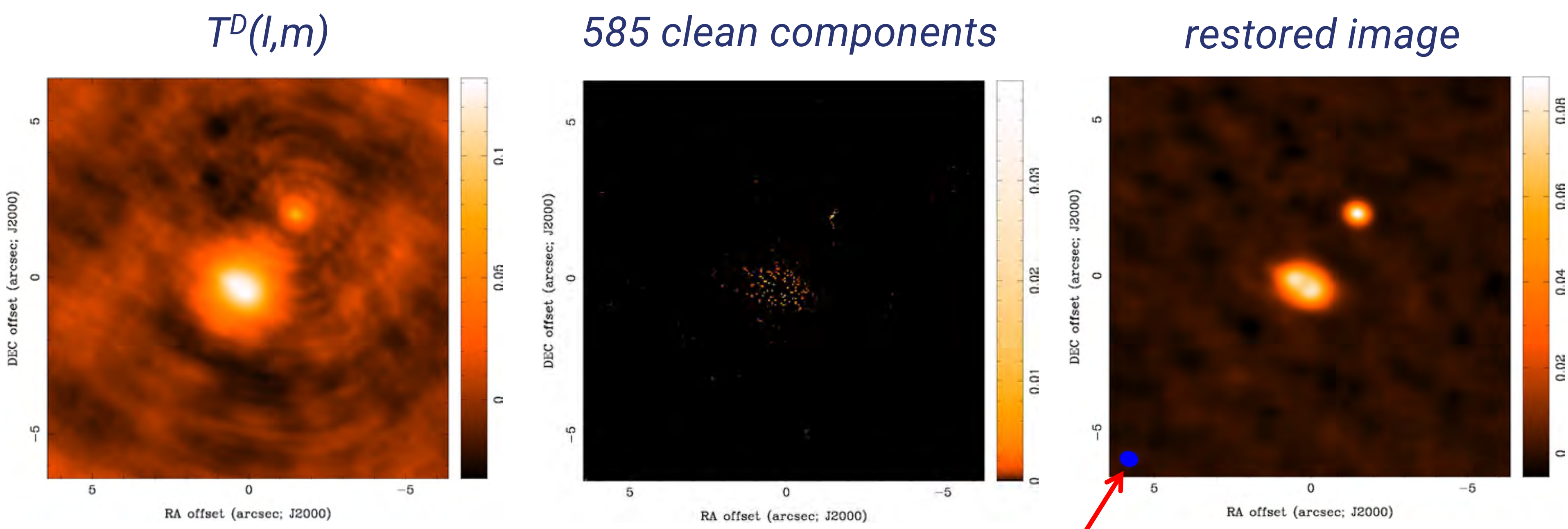


residual map



threshold reached

Clean Example: Restored Image



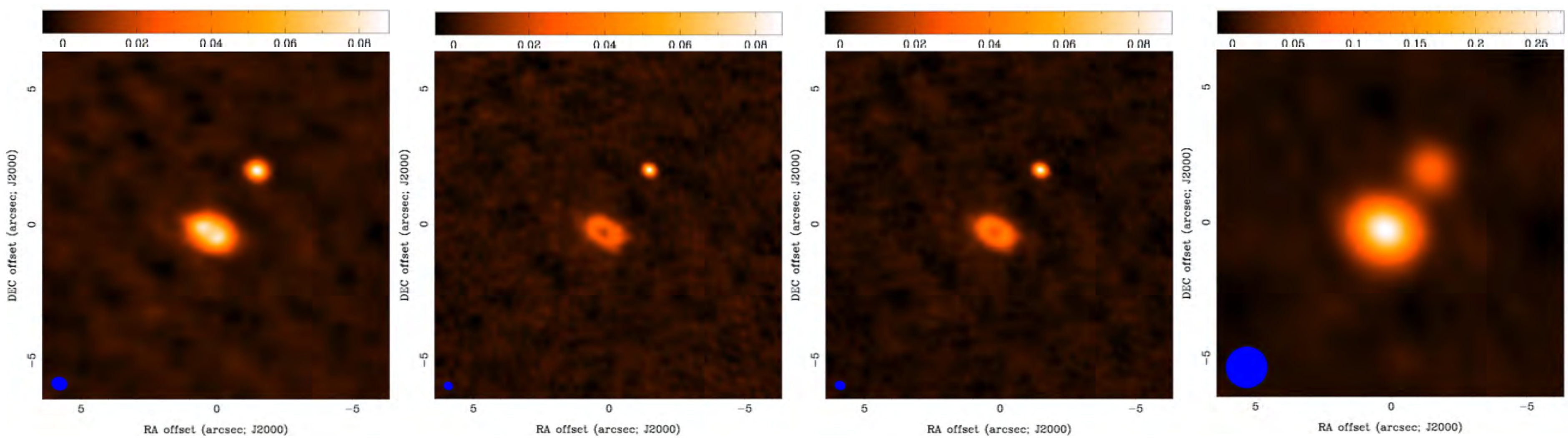
ellipse = restoring beam fwhm

final image depends on

- imaging parameters: pixel size, visibility weighting scheme, gridding, …)
- deconvolution: algorithm, iterations, stopping criterion, …)

Clean: Different Weighting Schemes

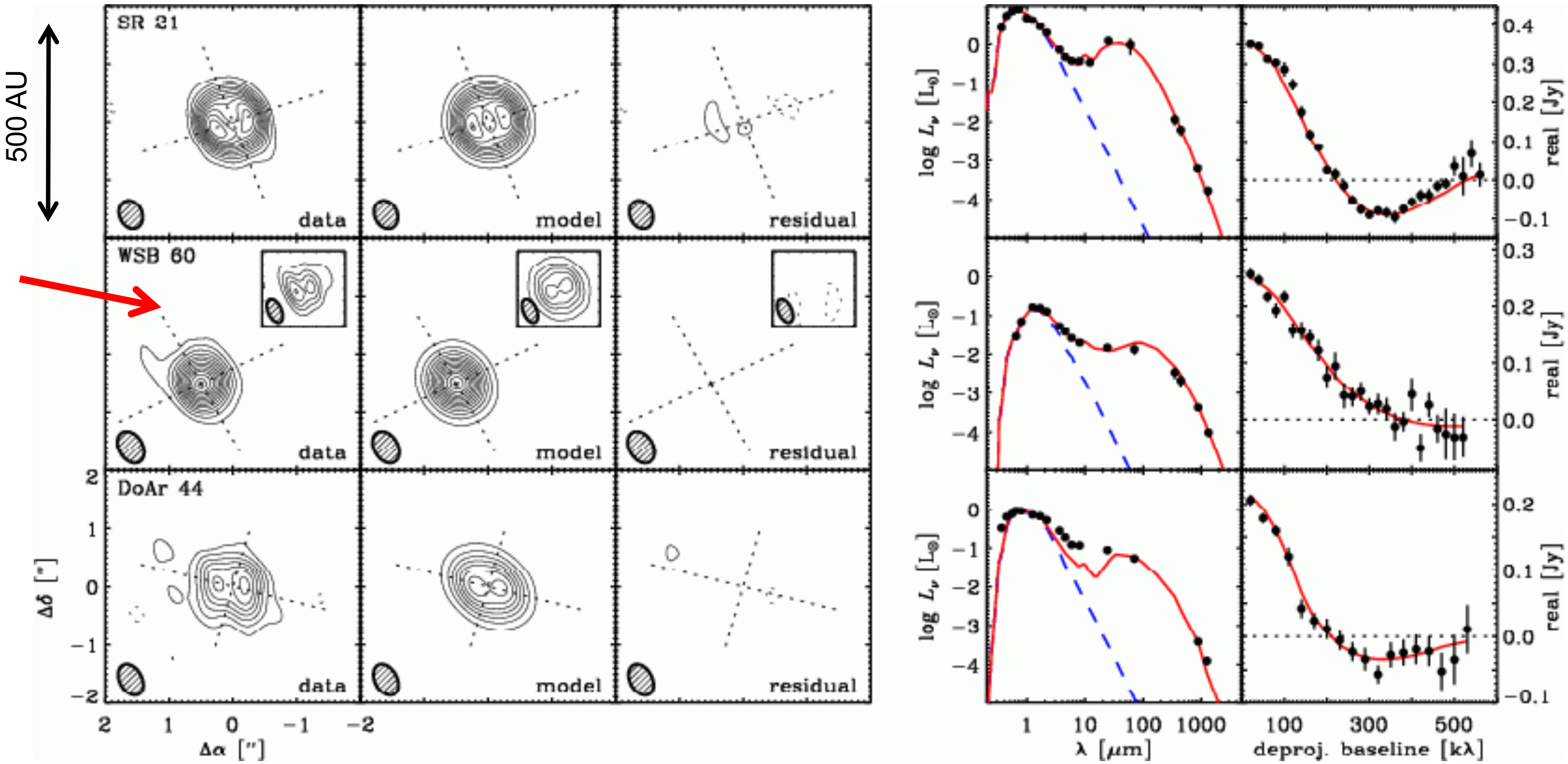
- emphasize different angular scales from the *(u,v)* data in the image



natural	uniform	robust=0	robust=0 + taper
0.59x0.50	0.35x0.30	0.40x0.34	1.5x1.5
ΔS = 1.0 mJy	ΔS = 2.1 mJy	ΔS = 1.3 mJy	ΔS = 1.4 mJy

Tune Imaging Parameters to Science

- example: SMA 870 μm images of protoplanetary disk dust continuum emission with resolved inner cavities (Andrews et al. 2009, ApJ, 700, 1502)



Scale Sensitive Methods

- standard clean often works poorly on very extended structure
  - and many point source components needed (slow)
- adjacent pixels in image are not independent
  - oversampled resolution limit
  - intrinsic source size: an extended source covering 1000 pixels might be better characterized by a few parameters than by 1000 parameters, e.g. 6 parameters for a Gaussian
- scale sensitive deconvolution algorithms employ fewer degrees of freedom in solution to model plausible sky brightness distributions
- e.g. multi-scale clean
  - make a collection of dirty images at different resolutions via convolution with input “scales” (e.g. 0,3,10,30 pixels)
  - find peak across all scales, remove fraction of peak at that scale from all dirty images, add corresponding blob to model, iterate..

Maximum Entropy Algorithm

- find the least biased image that agrees with *V(u,v)* samples
- mathematically, minimize $J = H - \alpha\chi^2$

  where *H* is the image "entropy"

$$H = -\sum_k T_k \log\left(\frac{T_k}{M_k}\right)$$

requires positivity

α is a Lagrange multiplier

χ^2 is a measure of agreement with the data

$$\chi^2 = \sum_k \frac{|V(u_k,v_k) - \mathrm{FT}\{T\}|^2}{\sigma_k^2}$$

and M is the "default image"

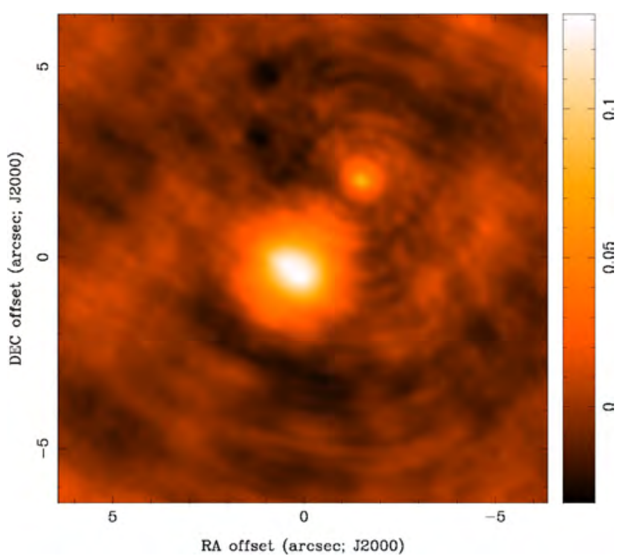
- can be solved using numerical techniques based on conjugate gradients, a fast (N logN) solver introduced by Cornwell & Evans (1983)

Maximum Entropy Algorithm

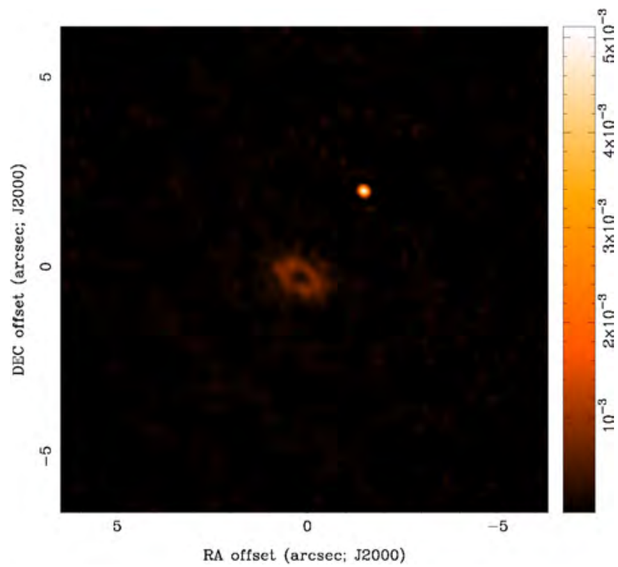
- in the absence of information in *V(u,v)*, returns the default image
- easy to include *a priori* information using the “default image” construct
  - If nothing known, then a flat default image is a good choice
  - a single dish image, if available, might make a good default image
- straightforward to generalize χ^2 to combine different observations
- alternative measures of entropy available
  - replace log with cosh for “emptiness” (does not require positivity)
- relatively fast on large images
- does not use direct information about dirty beam shape
  - can have trouble dealing with sidelobes of strong point sources
- effective angular resolution in image is signal-to-noise dependent
- for SMA, often more difficult to drive than clean

Maximum Entropy Example

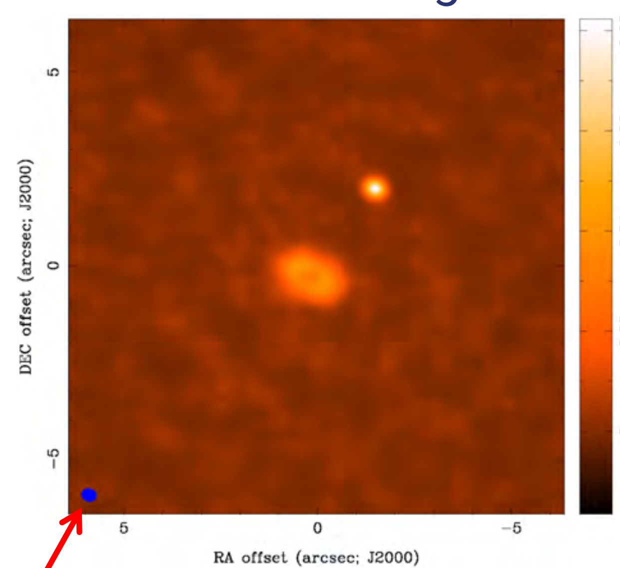
$T^D(l,m)$



maximum entropy model



maximum entropy restored image



ellipse = restoring beam fwhm

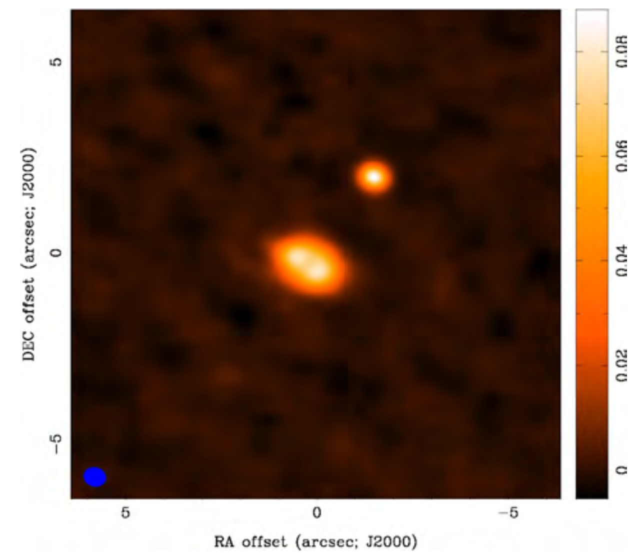
Measures of Image Quality

dynamic range

- ratio of peak brightness in image to rms noise in region devoid of emission
- easy way to calculate a *lower limit* to the error in brightness in a non-empty region

e.g. peak 88 and rms 1.0 mJy/beam

→ dynamic range = 88



fidelity

- difference between any reconstructed image and the correct image
- fidelity = input model/difference = inverse of the relative error

  = model * beam / abs (model * beam − reconstructed image)
- generally much lower than implied by dynamic range

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Invisible Large Scale Structure

- inevitable central hole in *(u,v)* plane coverage
- extended structure may be missed, attenuated, or distorted
- to estimate if the lack of short baselines is a problem for science
  - simulate the observations using a model of the source
  - check simple expressions for a Gaussian or uniform disk

Homework Problem

- Q: by what factor is the central brightness reduced as a function of source size due to missing short spacings for a Gaussian with fwhm $\theta_{1/2}$?
- A: a Gaussian source central brightness is reduced 50% when

$$\theta_{1/2} = 18'' \left(\frac{\nu}{100\ GHz}\right)^{-1} \left(\frac{B_{min}}{15\ meters}\right)^{-1}$$

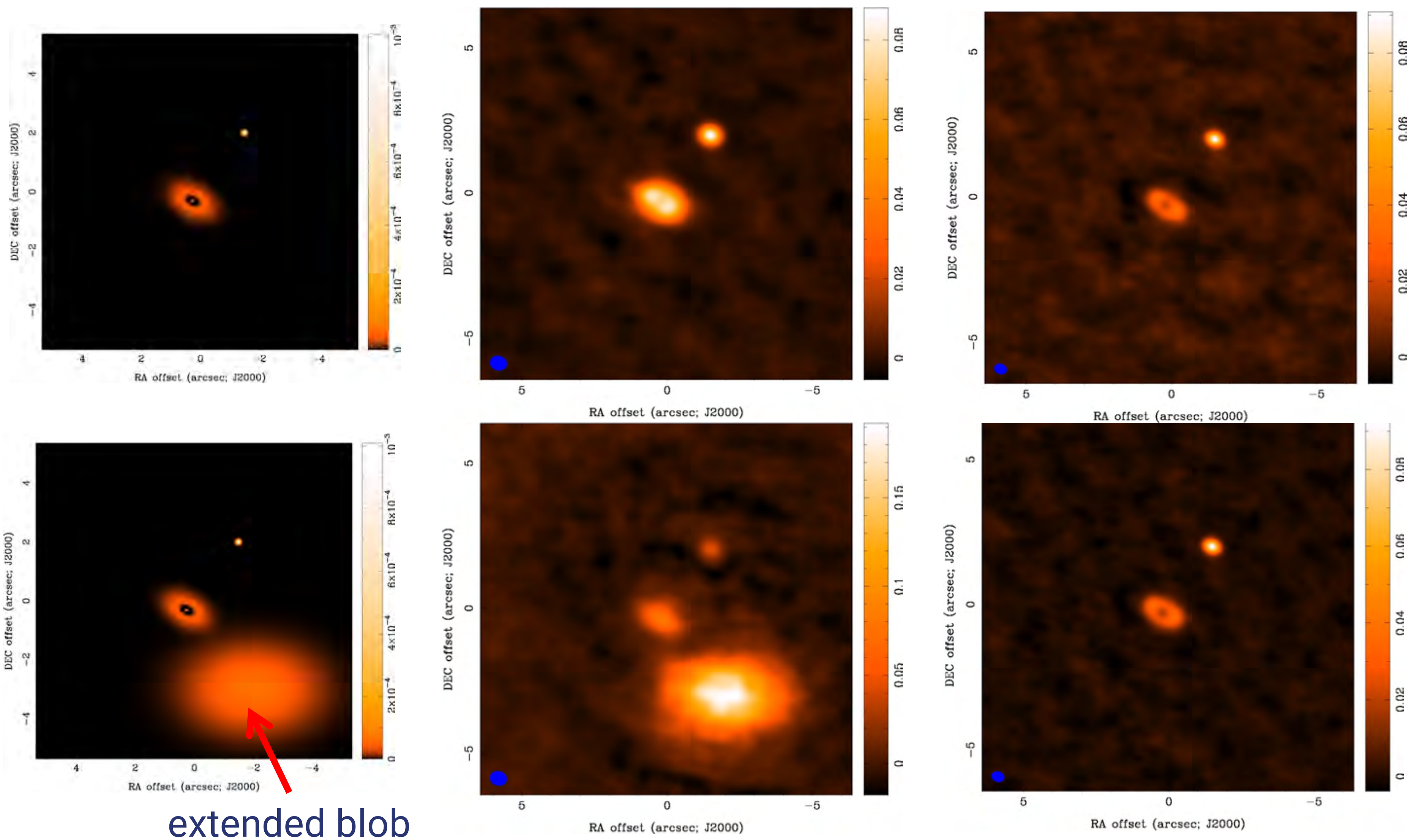
where B_{min} is the shortest baseline and ν is the frequency

(see derivation in appendix of Wilner & Welch 1994, ApJ, 427, 898)

Missing Short Baselines Example

natural weight

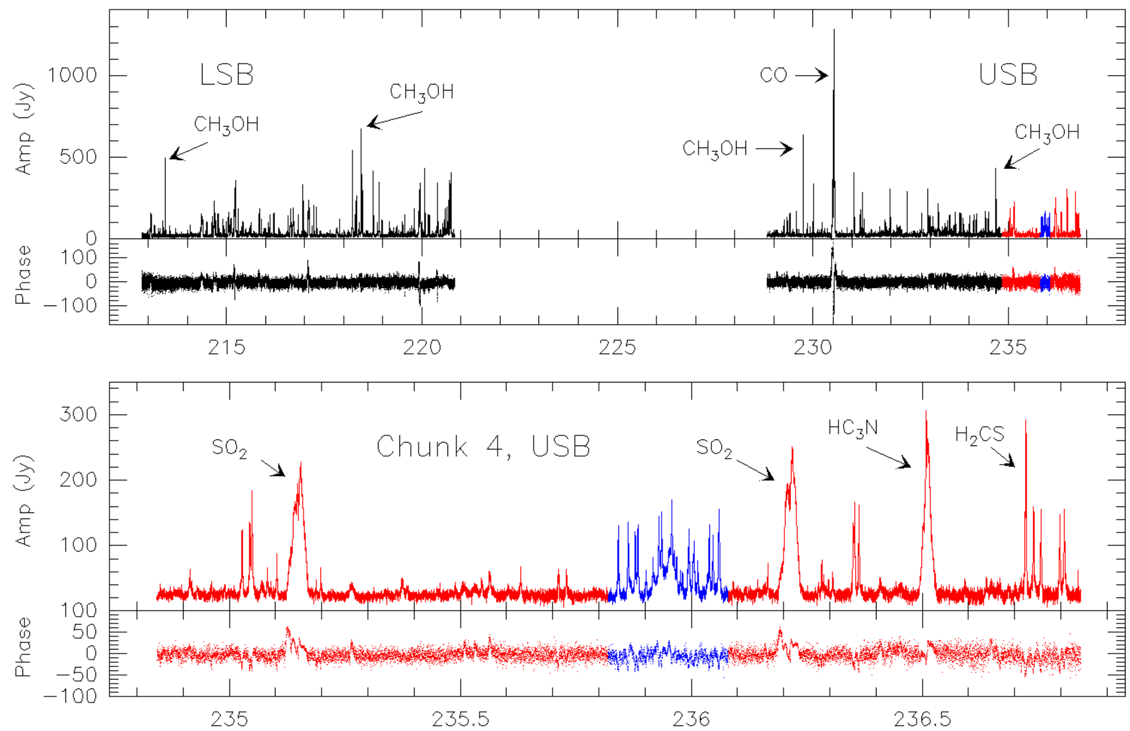
>75 kλ natural weight



extended blob
(10x disk flux)

Spectral Line Considerations: Science

- most of discussion applies to a single spectral channel of width $\delta\nu$
- science may require many such channels across total bandwidth $\Delta\nu$
- emission/absorption lines from molecular or atomic transitions
- a significant continuum spectral slope or curvature



CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Spectral Line Considerations: Technical

- **technical** reasons to divide $\Delta\nu$ into many channels
- avoid "bandwidth smearing"

$$V\left(\frac{\nu_0}{\nu}u, \frac{\nu_0}{\nu}v\right) = \left(\frac{\nu}{\nu_0}\right)^2 T\left(\frac{\nu}{\nu_0}l, \frac{\nu}{\nu_0}m\right)$$

average over $\Delta\nu$ results in radial smearing in image plane

$$\sim \frac{\Delta\nu}{\nu_0}\sqrt{l^2+m^2}$$

$\rightarrow$ constraint on field of view: $\sqrt{l^2+m^2} < 0.1\frac{\theta_s\nu_0}{\Delta\nu}$

- also, edit out narrow band interference

Concluding Remarks

- interferometry samples Fourier components of sky brightness
- make an image by Fourier transforming sampled visibilities
- deconvolution attempts to correct for incomplete sampling
- remember
    - there are an infinite number of images compatible with the visibilities
    - missing (or corrupted) visibilities affect the entire image
    - astronomers must make decisions in imaging and deconvolution
- it is fun and worth the trouble $\rightarrow$ high resolution images!

many, many imaging issues not covered in this talk, see references

END