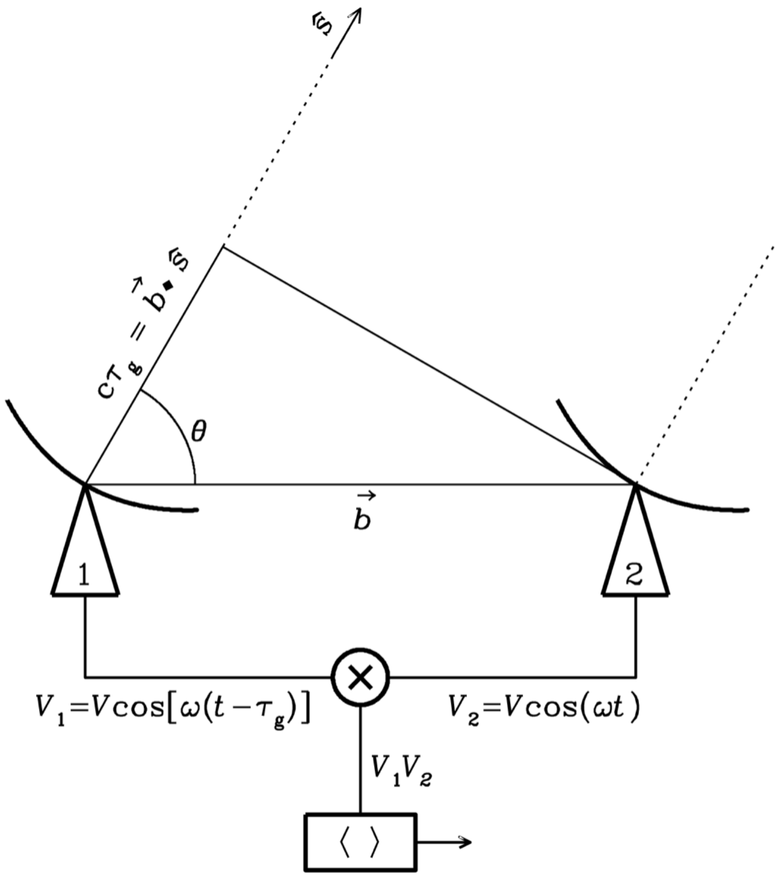
Essentials of
Radio and (Sub-)Millimeter Astronomy

Fabian Walter (MPIA)

Simple Interferometer



note: geometric delay removed through electronics
we are interested in delays due to different positions on sky

$$\tau_g = \frac{\vec{b} \cdot \hat{s}}{c}.$$

voltages:

$$V_1 = V\cos[\omega(t - \tau_g)]$$

$$V_2 = V\cos(\omega t).$$

multiplication: $$V_1 V_2 = V^2 \cos[\omega(t-\tau_g)]\cos(\omega t) = \left(\frac{V^2}{2}\right)[\cos(2\omega t - \omega\tau_g) + \cos(\omega\tau_g)]$$

$$\cos x \cos y = [\cos(x+y) + \cos(x-y)]/2$$

time averaging: $$R = \langle V_1 V_2 \rangle = \left(\frac{V^2}{2}\right)\cos(\omega\tau_g).$$

for $(\Delta t \gg (2\omega)^{-1})$

$\cos(2\omega t - \omega\tau_g)$-> 0

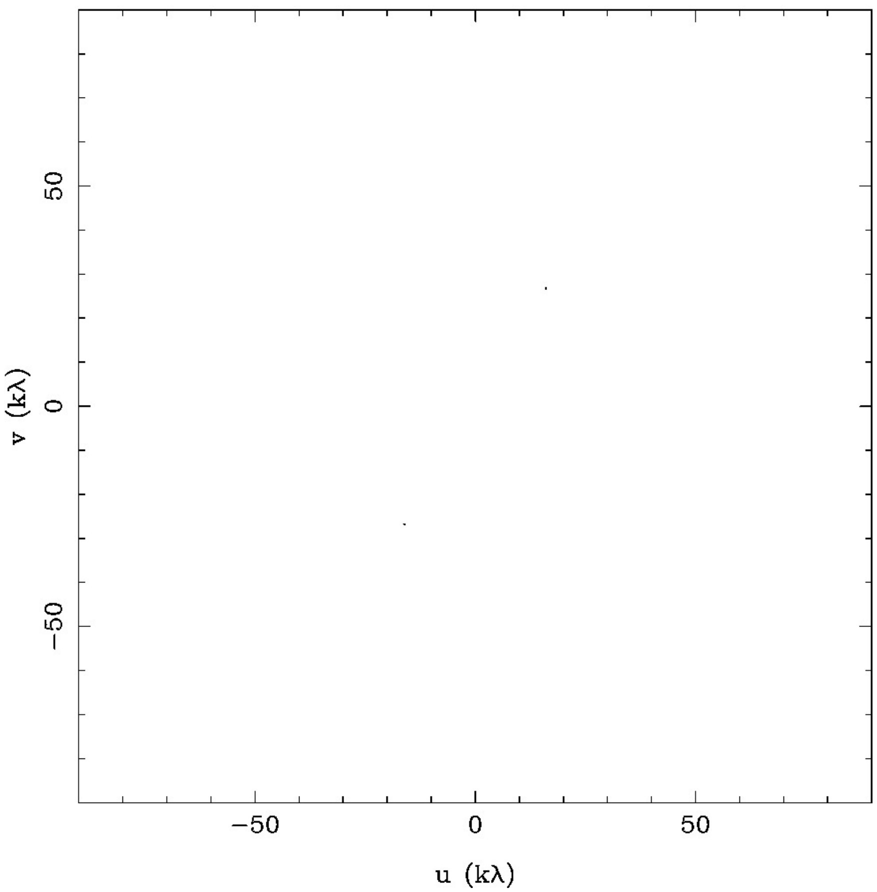
Lecture 7

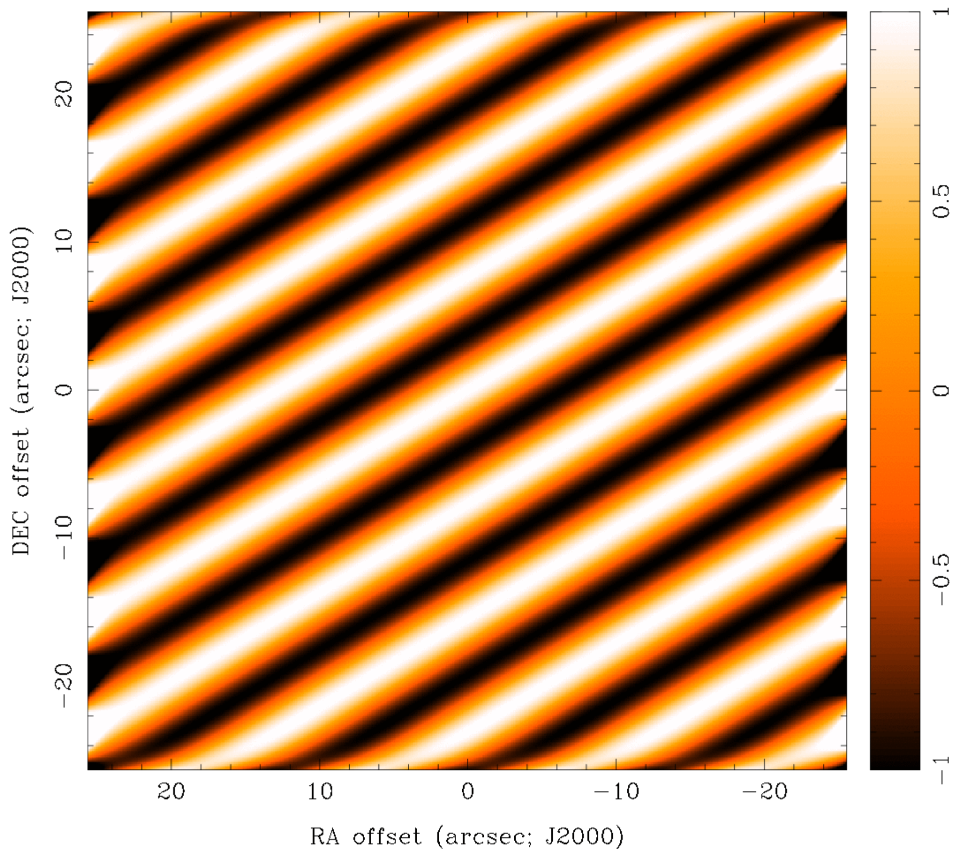
Basics of Interferometry: Imaging I

slides: Essential Radio Astronomy by NRAO (Condon & Ransom)
Dr. Michael Wise (ASTRON)
Prof. David Wilner (Harvard)
see also SMA summer school 2021: https://lweb.cfa.harvard.edu/sma-school/program/

Example SMA (u,v) Plane Sampling

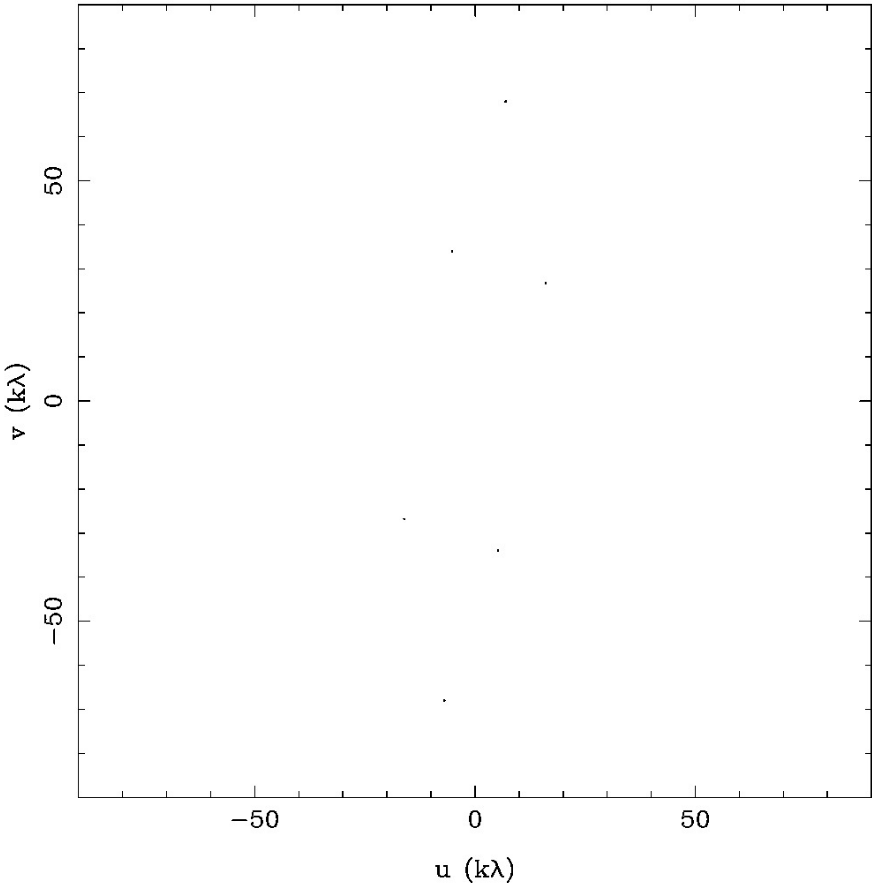
2 antennas, 1 x 30s sample

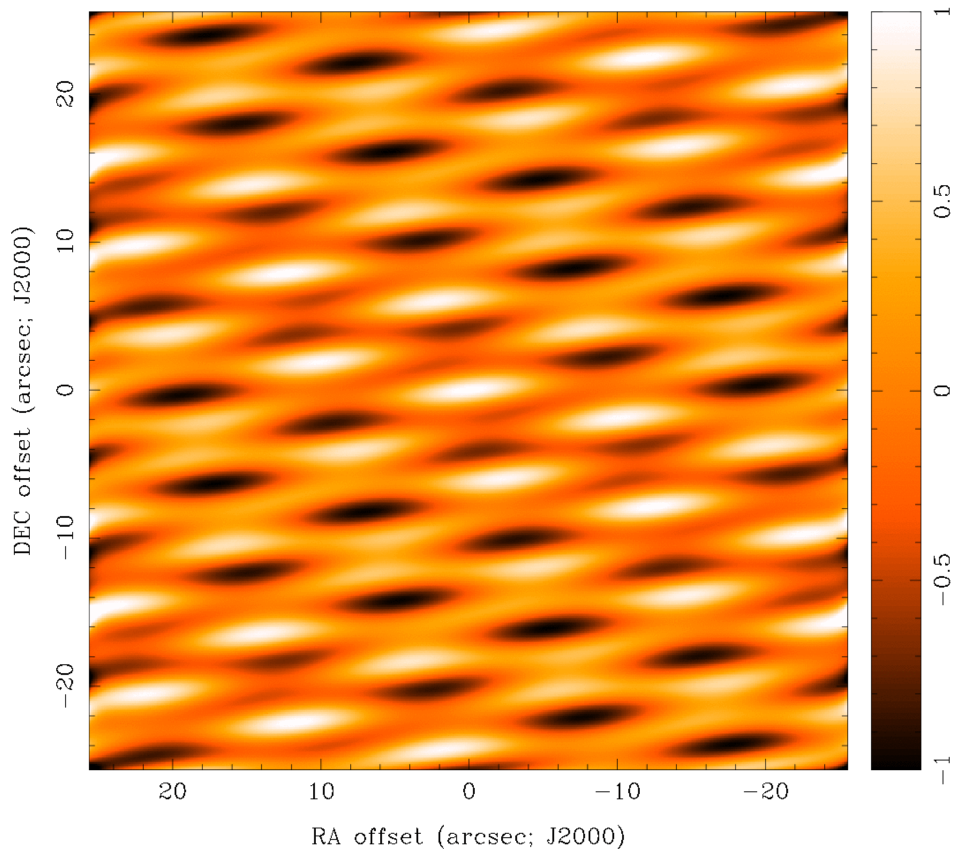




Example SMA (u,v) Plane Sampling

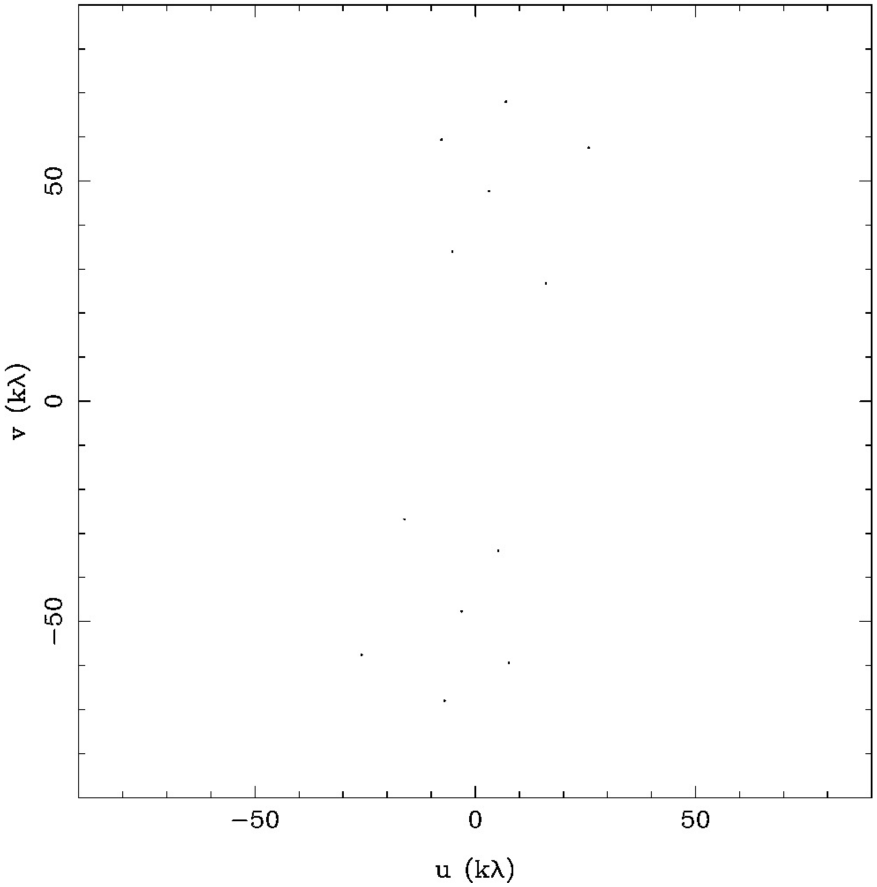
3 antennas, 1 x 30s sample

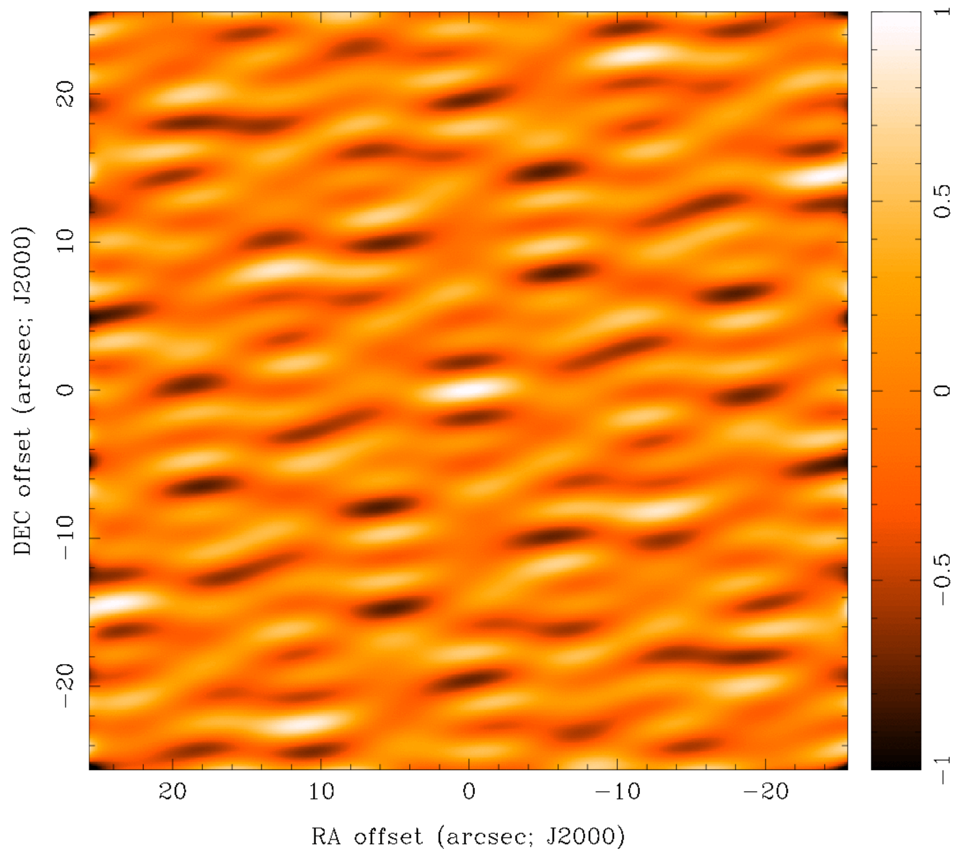




Example SMA (u,v) Plane Sampling

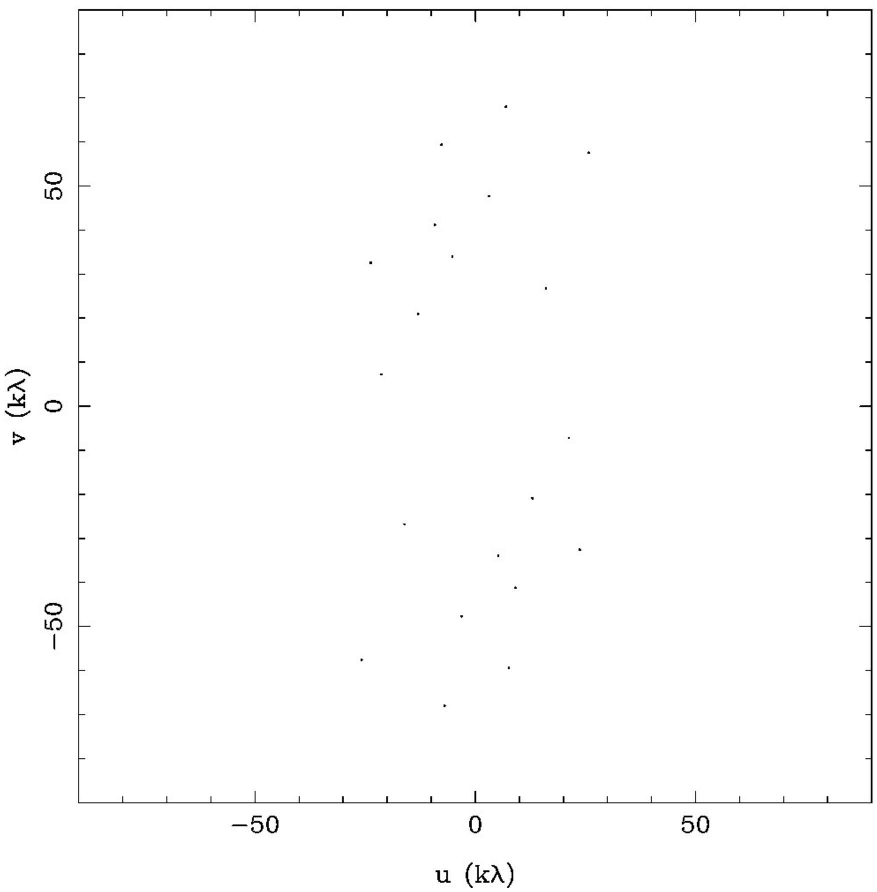
4 antennas, 1 x 30s sample

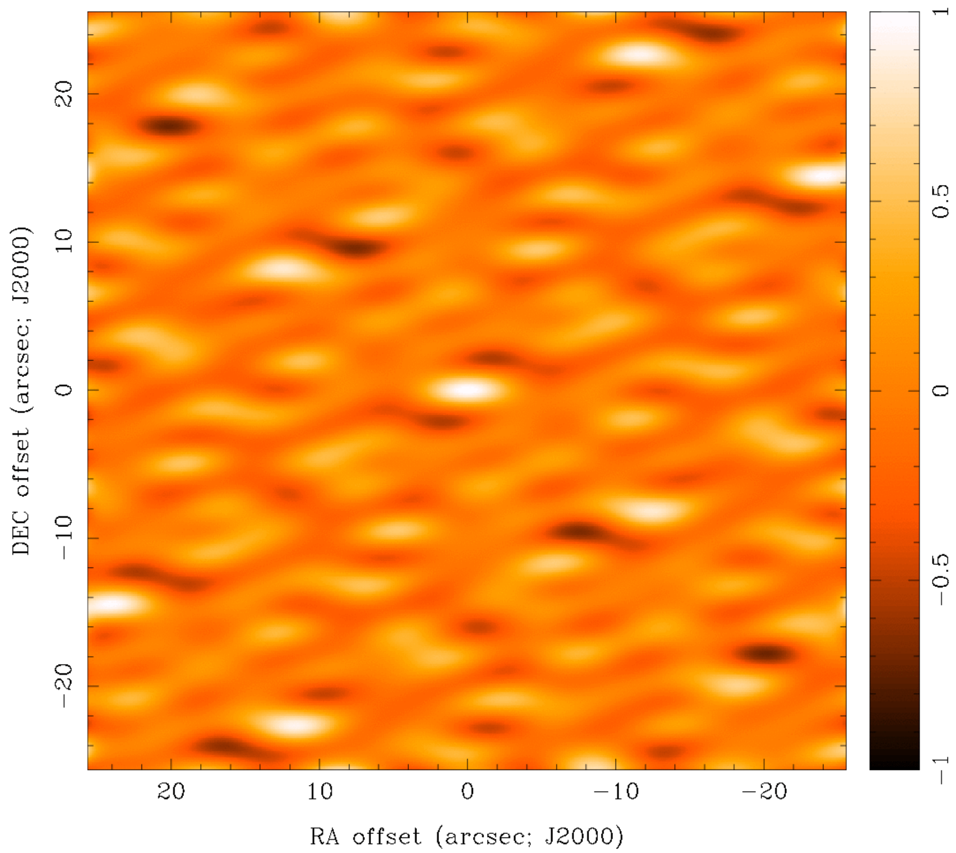




Example SMA (u,v) Plane Sampling

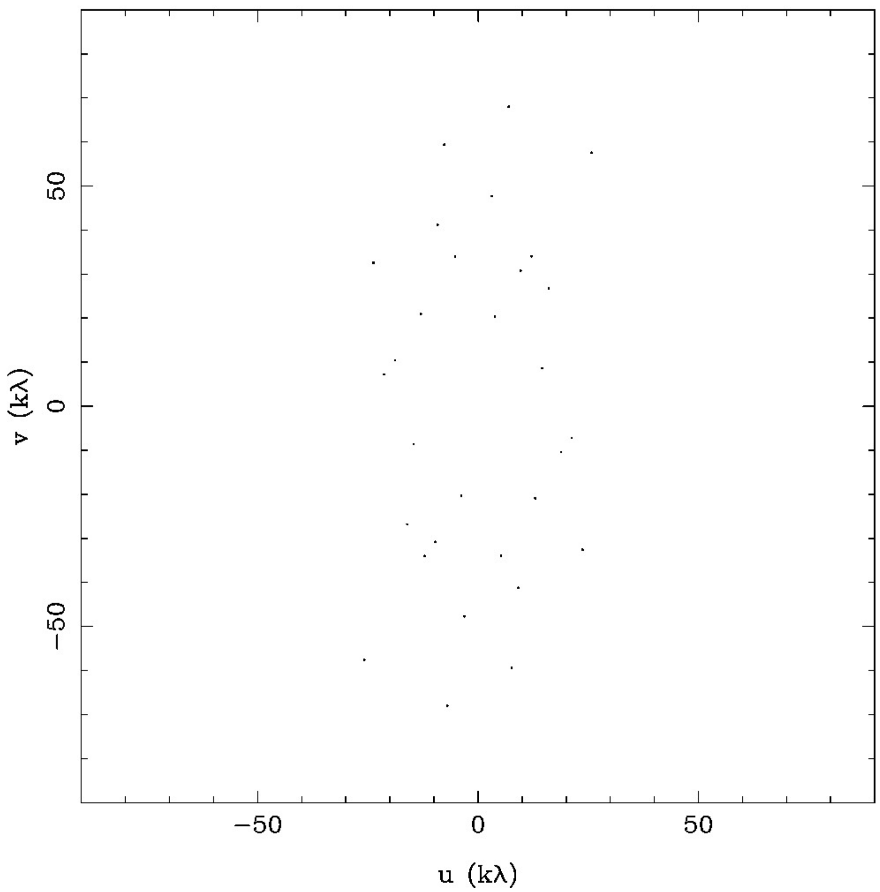
5 antennas, 1 x 30s sample

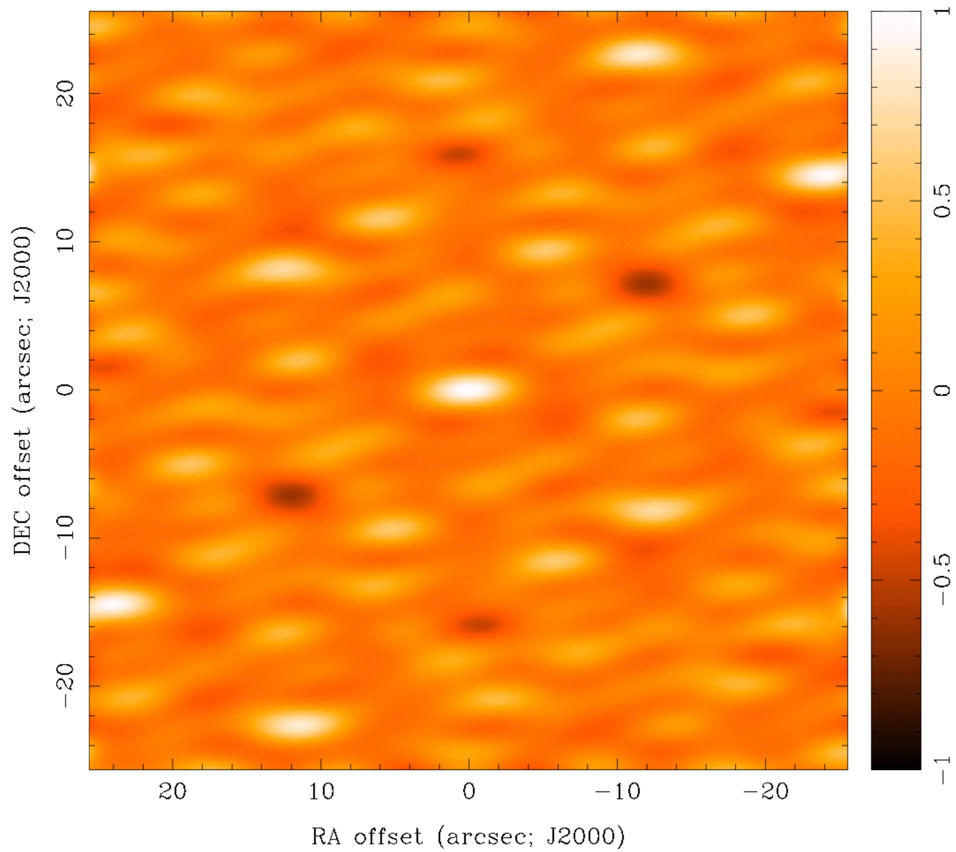




Example SMA (u,v) Plane Sampling

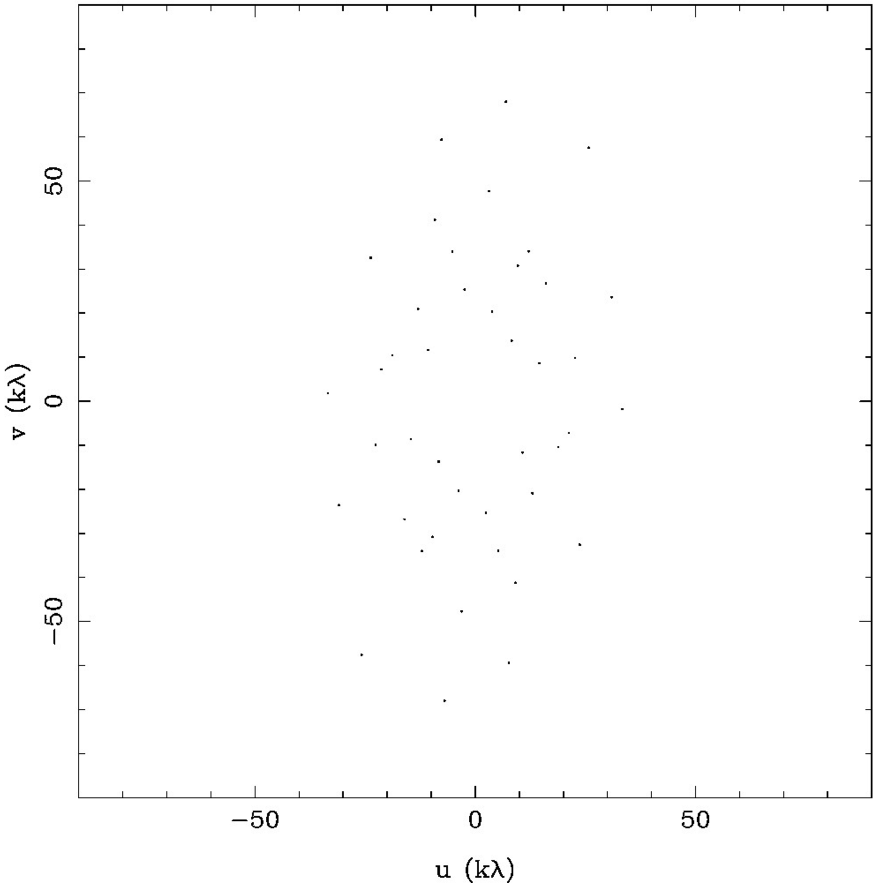
6 antennas, 1 x 30s sample

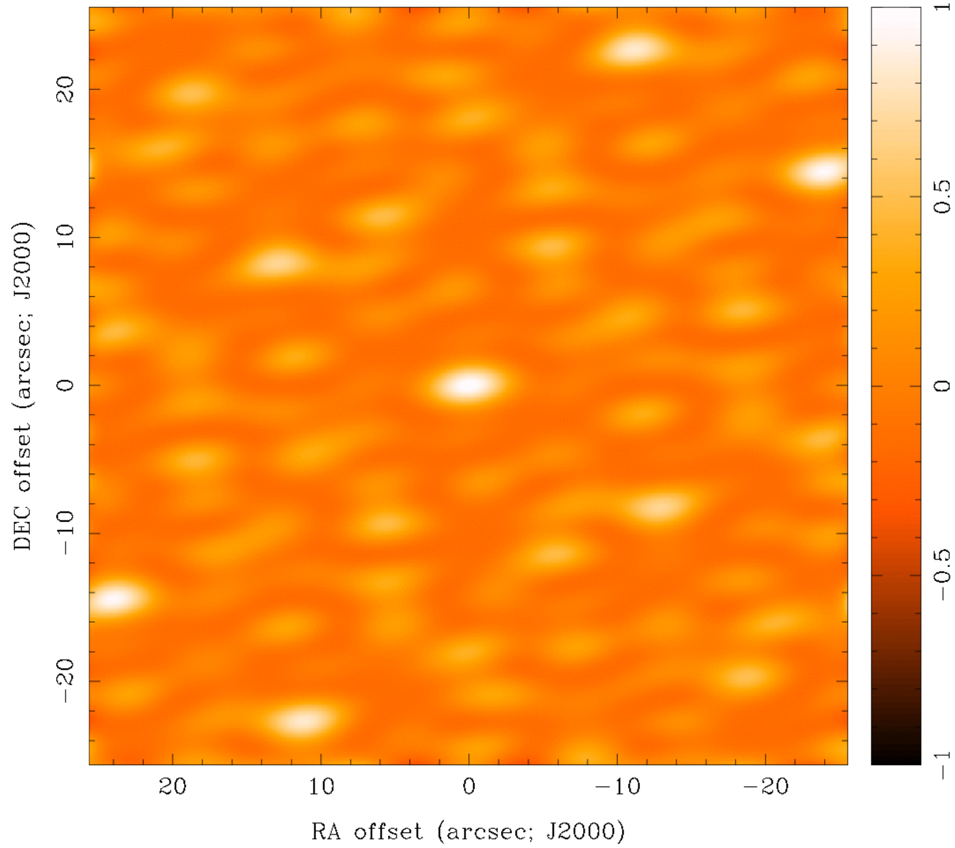




Example SMA (u,v) Plane Sampling

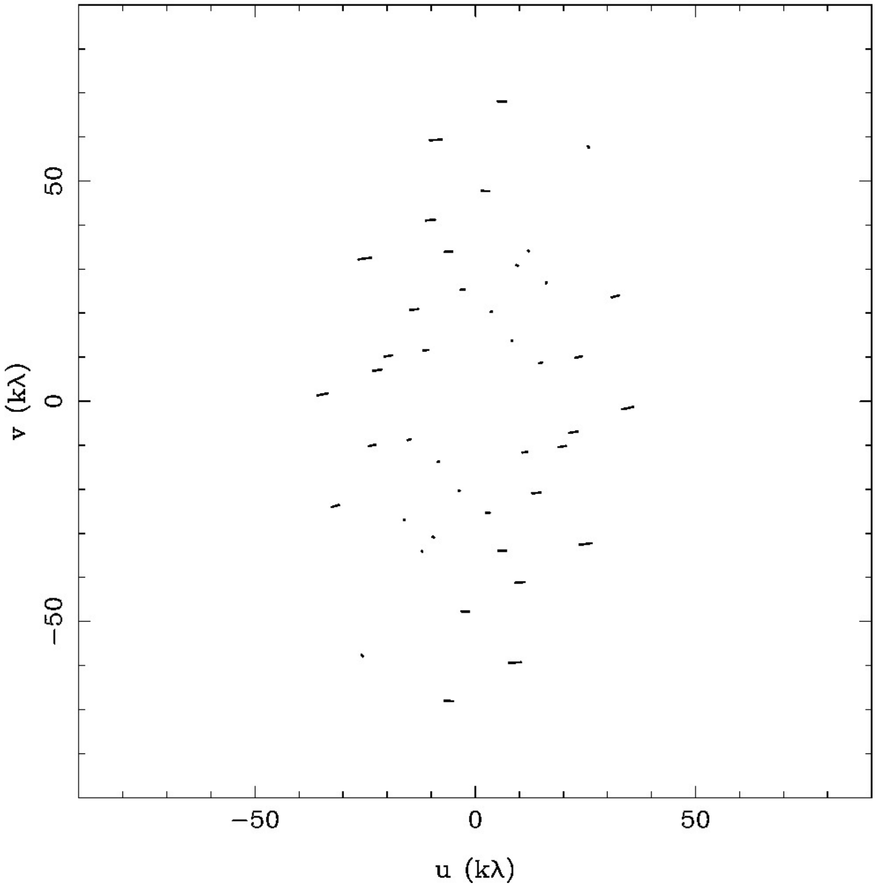
7 antennas, 1 x 30s sample

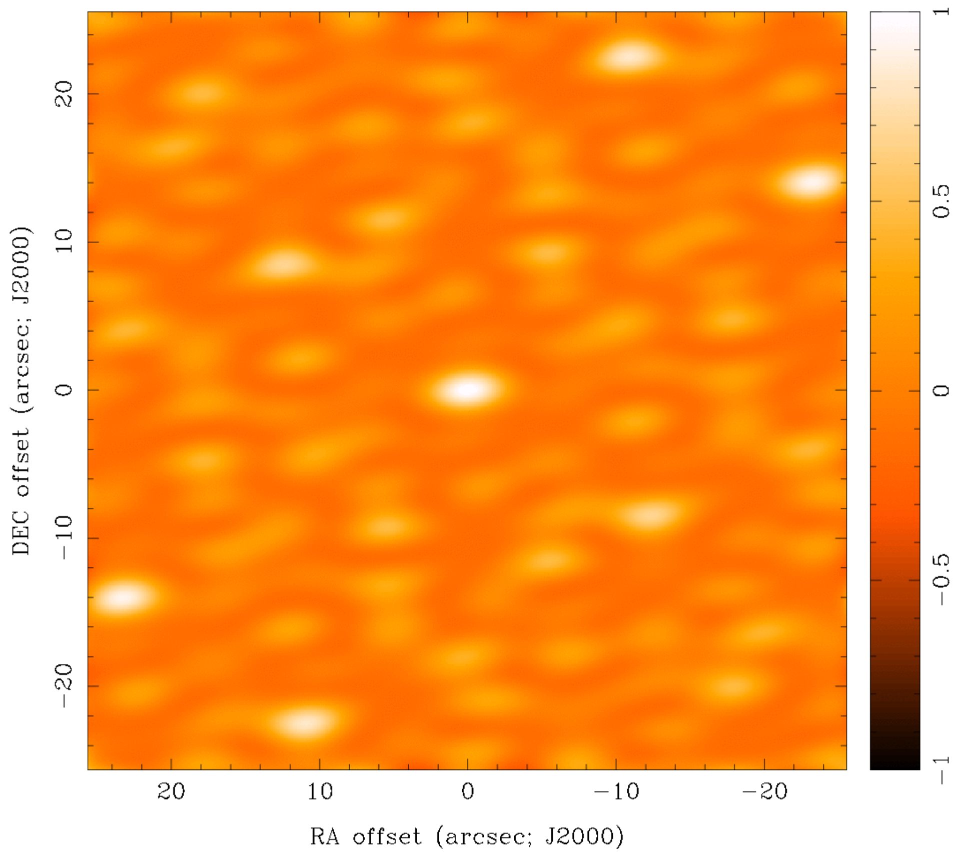




Example SMA (u,v) Plane Sampling

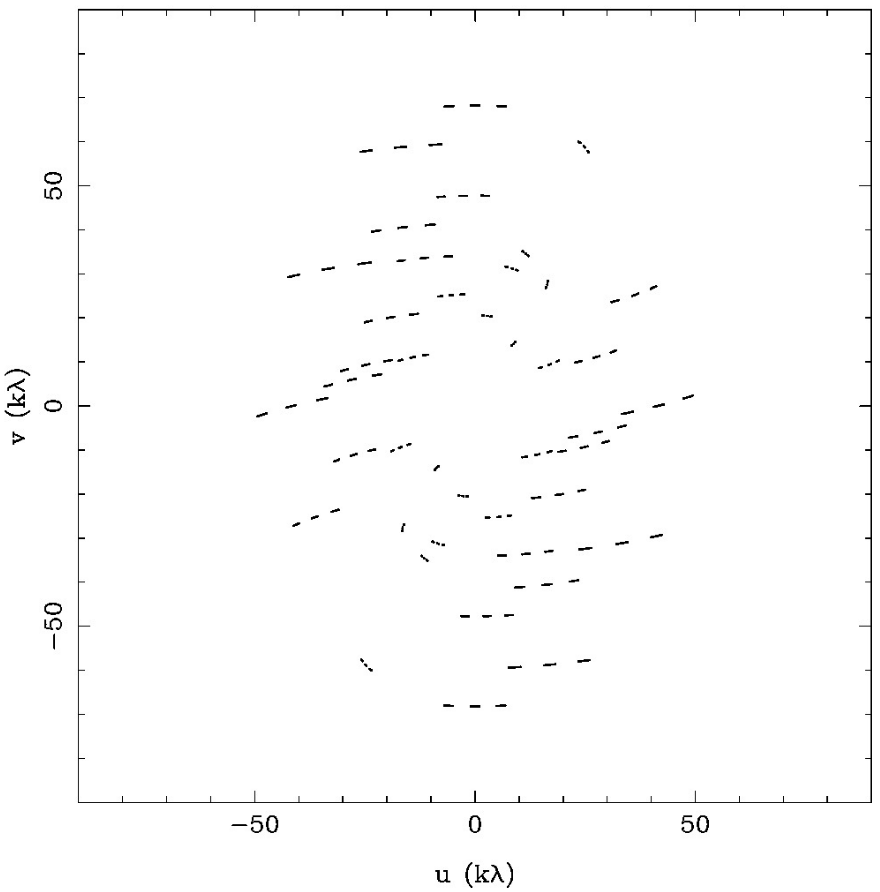
7 antennas, 10 x 30s samples

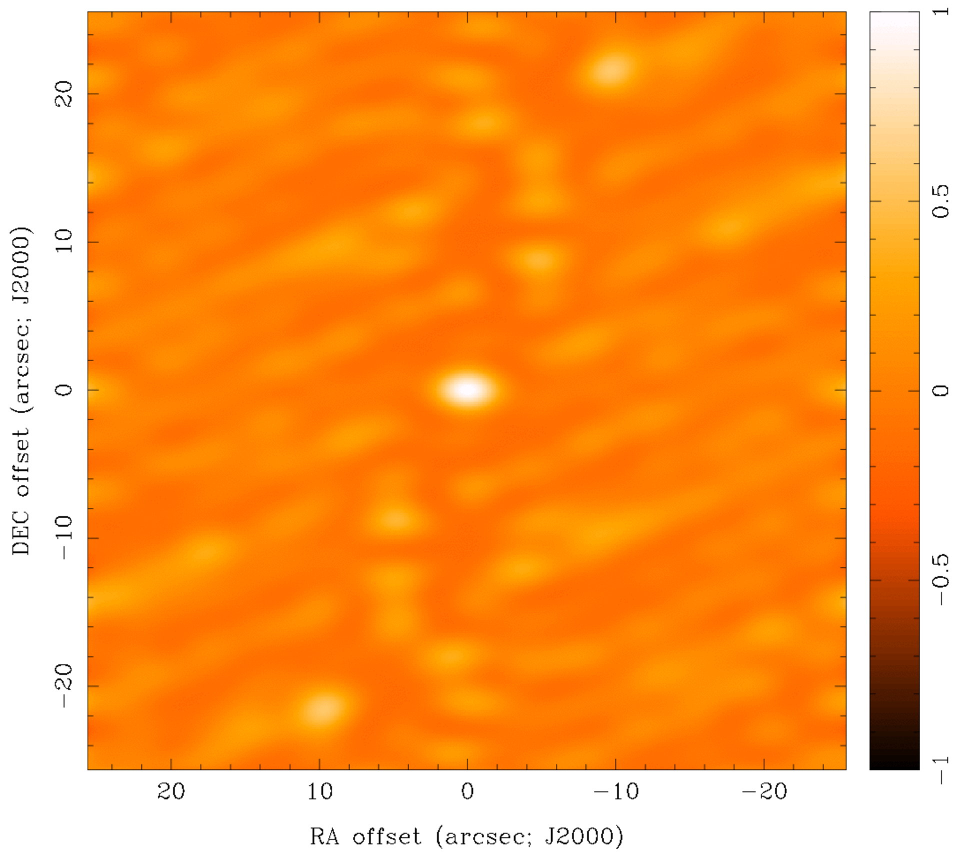




Example SMA (u,v) Plane Sampling

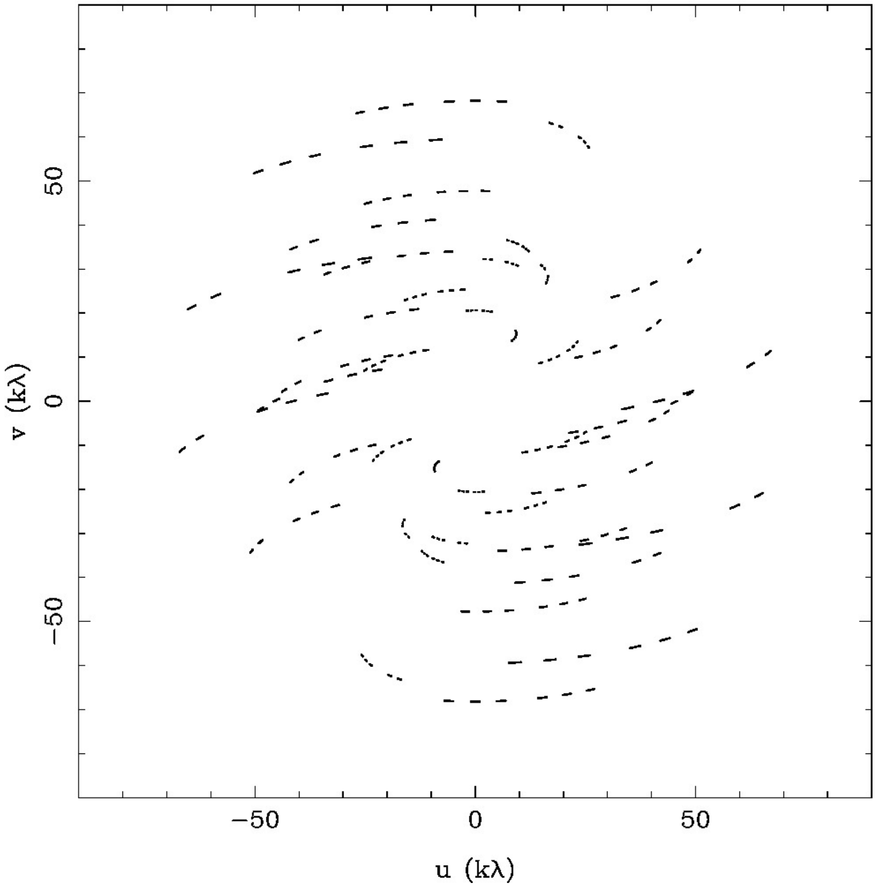
7 antennas, 1 hour

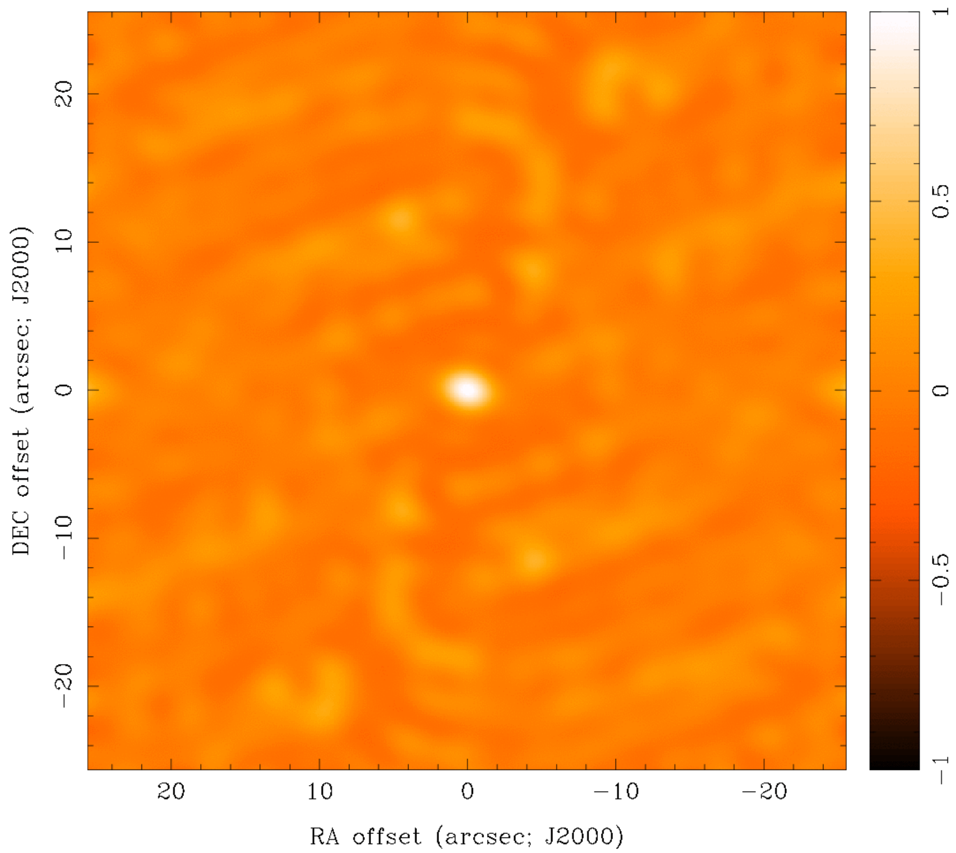




Example SMA (u,v) Plane Sampling

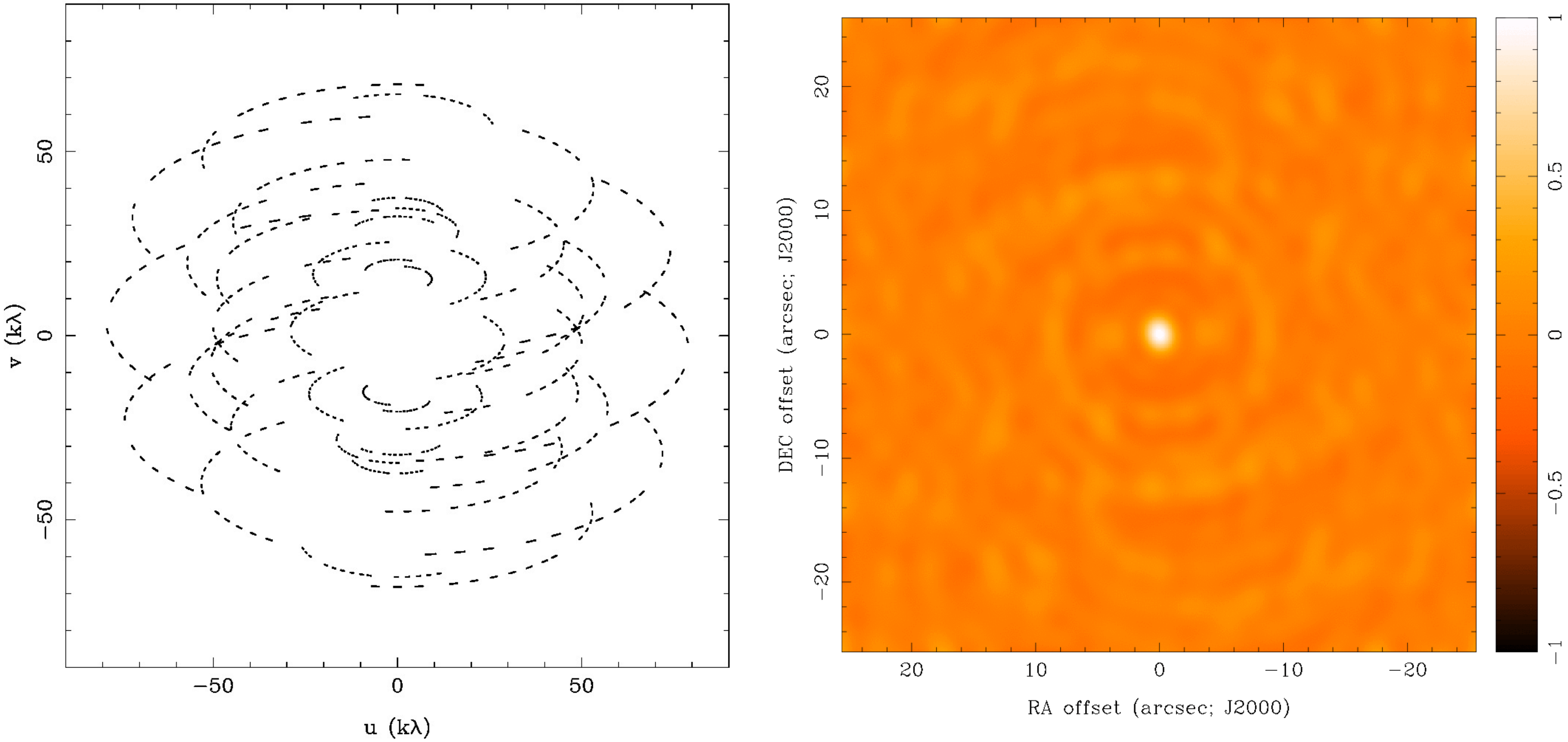
7 antennas, 3 hours



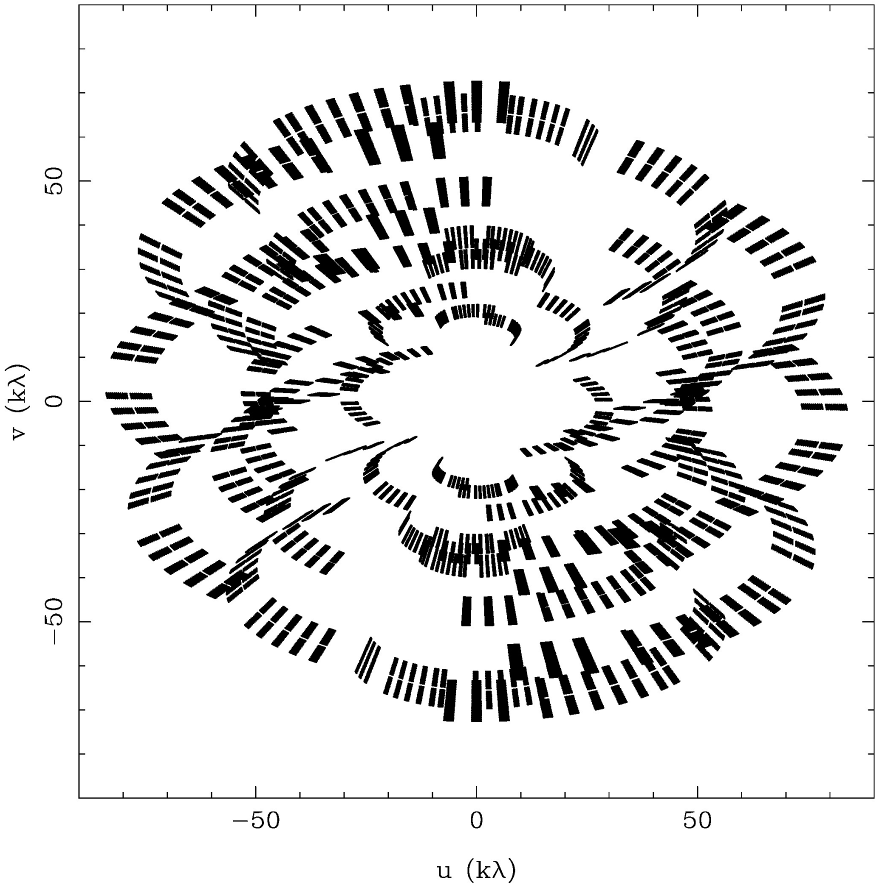


Example SMA (u,v) Plane Sampling

7 antennas, 7 hours



SMA Multi-frequency Synthesis



for continuum

"multi-frequency synthesis"

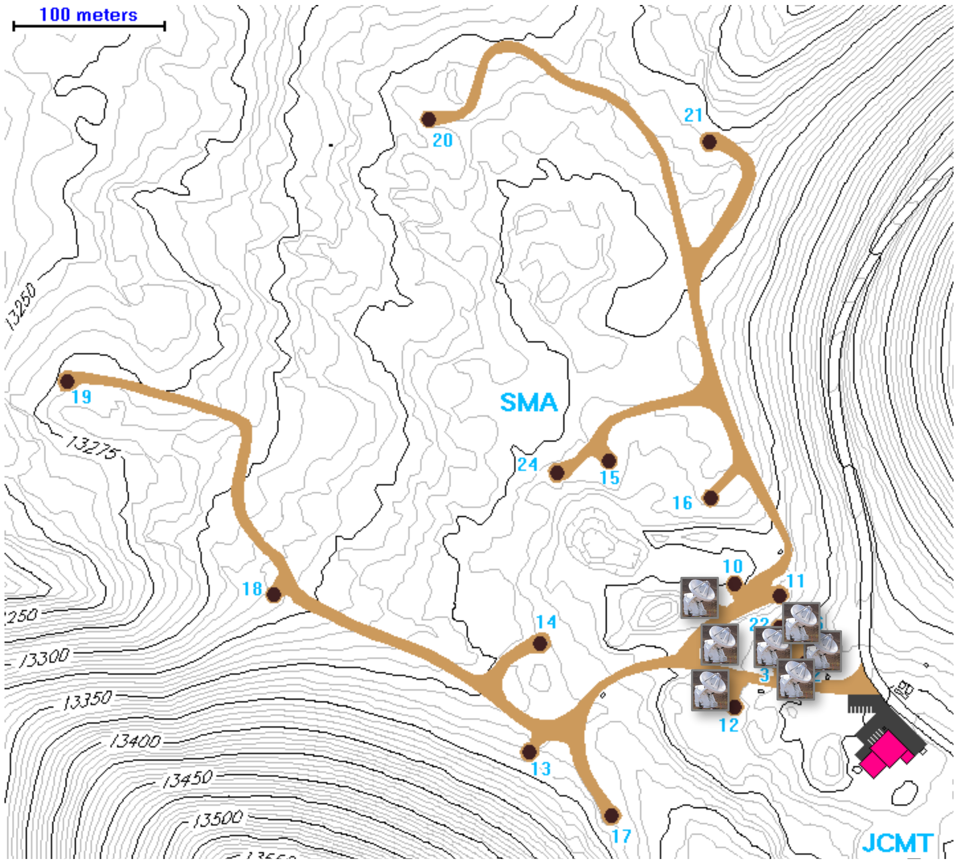
e.g. SWARM 44 GHz coverage

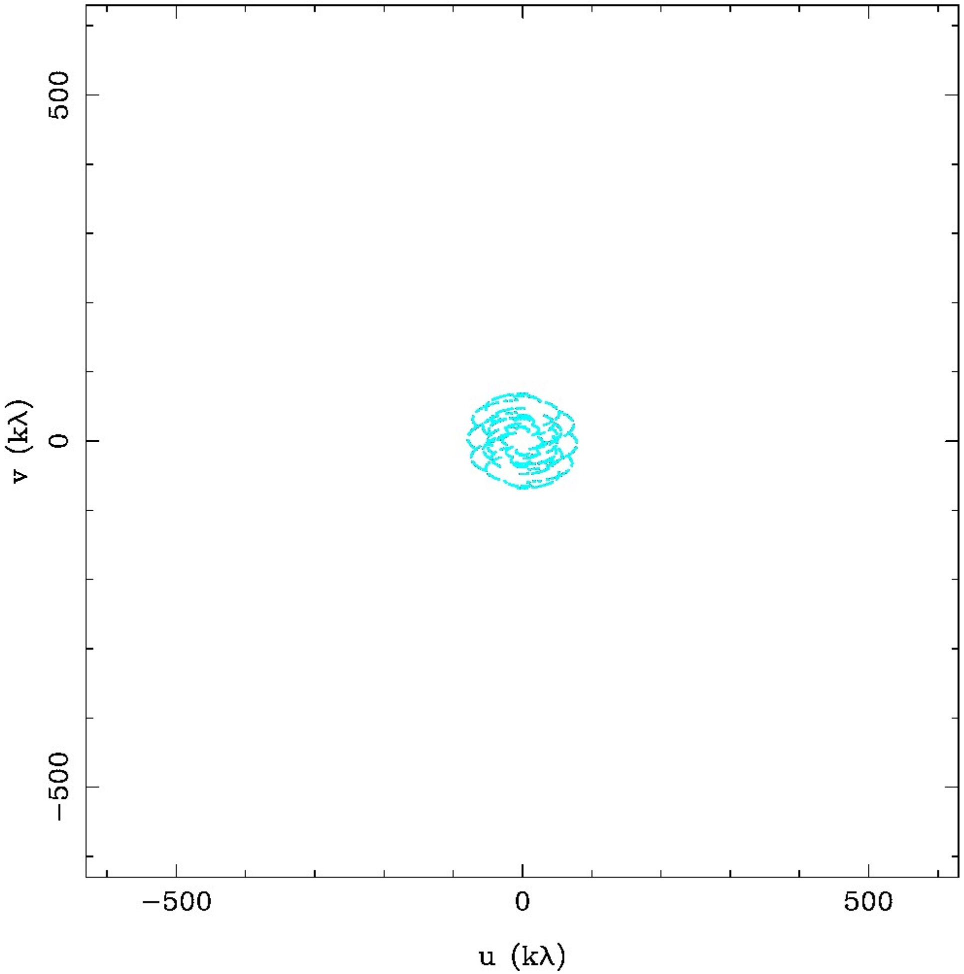
12 GHz x 2 SB x 2 pol, 4 GHz overlap

→ *(u,v)* samples spread radially

Example SMA (u,v) Plane Sampling

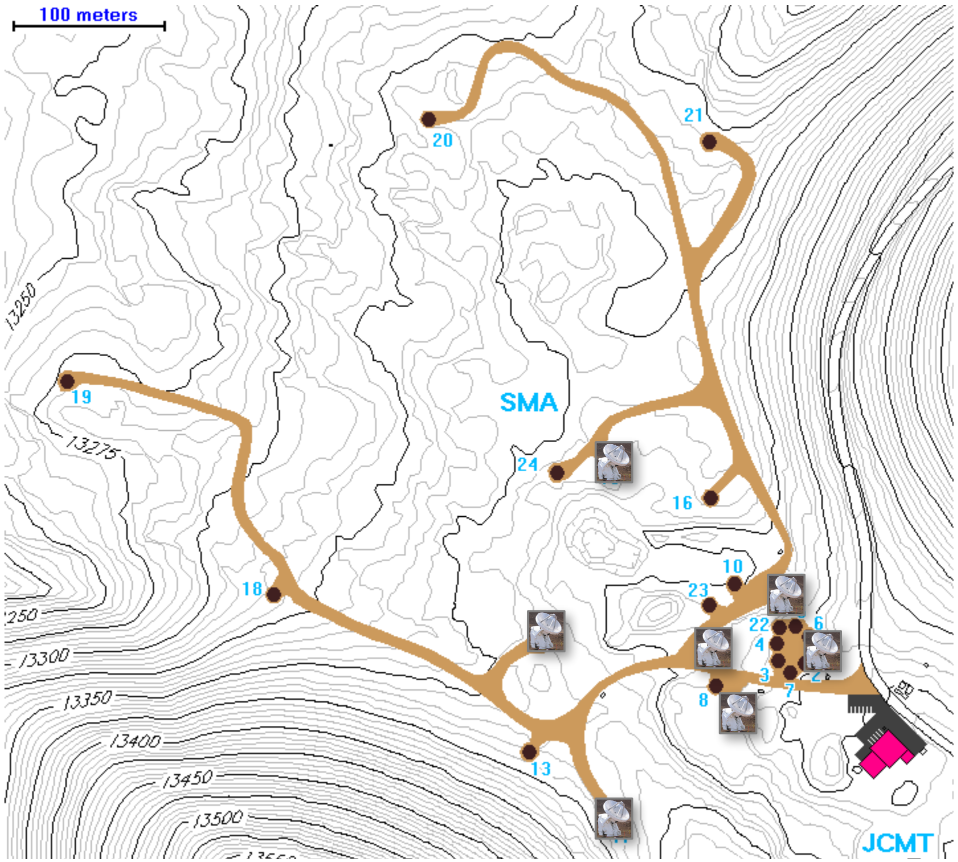
COM configuration of 7 SMA antennas, ν = 345 GHz, dec = + 22 deg

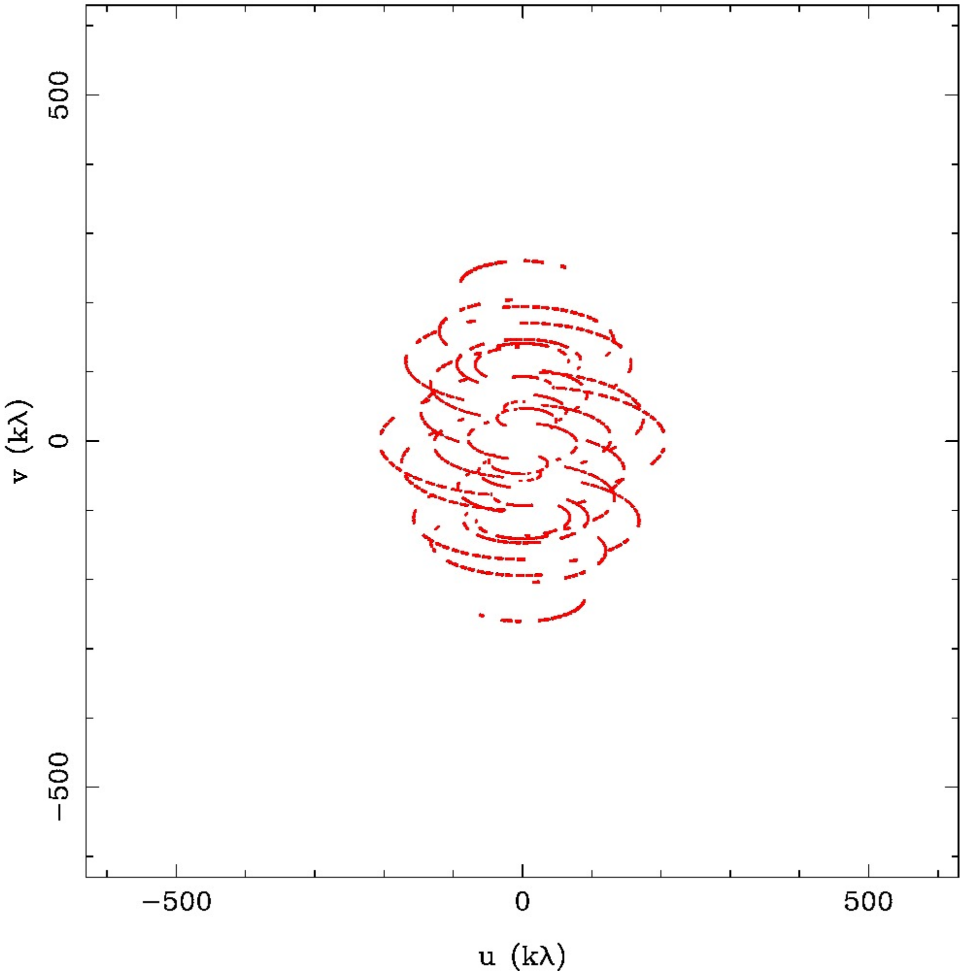




Example SMA (u,v) Plane Sampling

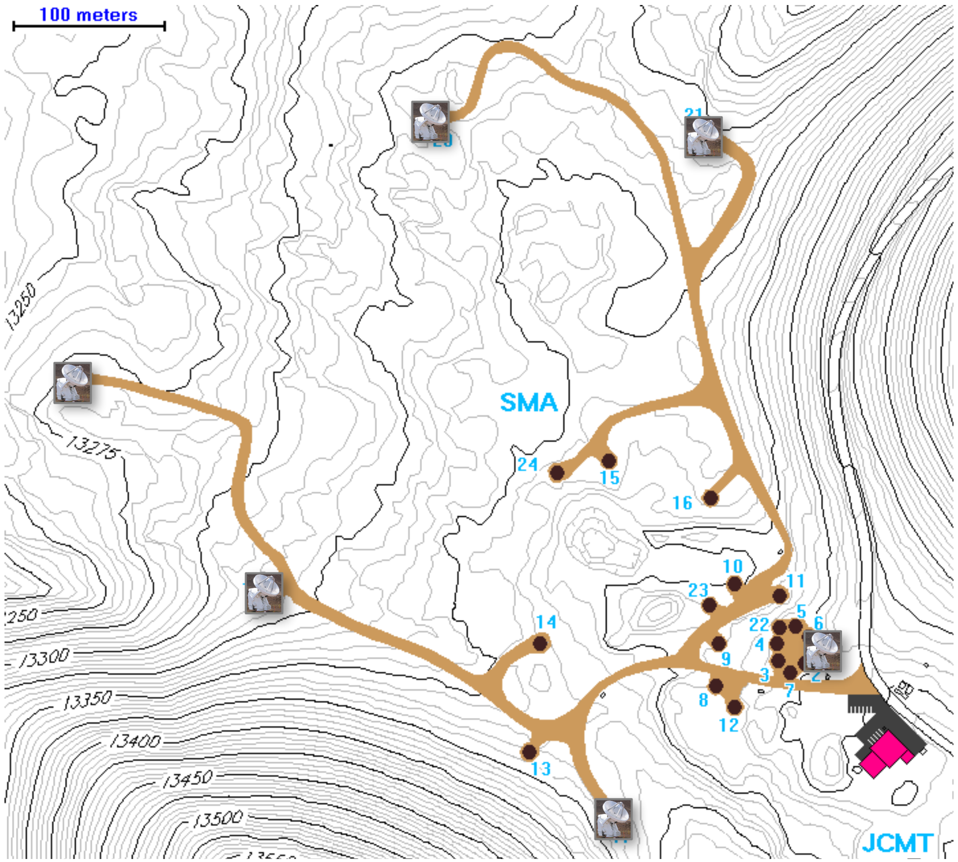
EXT configuration of 7 SMA antennas, ν = 345 GHz, dec = + 22 deg

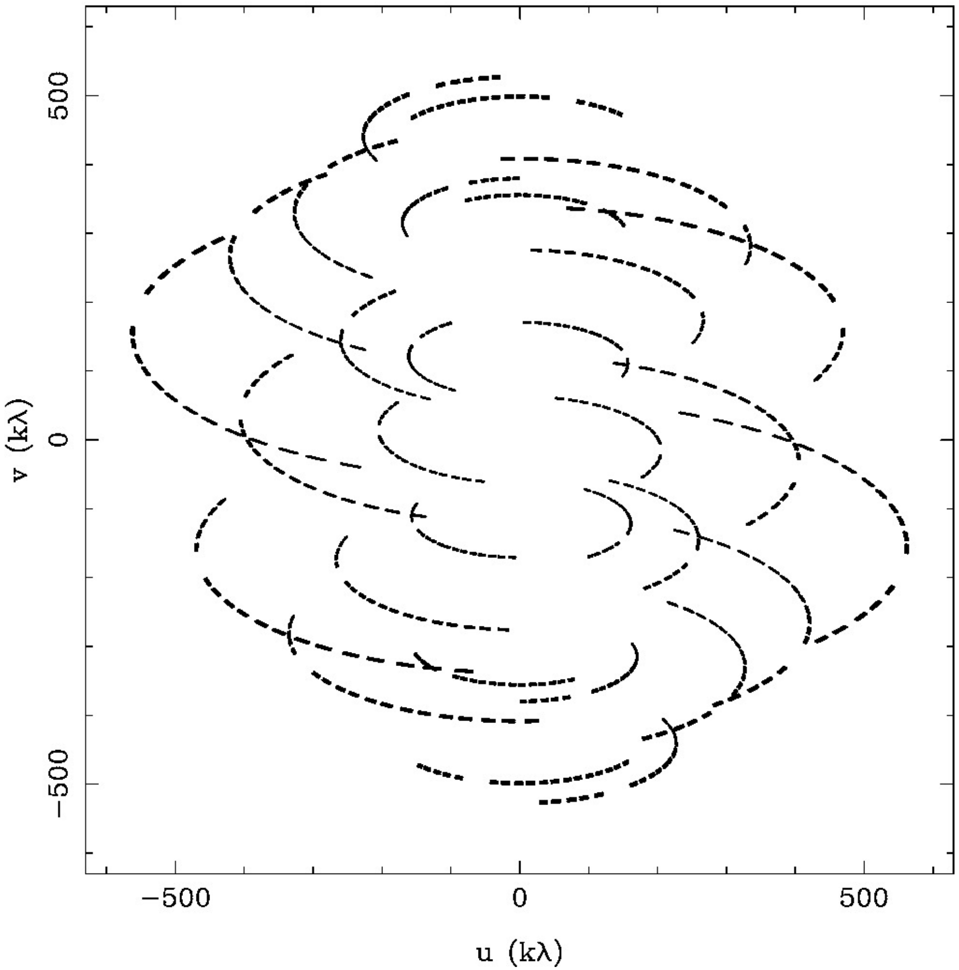




Example SMA (u,v) Plane Sampling

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





moving antennas: VLA

is a very significant operation!

VLA antennas on track

VLA configurations

VLA D configuration
B_{max} = 1 km

VLA D configuration
B_{max} = 36 km

moving antennas: ALMA

Scheuerle Fahrzeugfabrik GmbH (www.scheuerle.com)
Pfedelbach - just around the corner!

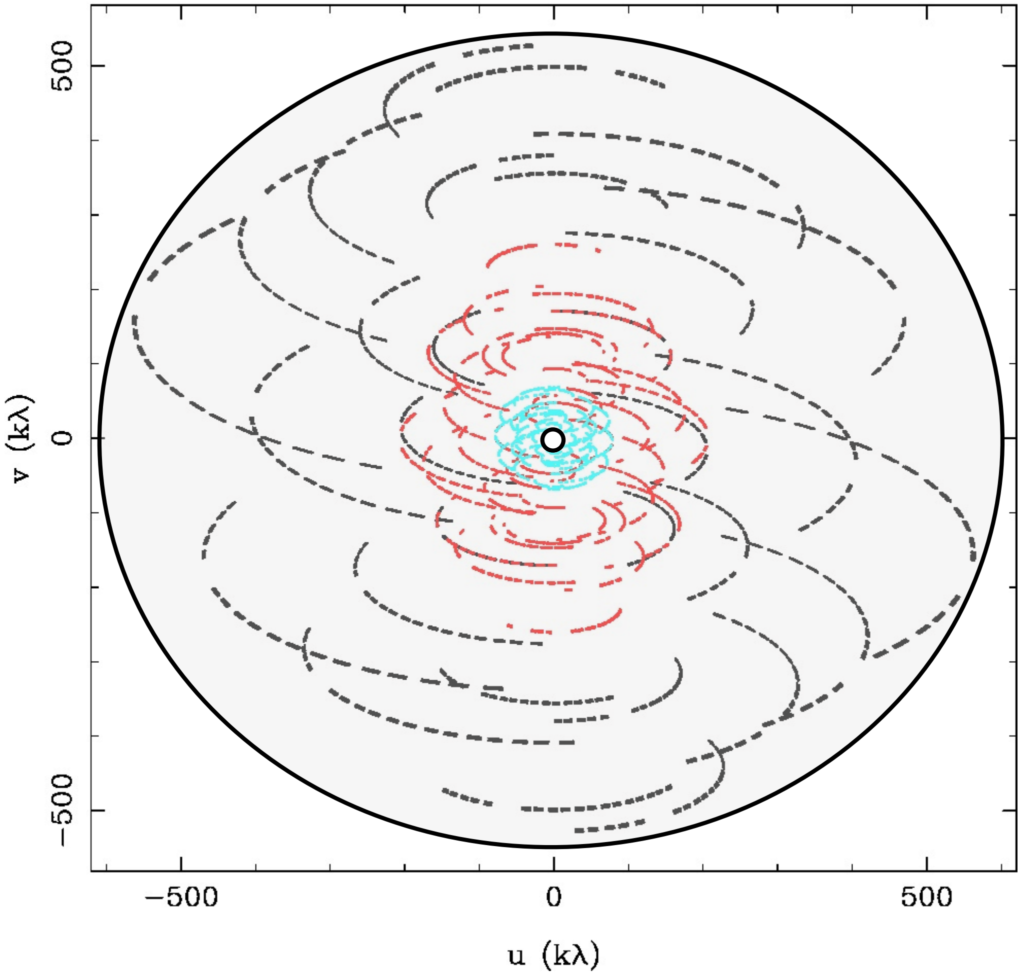
10 metres wide, 20 metres long and 6 metres high

moving antennas: ALMA

note: no tracks, and no actual configurations. antennas are always moved ‘breathing array’

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

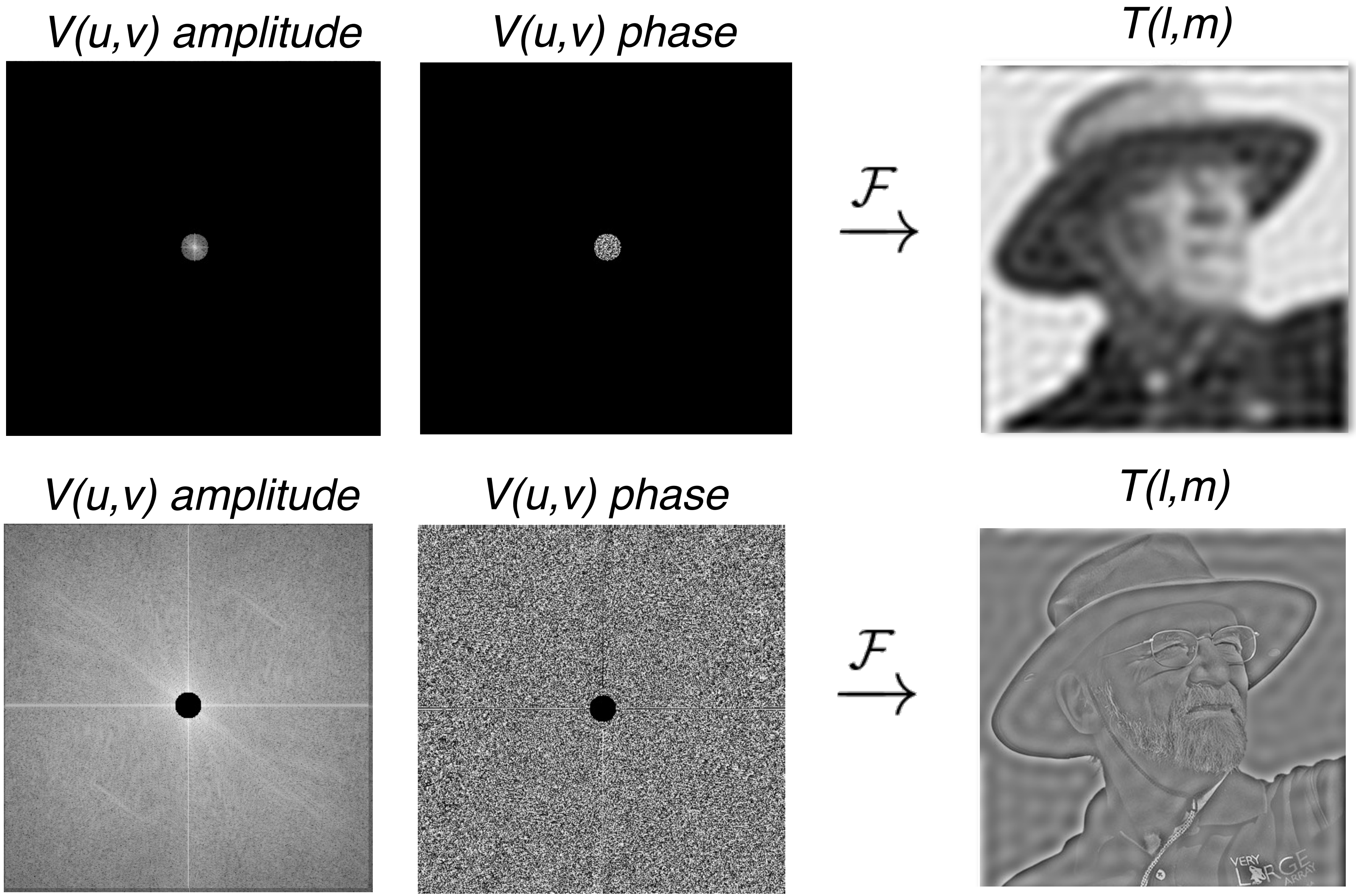
inner boundary

- no info on larger scales
- extended sources invisible

irregular coverage in between

- sampling theorem violated
- information missing

Inner and Outer (u,v) Boundaries

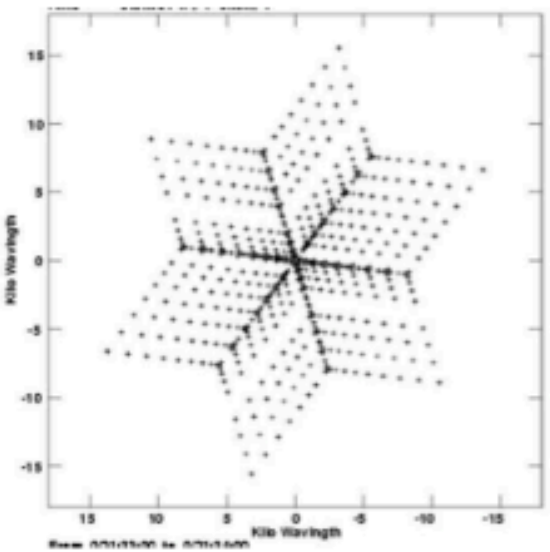


snapshot of uv coverage

Projection off Meridian

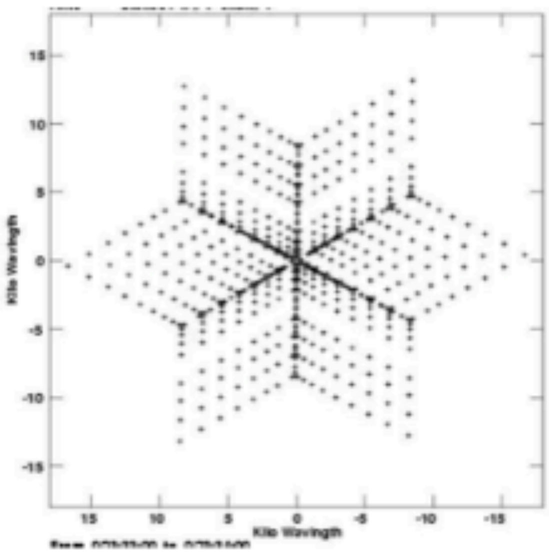
u

v



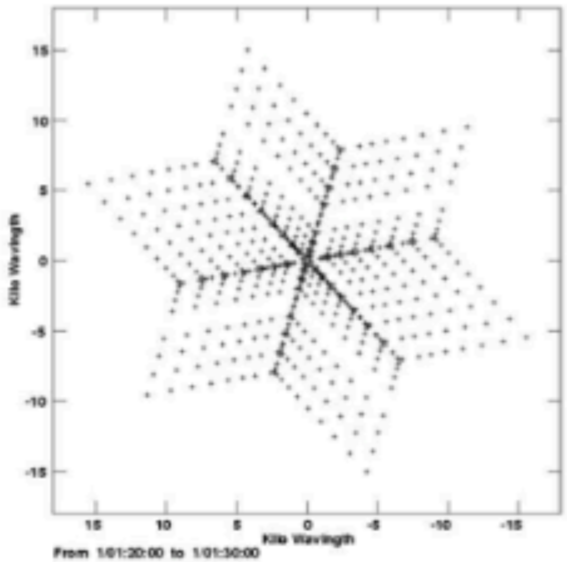
HA = -2h

Meridian



HA = 0h

Projection off Meridian



HA = 2h

A Preprint
from the

OWENS VALLEY RADIO OBSERVATORY

California Institute of Technology
Pasadena, California

1972

I. APERTURE SYNTHESIS STUDY OF NEUTRAL HYDROGEN IN THE GALAXIES NGC 6946 AND IC 342

by

D. H. Rogstad, G. S. Shostak and A. H. Rots

can get uv coverage with just 2 antennas!

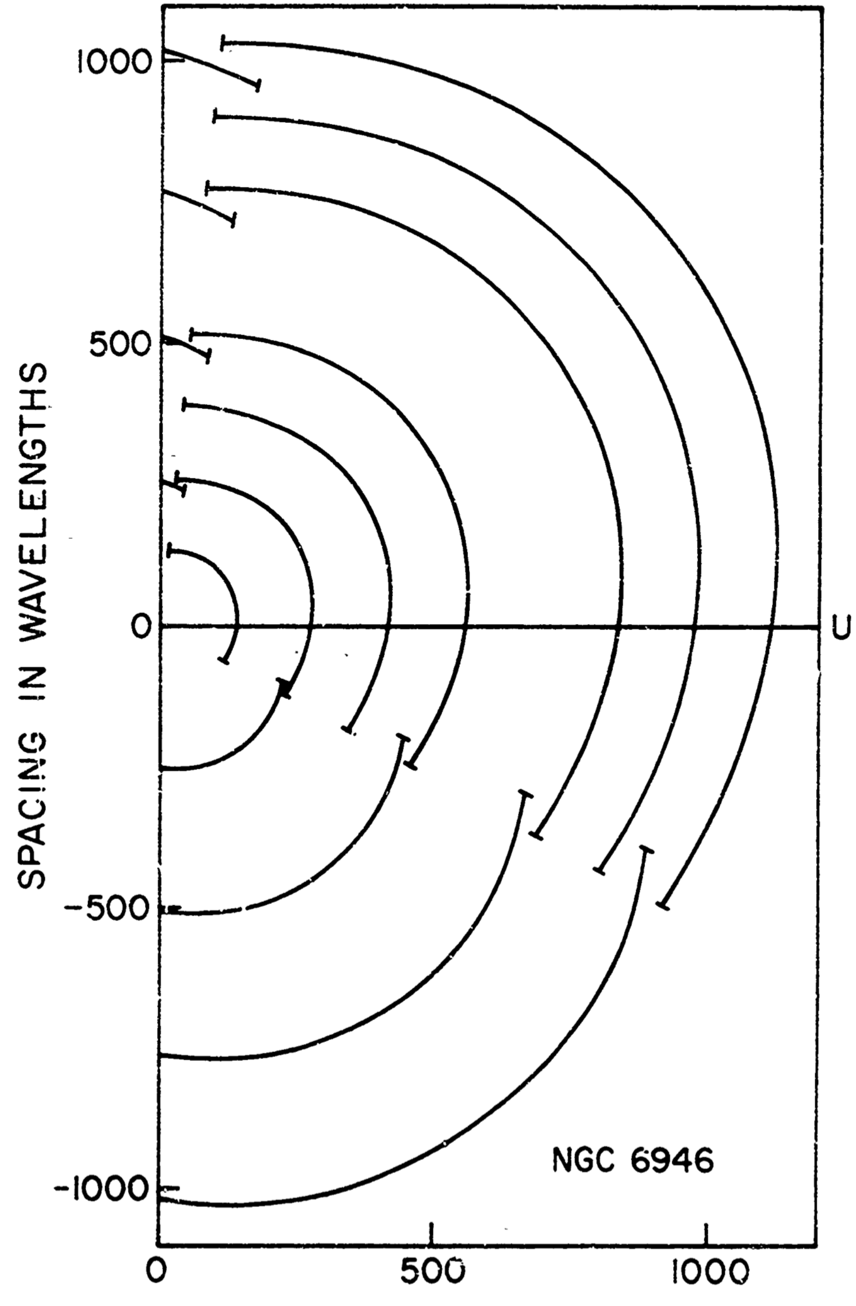


Figure 1a -16-

one antenna pair - antennas moved over time!

2D high-resolution map with 2 antennas

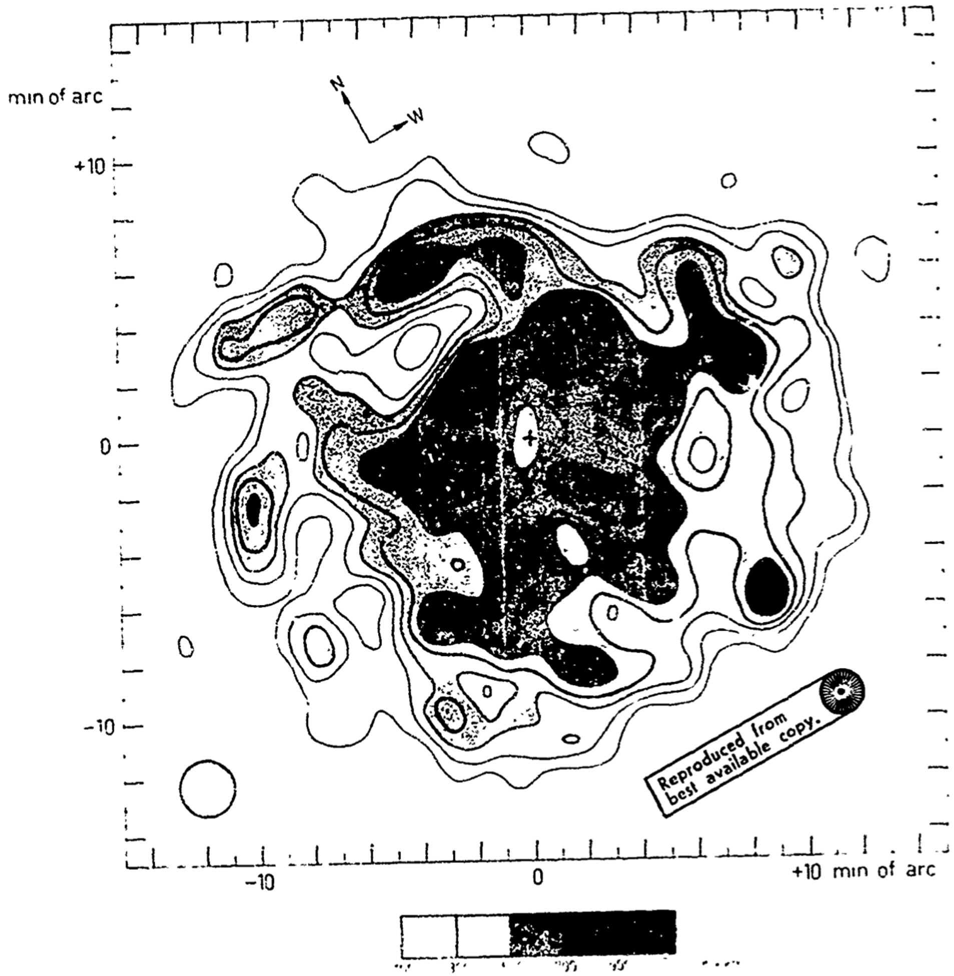
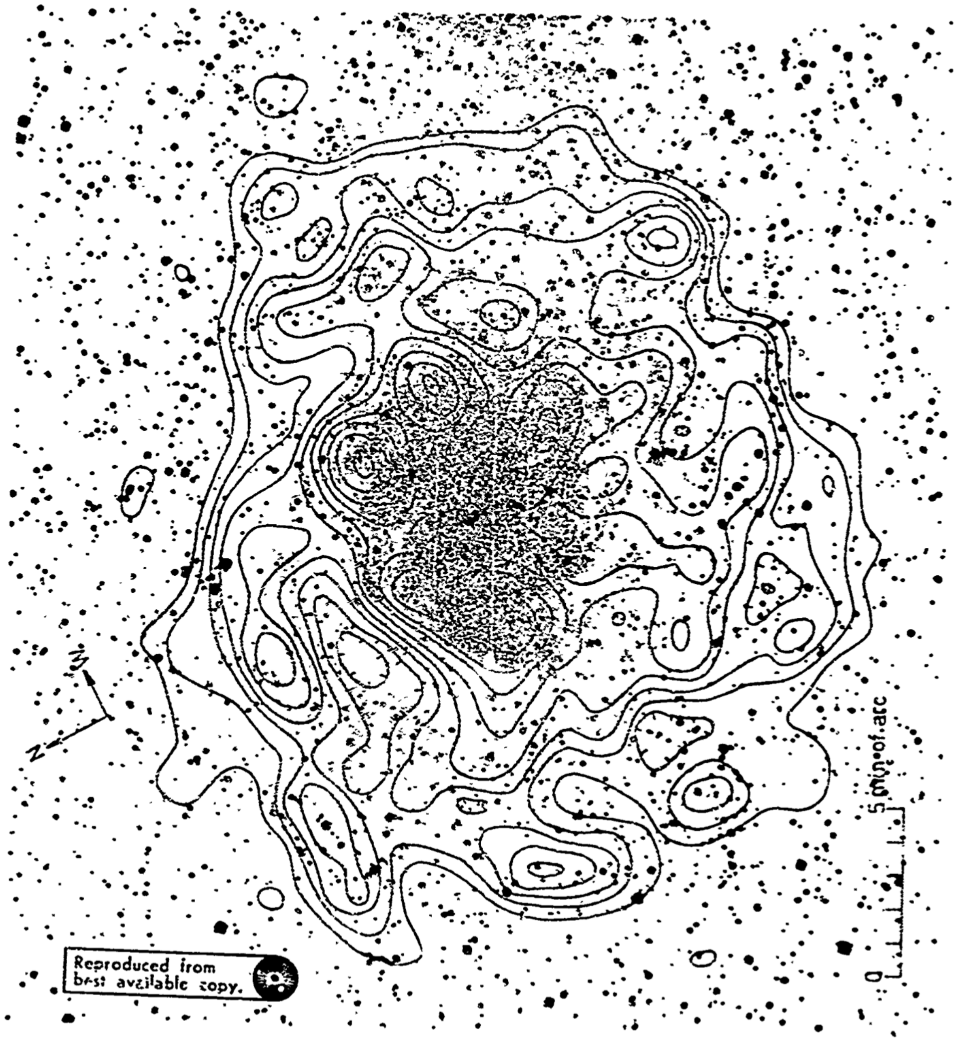


Figure 2

-18-

atomic hydrogen

2D high-resolution map with 2 antennas



-28-

Ficure 11

overplotted on optical emission

sensitivity of an interferometer

$$S_{\rm rms} = \frac{2kT_S}{A\eta_A \sqrt{n_a(n_a-1)\Delta\nu_{\rm IF}\tau_0}}$$

large number of antennas n_a —> $S \sim 1/n_a$

as in the case of single dish: sqrt-dependence on frequency width and integration time

interferometry vs. single dish

pros

- 100-m baseline is a lot cheaper than building a 100-m antenna
- Capability for reconfigurable spatial and spectral resolution
- Lots of things are better when your data integrates down to zero
- Get multi-pixel images with (b/D)^2 pixels
- get positions of sources with subarcsec position (tracking errors of telescopes much bigger)

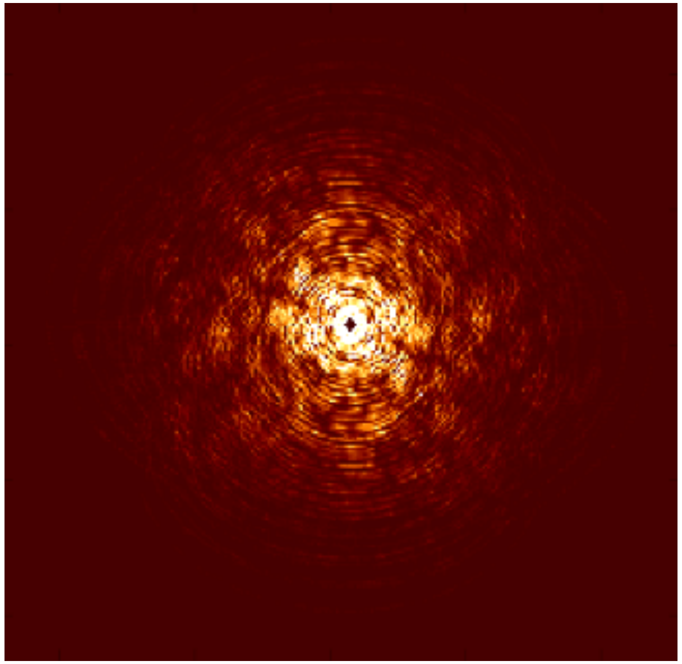
interferometry vs. single dish

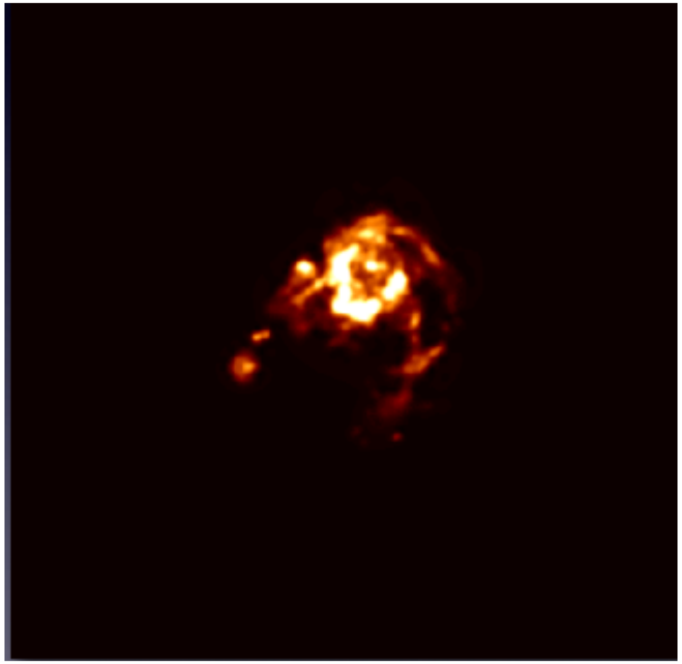
cons

- 100-m baseline doesn't have the sensitivity of a 100-m single dish
- Can't use incoherent detectors (e.g., bolometers)
- Data processing tends to be much more complex (both in realtime and offline)
- Interferometers can't recover very largest spatial scales (roughly limited to primary beam size)

basic imaging

How do we go from the measurement of the visibility function to images of the sky?





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

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

- apply the convolution theorem

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

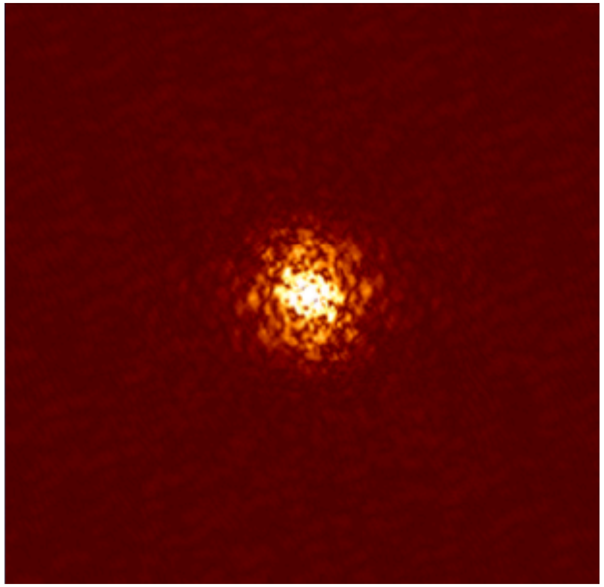
where the Fourier transform of the sampling pattern $s(l,m) \xrightarrow{\mathcal{F}} S(u,v)$

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

- radio astronomy jargon:

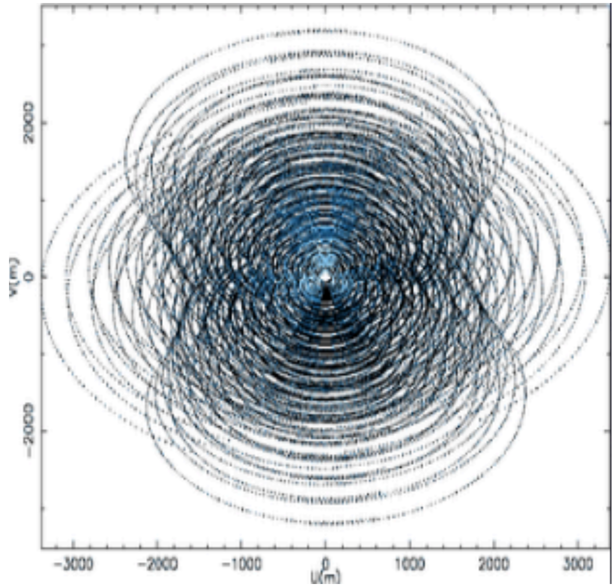
  the “dirty image” is the true image convolved with the “dirty beam”

uv plane only sampled at discrete points



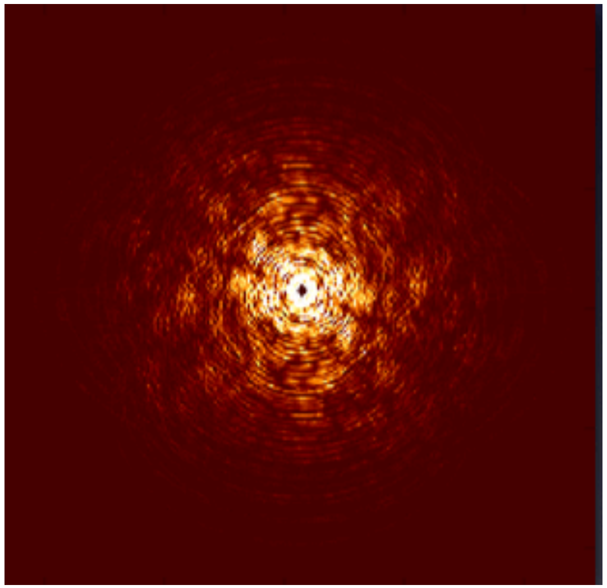
ideal case — V(u,v)

X



uv coverage (S(u,v))

=



measured visibilities — $V_M(u,v)$

snapshot of uv coverage

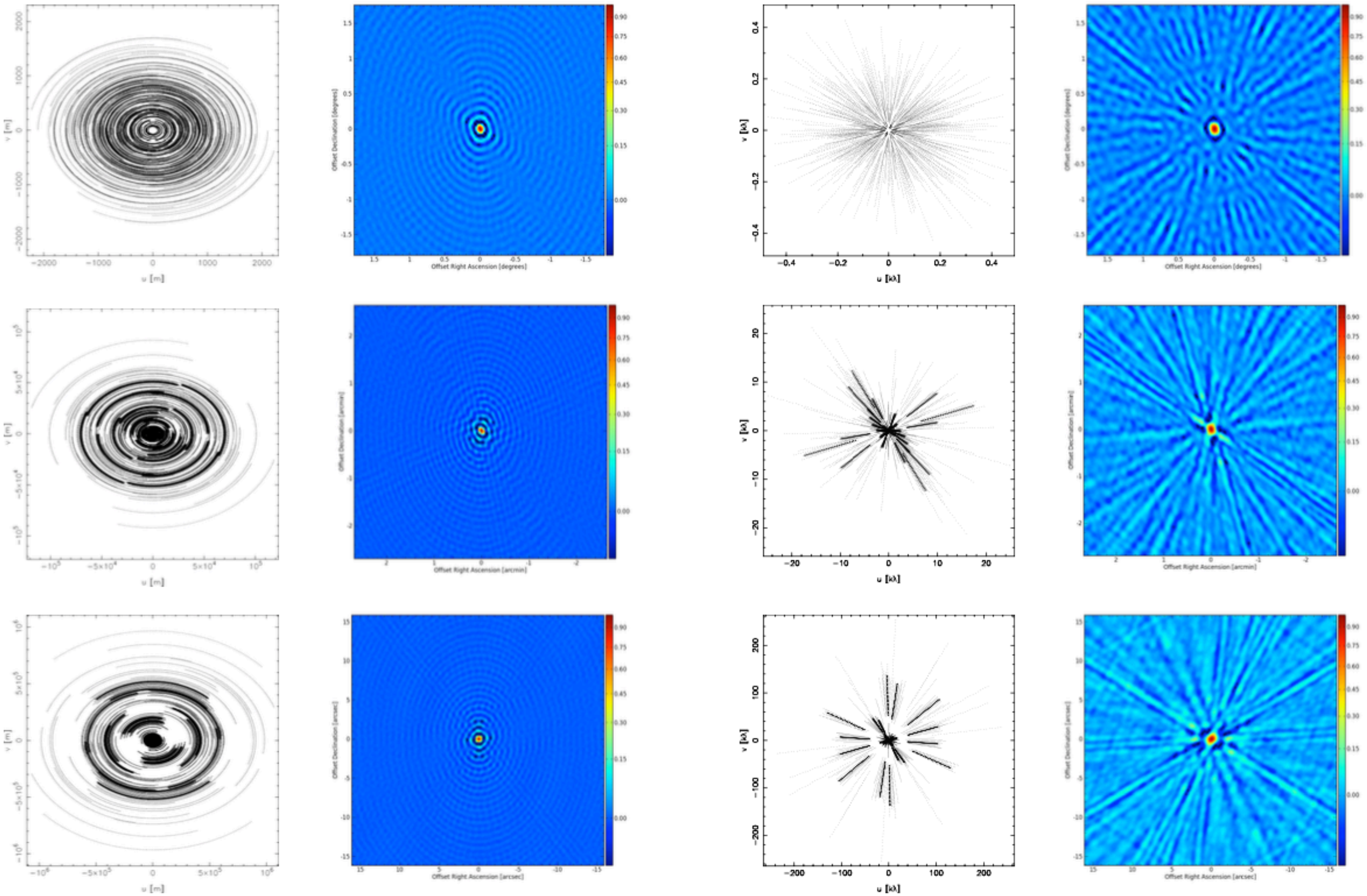
6 hr track | *Instantaneous*

LOFAR

10 km

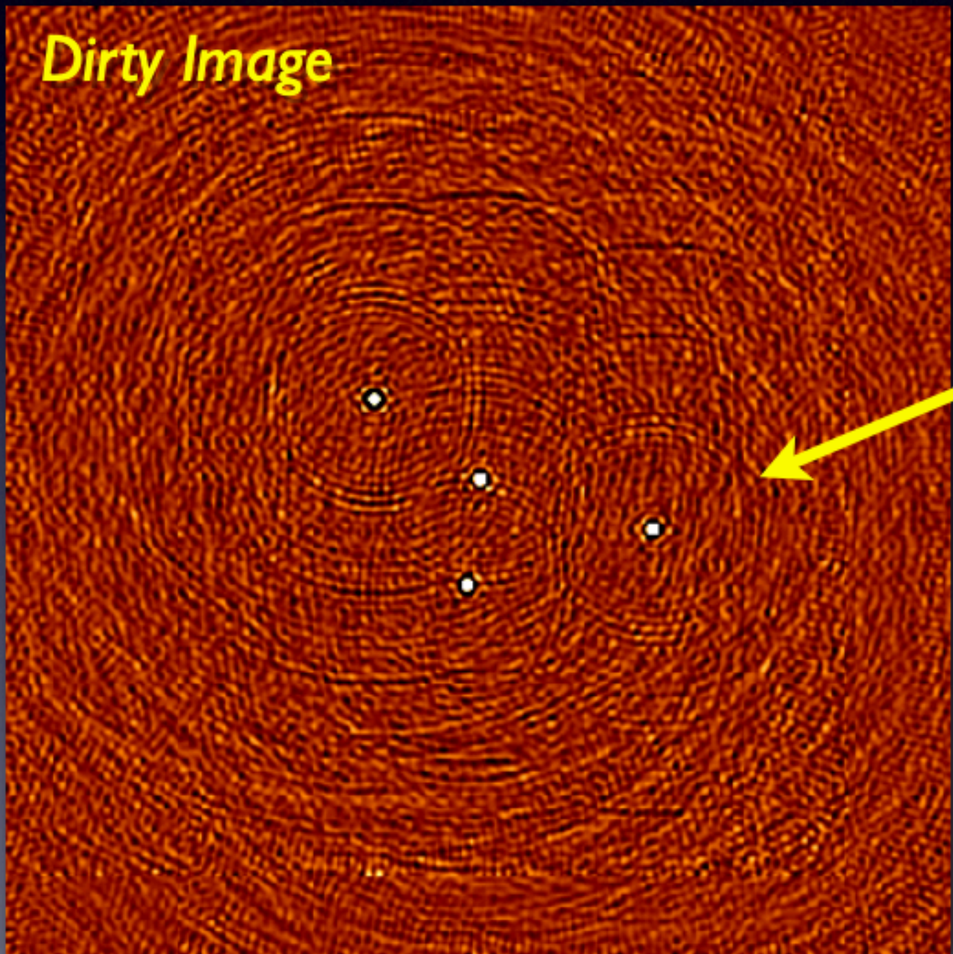
100 km

1000 km



convolution with Dirty Beam

$$I_D(x,y) = \sum_i B(x - x_i, y - y_i) * I(x_i, y_i)$$



$$= I(x_0, y_0) * B(x - x_0, y - y_0) + I(x_1, y_1) * B(x - x_1, y - y_1) + \ldots$$

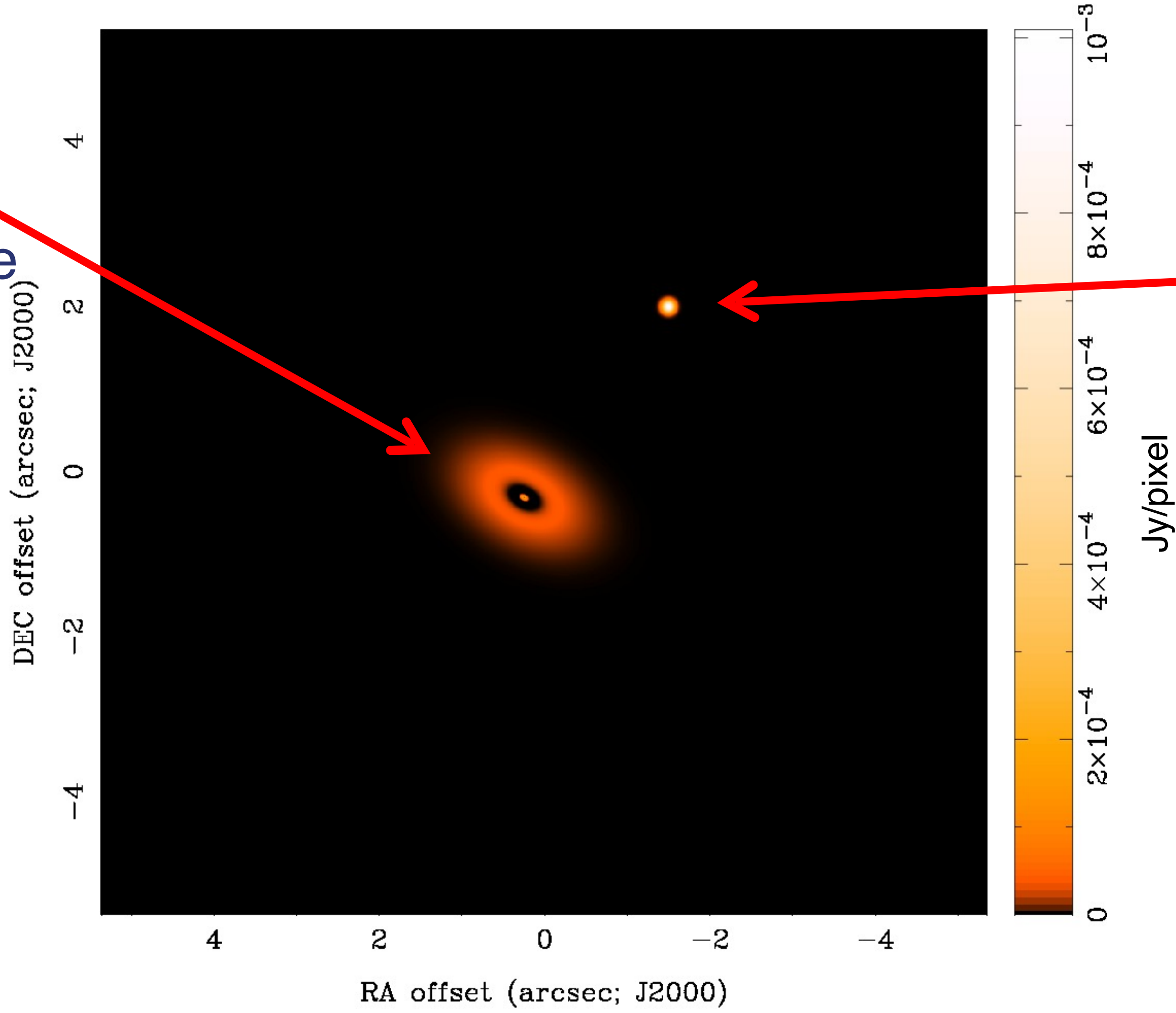
Dirty beam can vary with time and position across the field

Example model sky brightness

disk

+

central source

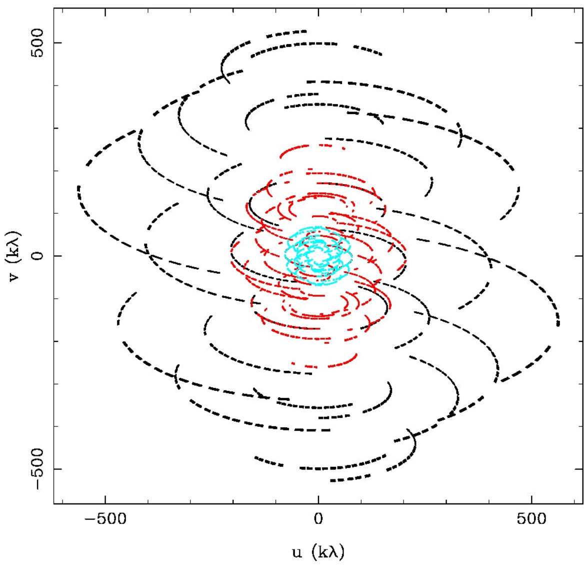


compact blob

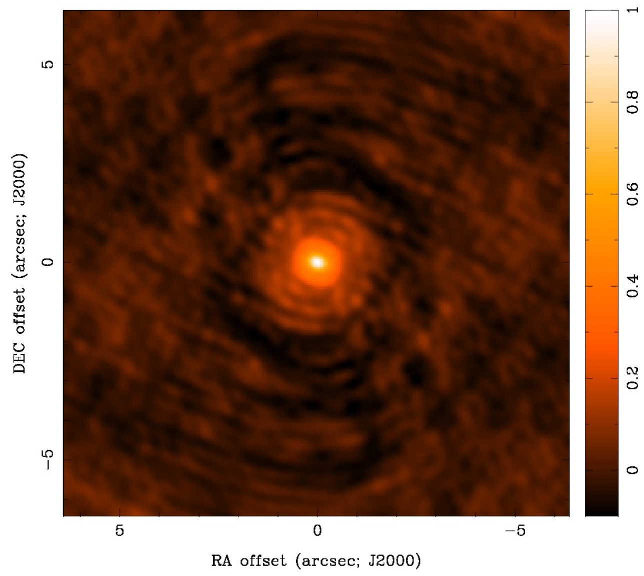
(0.25x disk flux)

Dirty beam and dirty image

$S(u,v)$

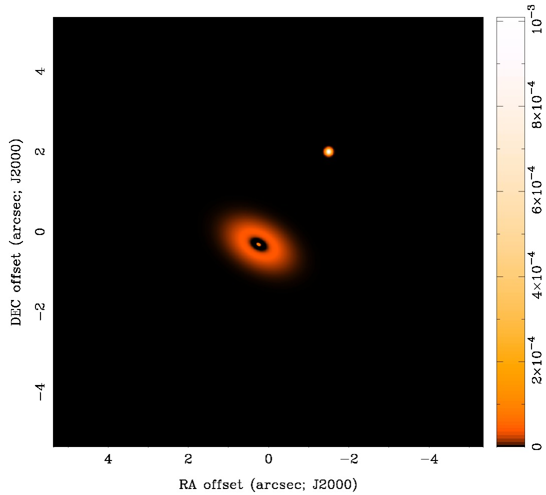


$\xrightarrow{\mathcal{F}}$



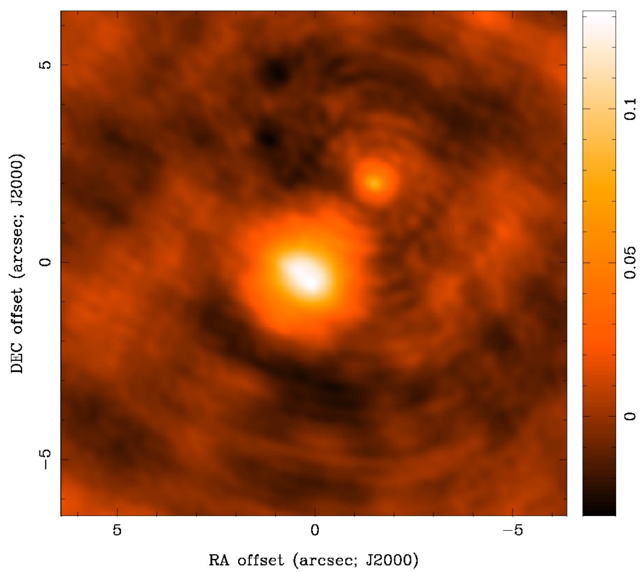
$s(l,m)$
"dirty beam"

$*$



$T(l,m)$

$=$



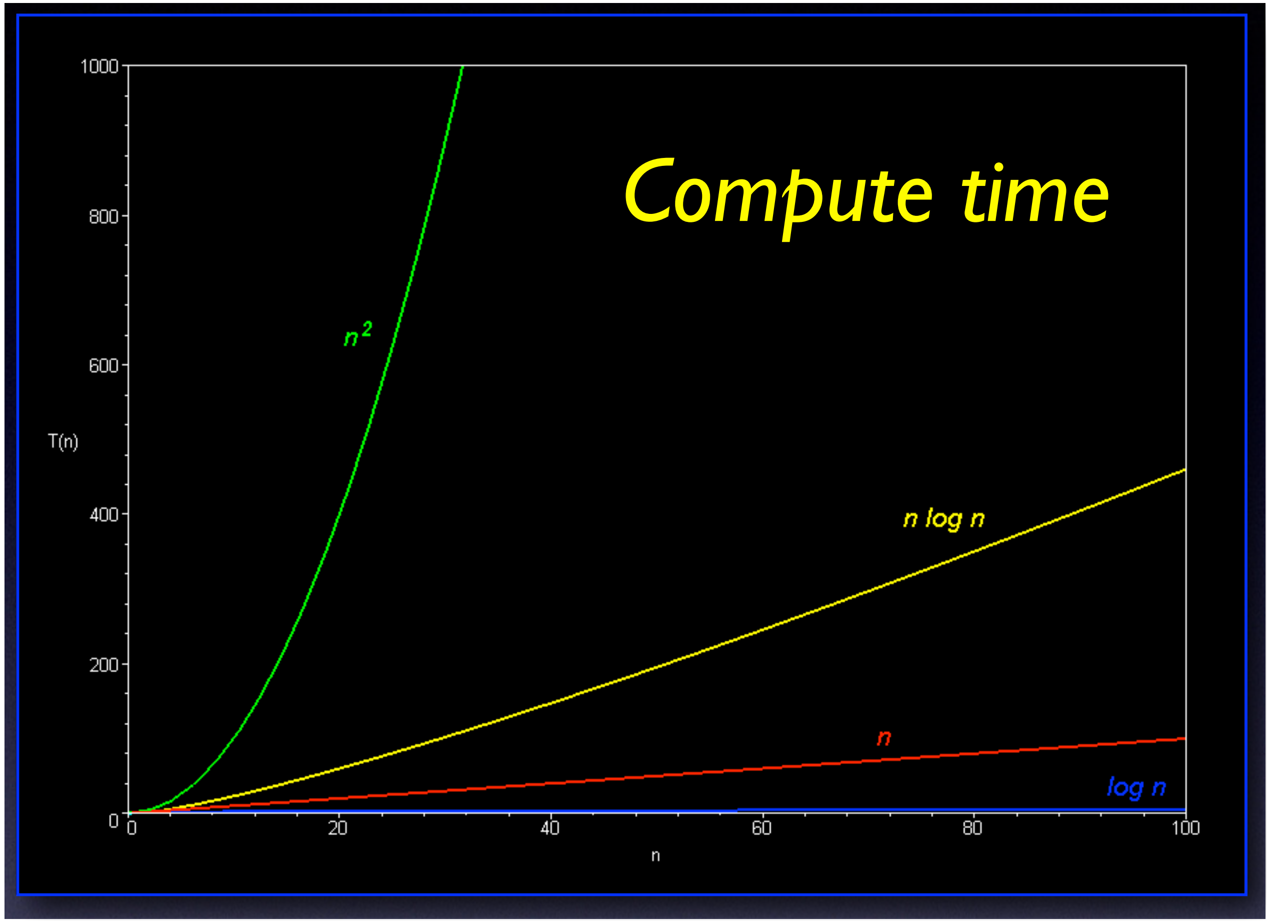
$T^D(l,m)$
"dirty image"

Fast Fourier Transform

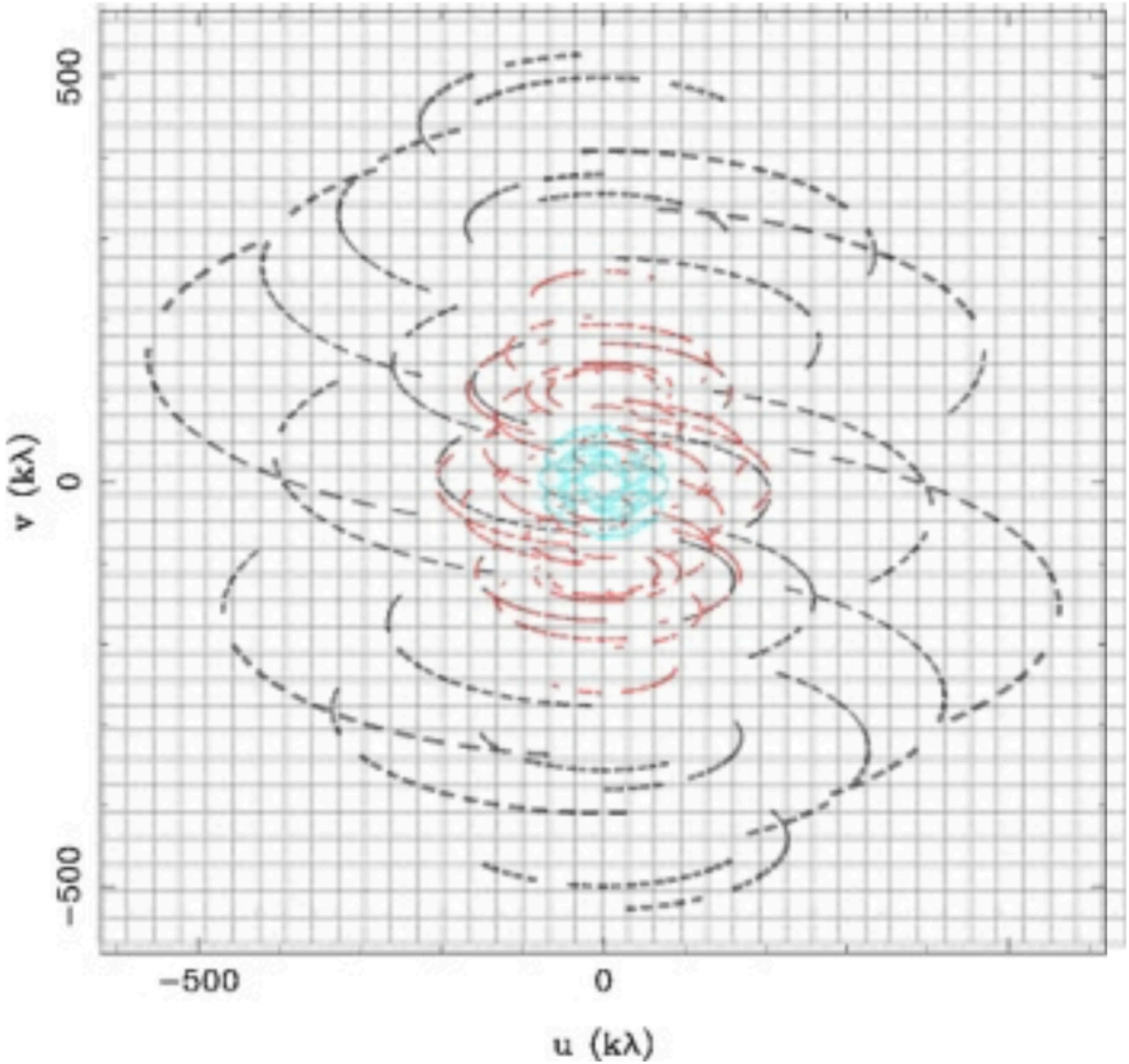
Fast Fourier Transform (FFT) is used to compute the Fourier integral

- 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(NlogN)
  - but speed 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

Fast Fourier Transform: Speed



FFT: the need for gridding

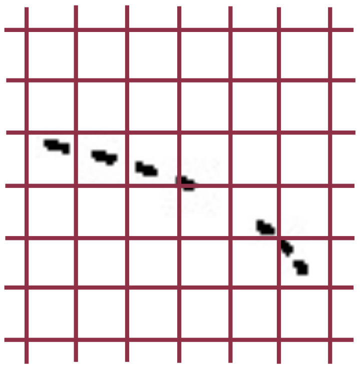


FFT: the need for gridding

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

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



- 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”)

gridding: 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)

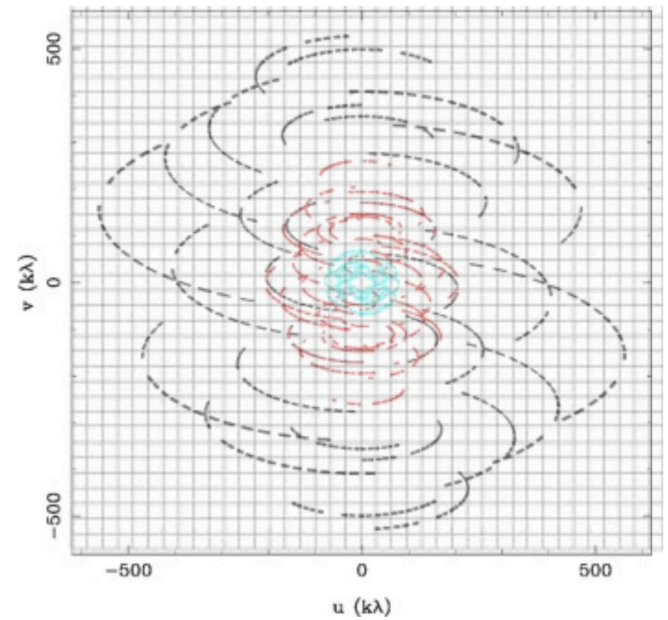
pixel size

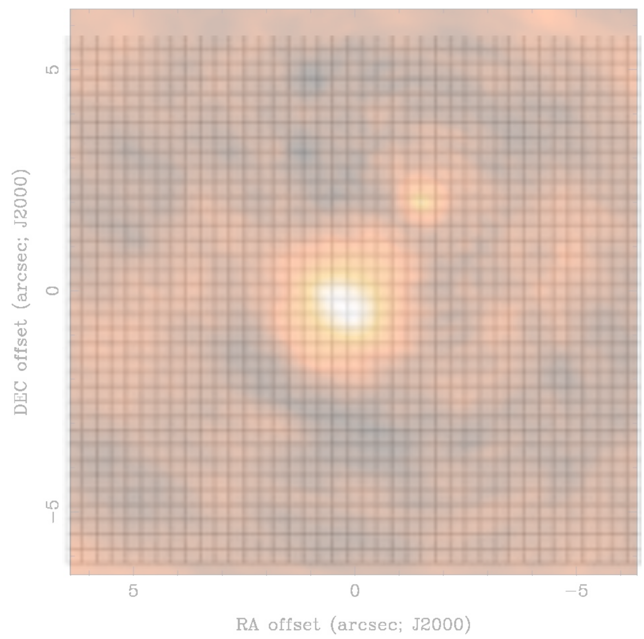
- satisfy Nyquist-Shannon sampling theorem for longest baselines

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

- in practice, use 3 to 5 pixels across dirty beam main lobe to aid deconvolution process

grids in UV space vs RA / DEC





u,v

l,m (RA/DEC)

pixel size —> image size
and vice versa!!!

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

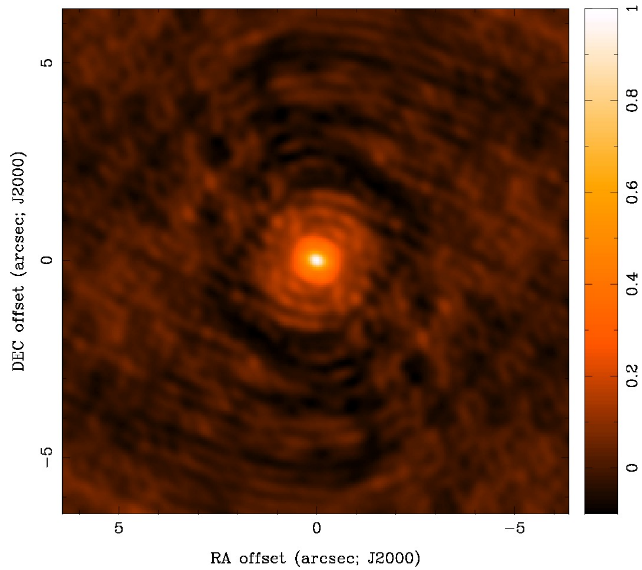
- Another way is to introduce a weighting function, $W(u,v)$, into the visibility gridding process
  - $W(u,v)$ modifies the sampling function: $S(u,v) \rightarrow S(u,v)W(u,v)$
  - changes the dirty beam shape

- $W(u,v)$ 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 visibility noise)

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



Gaussian fit to central core

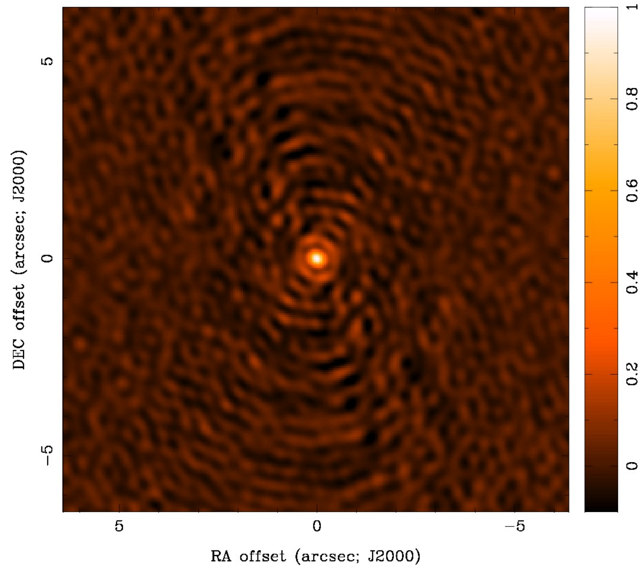
0.59x0.50 arcsec

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

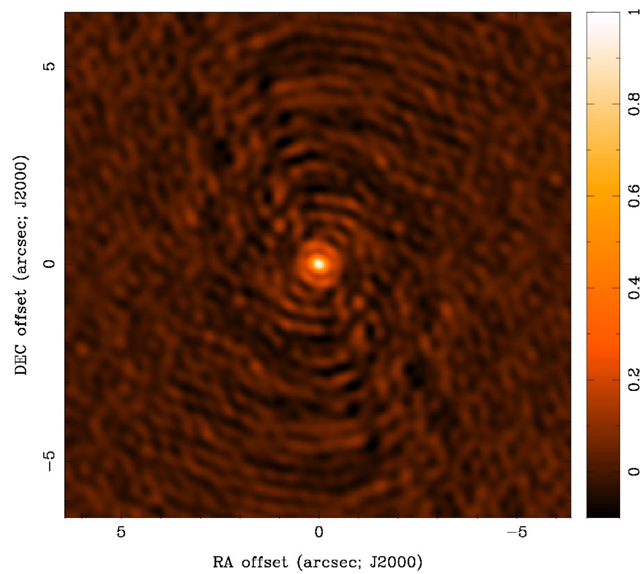
- software implementations differ
- e.g. $$W(u,v) = \frac{1}{\sqrt{1 + S_N^2 / S_{thresh}^2}}$$

  S_N is cell natural weight

  S_{thresh} is a threshold parameter

  high S_{tresh} -> NA, low S_{tresh} -> UN

- 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

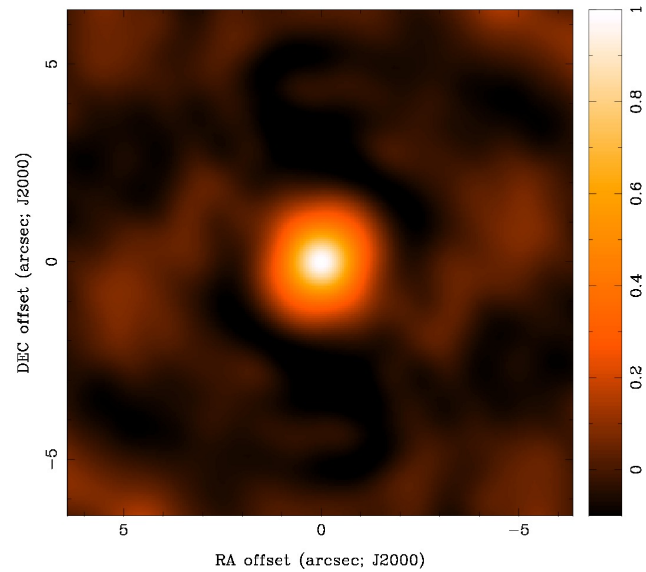
Tapering

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
    → 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



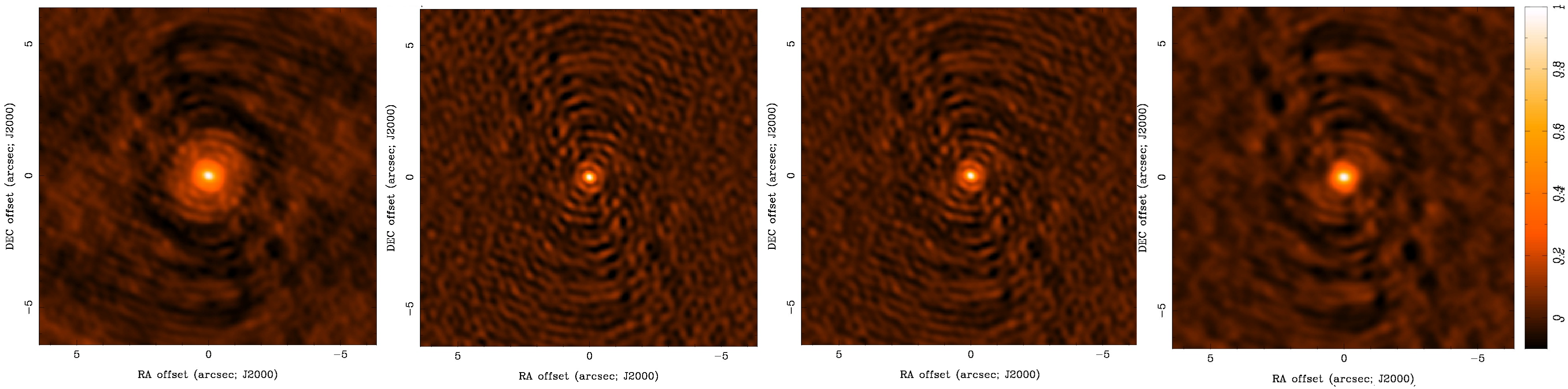
Gaussian fit to central core

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



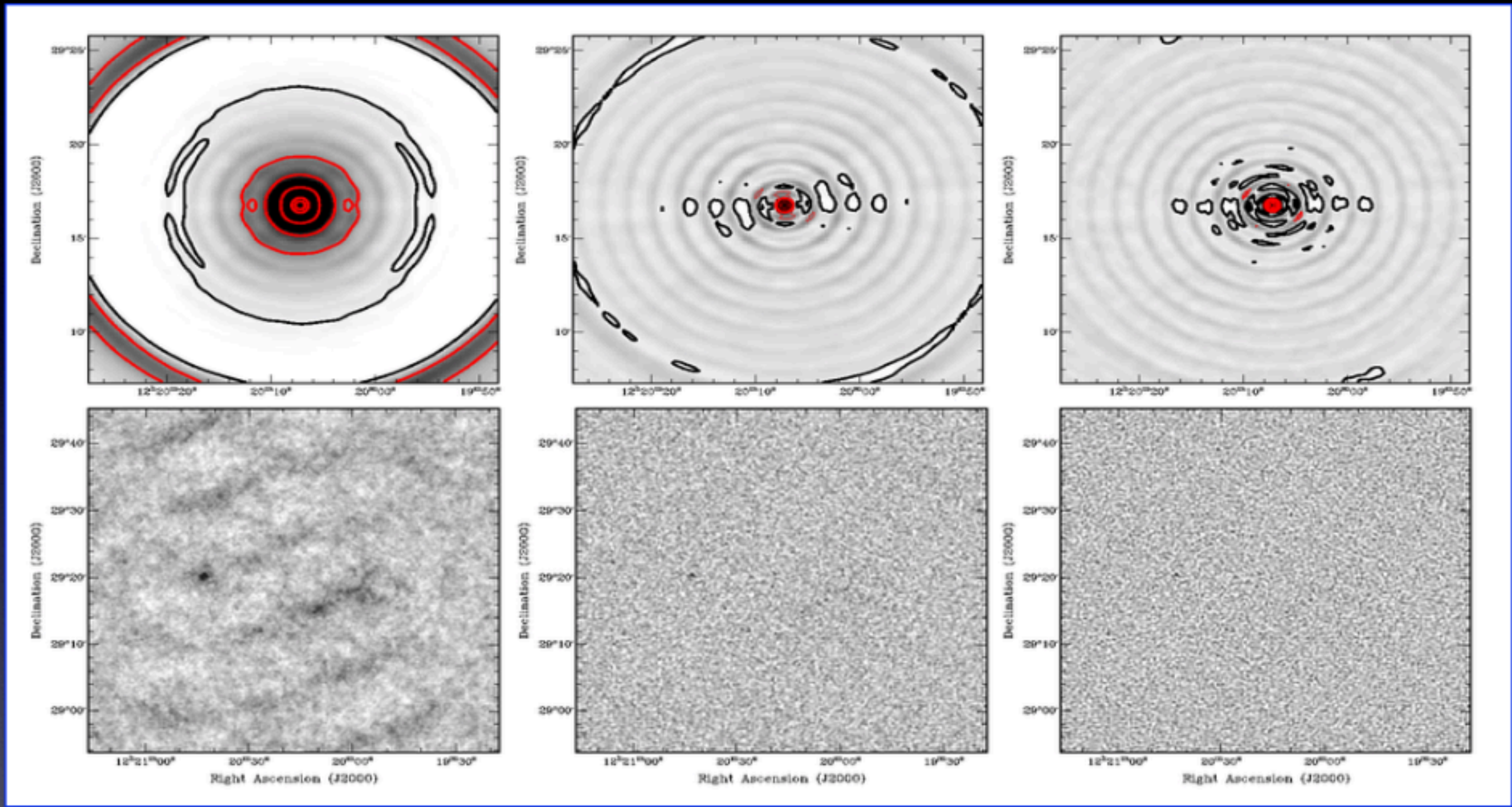
natural
0.59x0.50
ΔS = 1.0 mJy

uniform
0.35x0.30
ΔS = 2.1 mJy

robust=0
0.40x0.34
ΔS = 1.3 mJy

robust=0 + taper
0.59x0.50
ΔS = 1.2 mJy

Example: WSRT



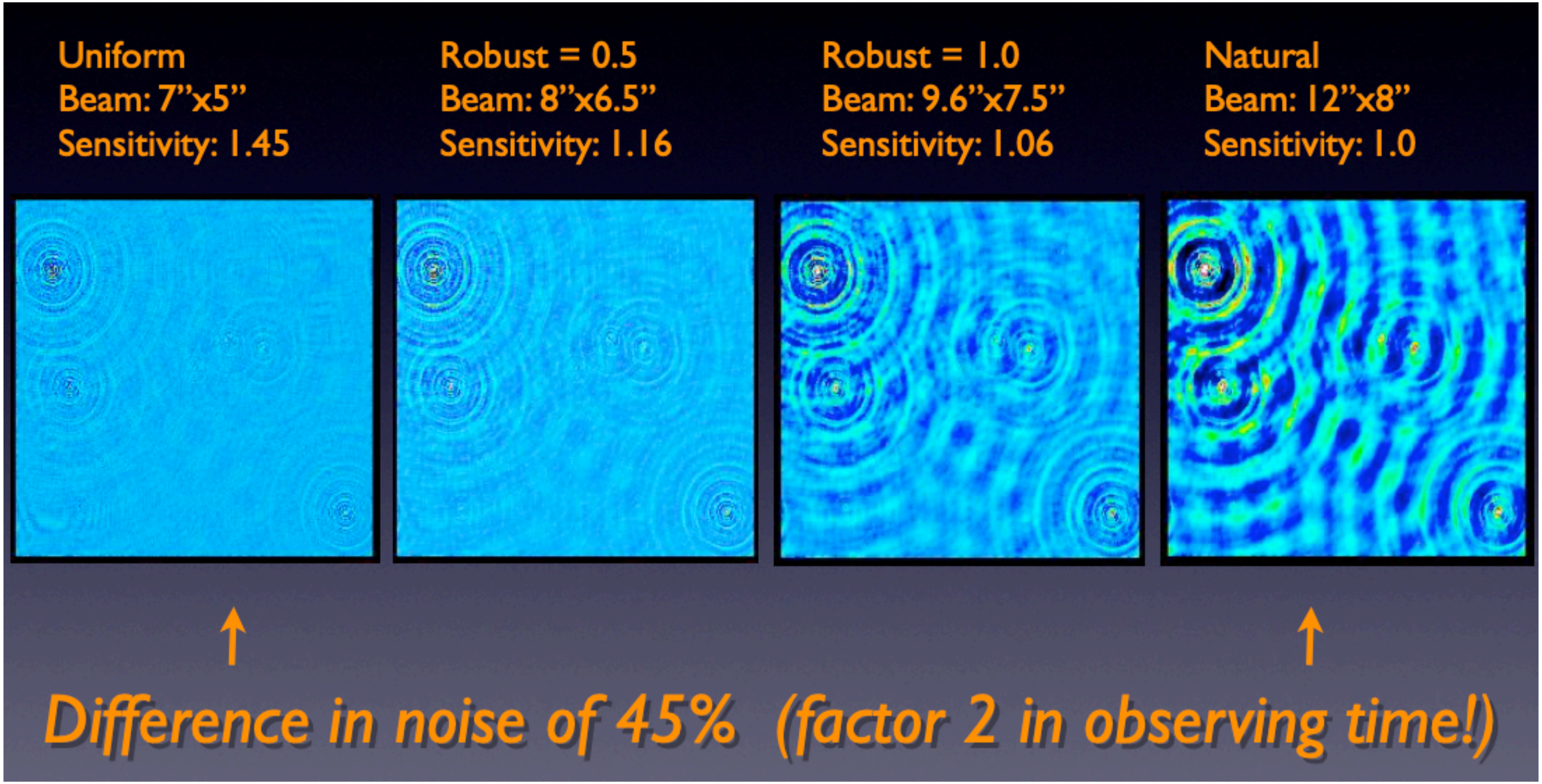
Natural weighting
σ = 0.5

Robust = 0
σ = 0.6

Uniform weighting
σ = 0.7

Difference in noise of 40% (factor 2 in observing time!)

Example: JVLA



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

\+ **taper** to
improve sensitivity to extended emission, reduce dirty beam sidelobe structure, match resolution from different observations, …