Widefield Imaging II: Mosaic(k)ing + Short/Zero Spacing Correction

Thirteenth Synthesis Imaging Workshop

2012 May 29 - June 5

Juergen Ott (NRAO)

NRAO · Associated Universities, Inc. · New Mexico Tech SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY · The University of New Mexico · New Mexico CONSORTIUM

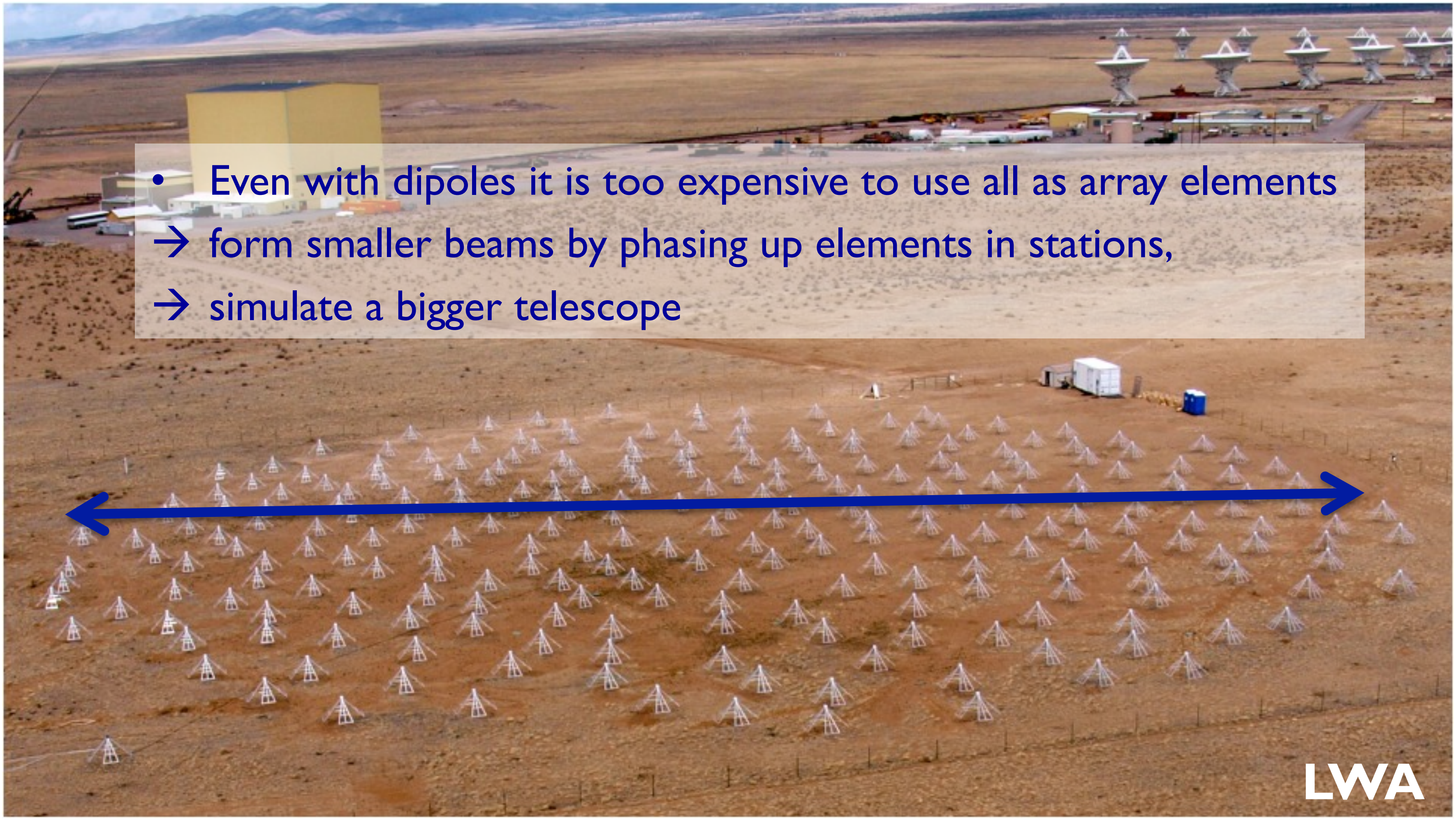
Mosaicing: What is it all about?

- Imaging large regions on the sky, much larger than the primary beam
- The primary beam/field of view depends on the size of an individual dish, not the array configuration
- Mosaicing unlocks some short spacing information
- Is this important? Yes! The entire sky has about 41,253 deg^2
- Primary beams:
  - JVLA (25m dishes): 20cm: 0.25 deg^2, 7mm: 0.0003 deg^2
  - ALMA (12m dishes): band 3 (3mm): 0.02 deg^2, band 9 (650GHz): 0.000005 deg^2
  - Nearby galaxies: M31 (@700kpc) : 3deg^2

    Arp 220 (@70Mpc): 0.004 deg^2
- **Solution 1:** go to smaller dishes (e.g. ATA, 6m dishes @20cm: 6.3 deg^2) but you will need a lot of dishes to gain sensitivity (ATA had planned hundreds)
- **Solution 2:** Mosaicing

Small Dishes: SKA

Dipoles: 2π



- Even with dipoles it is too expensive to use all as array elements

→ form smaller beams by phasing up elements in stations,

→ simulate a bigger telescope

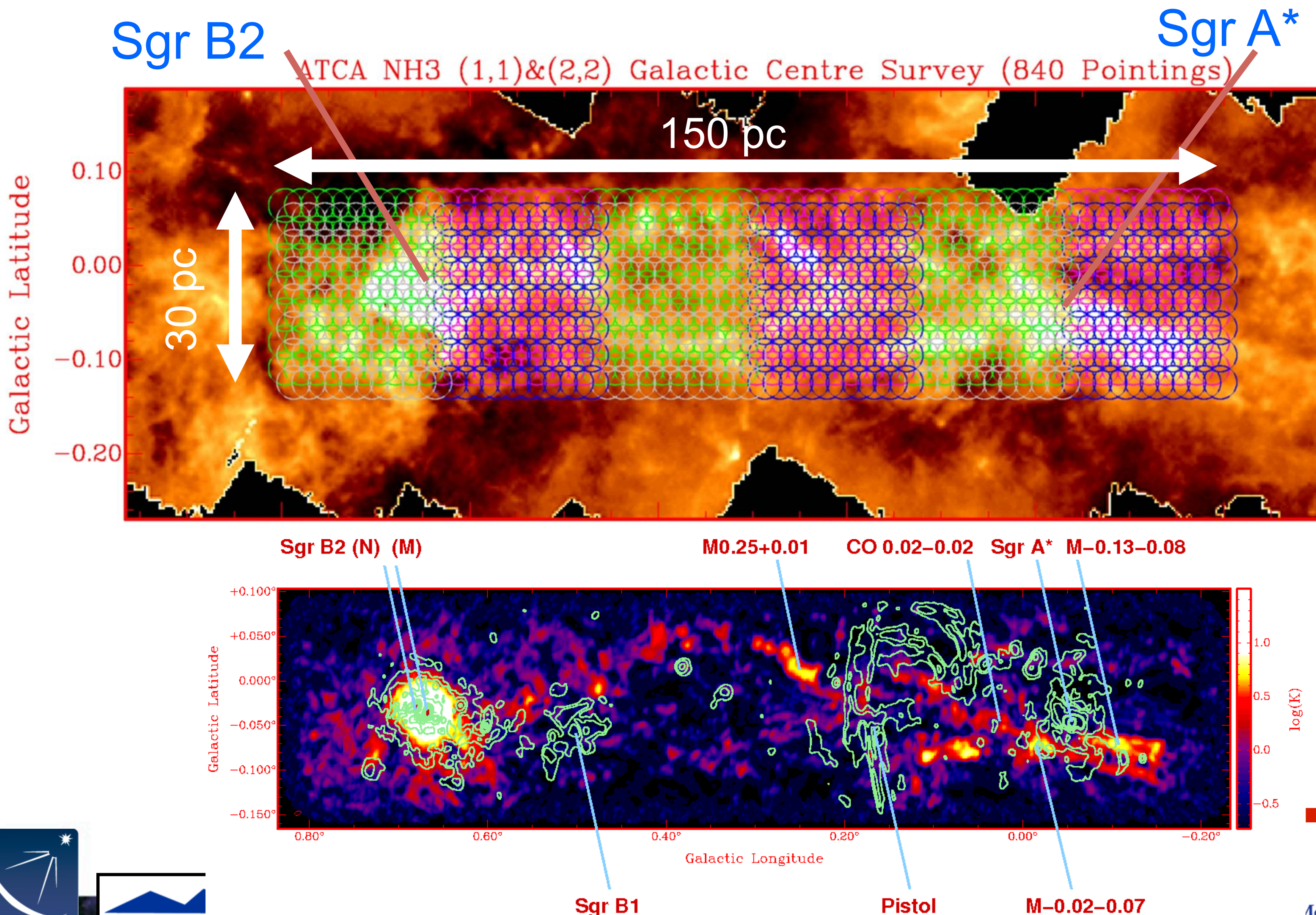
NRAO New Mexico Tech SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY The University of New Mexico New Mexico CONSORTIUM

Mosaicing

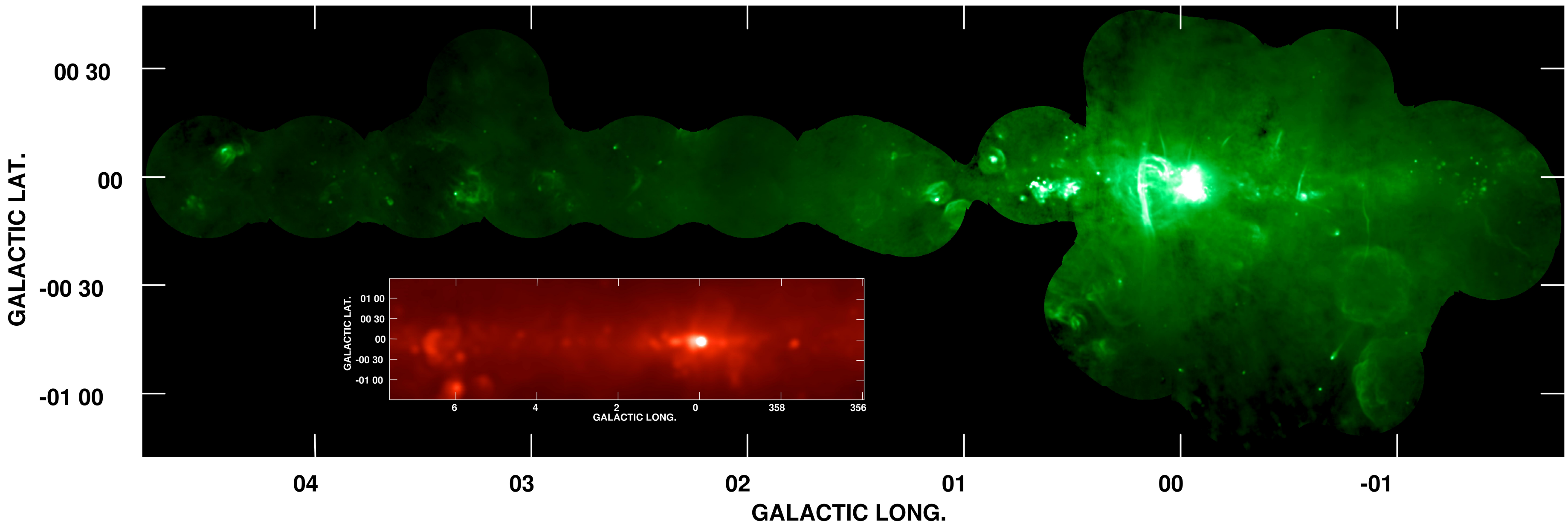
Single Dish Mapping

Interferometric Mosaicing

Galactic Center



Galactic Center



- Yusef-Zadeh et al.

Problems to solve

- Each primary beam is an attenuation which needs to be accounted for
- Pointings are in a time sequence:
  - Each pointing has a different uv-coverage
  - In addition: deconvolution in non-linear, so even identical uv-coverage but at a slightly different part on the sky with different sources in the field will result in a somewhat different image
  - Atmospheric water vapor/Ionospheric variations from pointing to pointing
- Adequate sky coverage: Best uv-coverage? Uniform sensitivity? Maximal sky coverage?
- Minimize drive time but maximize well spaced uv-coverage across map to retain information
- Mosaicing is frequently used for very extended structures, short/zero spacing correction may be required (Ekers-Rots theorem: mosaicing can gain back some of the shorter spacings)

The effect of the Primary Beam

PB defined by single antenna (SD). Not by the array configuration.

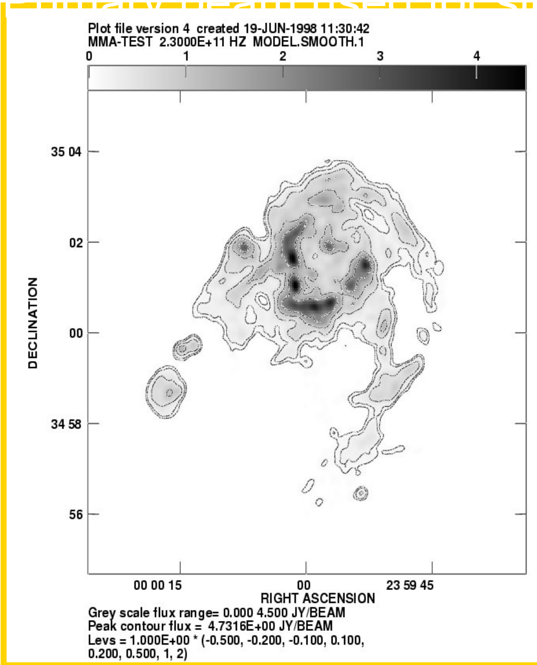
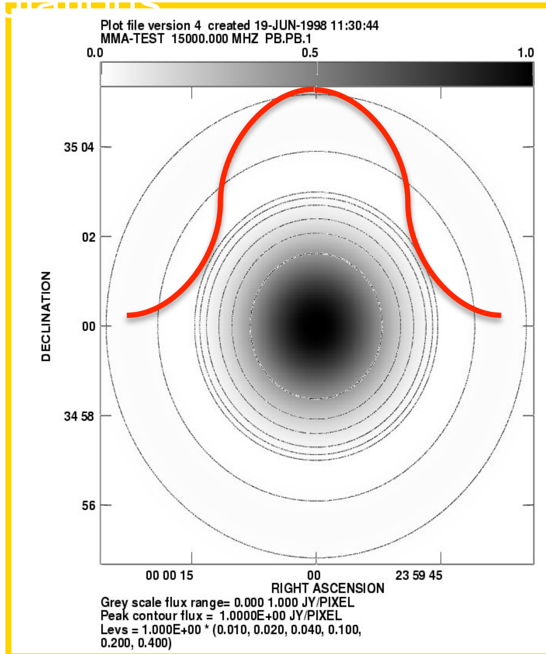
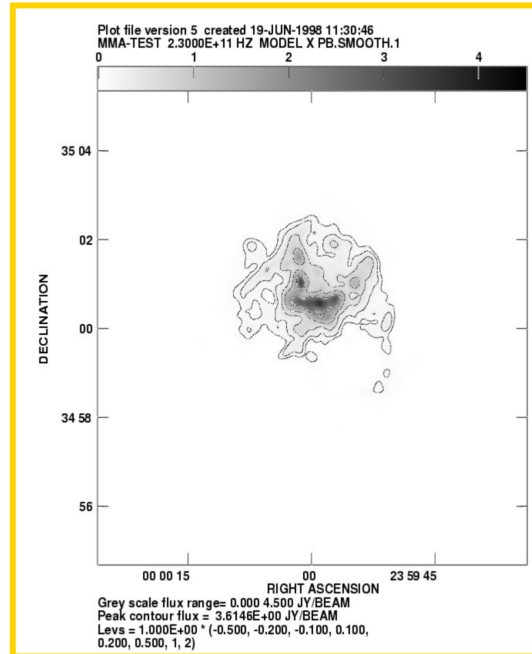


Image larger than PB



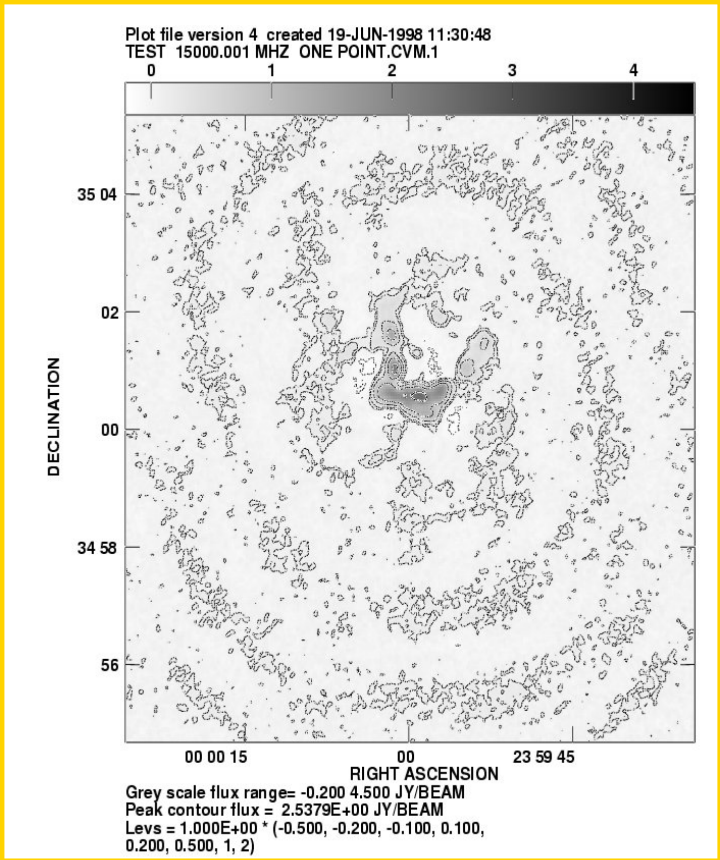
PB provides sensitivity pattern on sky



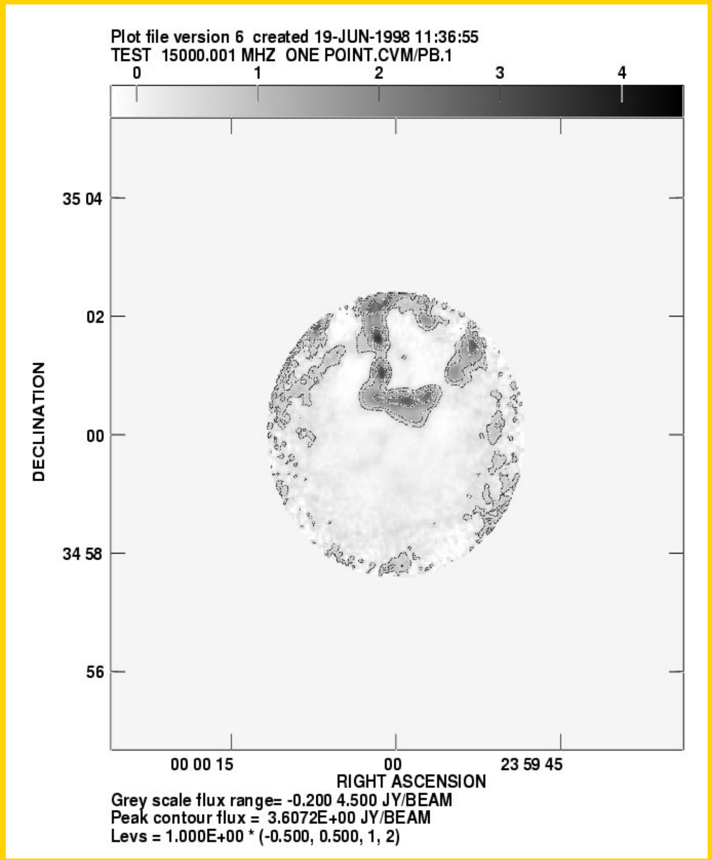
PB applied: sensitive to center only

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

The effect of the Primary Beam



Noise before PB correction



PB correction changes noise characteristics

Stitching the maps together: Image Reconstruction

Widely-used methods for mosaic image reconstruction:

- **Linear combination**

  Map points individually → deconvolve individually → combine

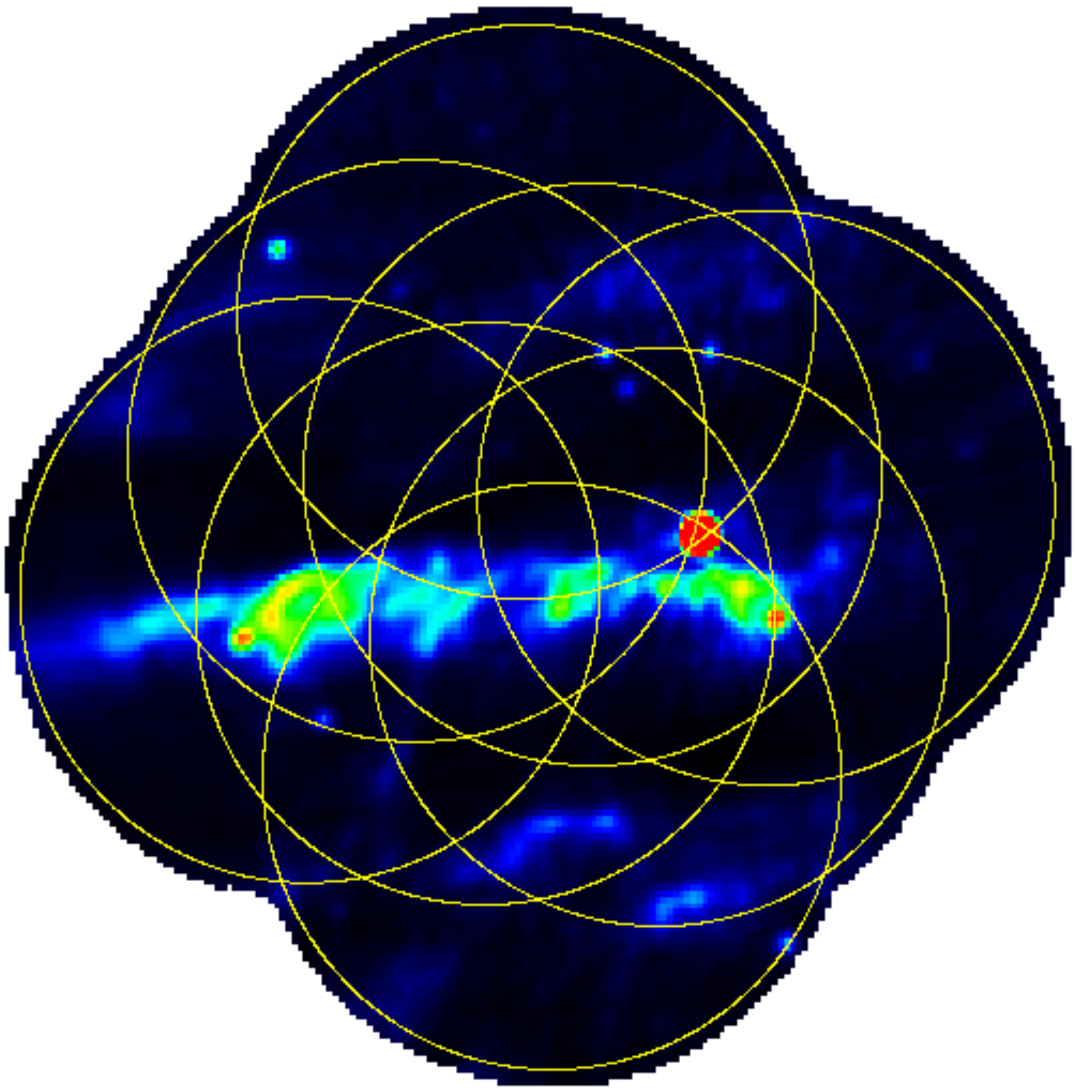
- **Joint deconvolution**

  Map points individually → combine → deconvolve together

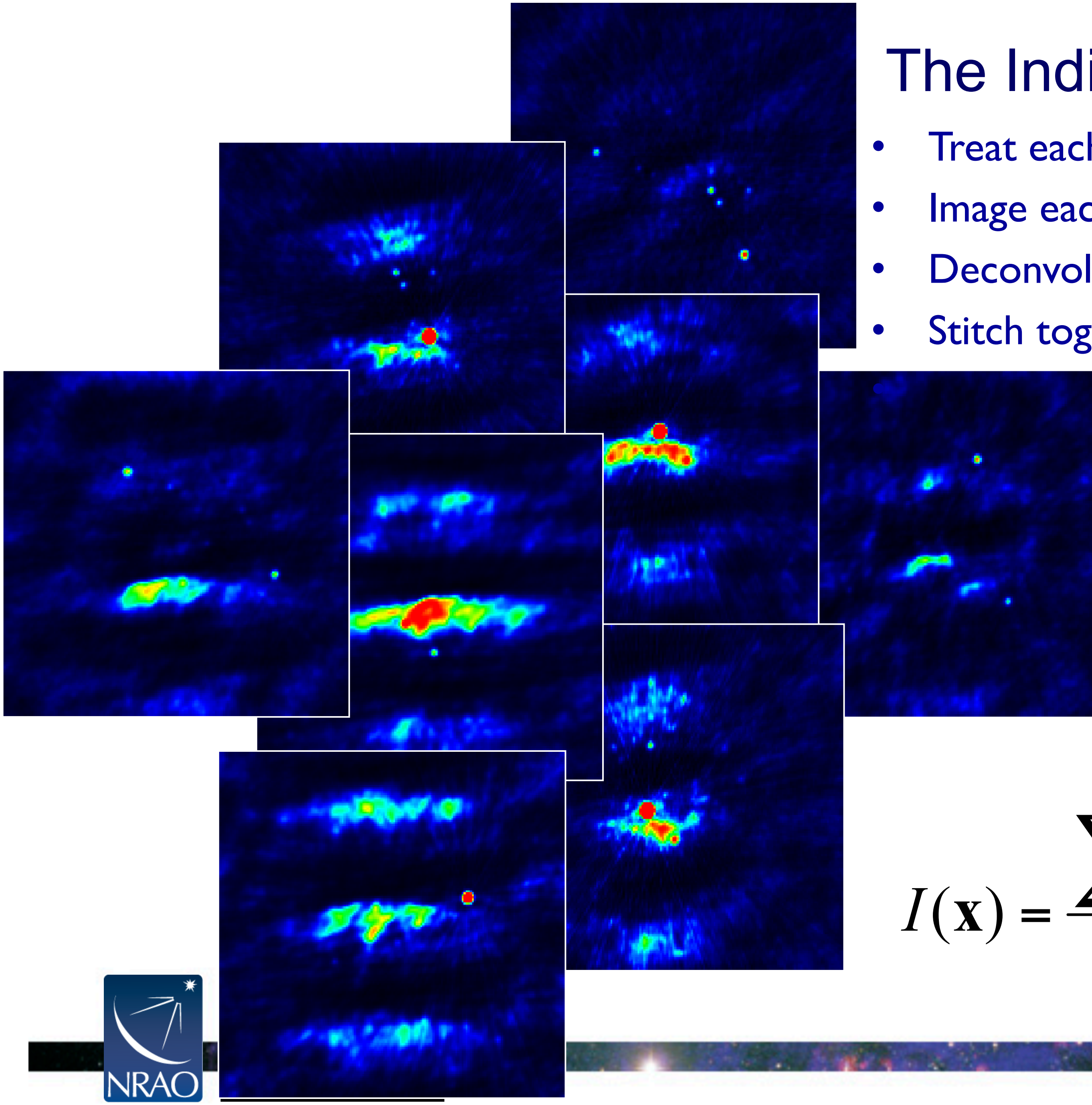
- **Widefield Imaging by regridding of all visibilities before FFT into a single map**

  Combine pointings in uv-space → single map → deconvolve

Mosaicing: Linear Combination of Images



The University of New Mexico New Mexico CONSORTIUM

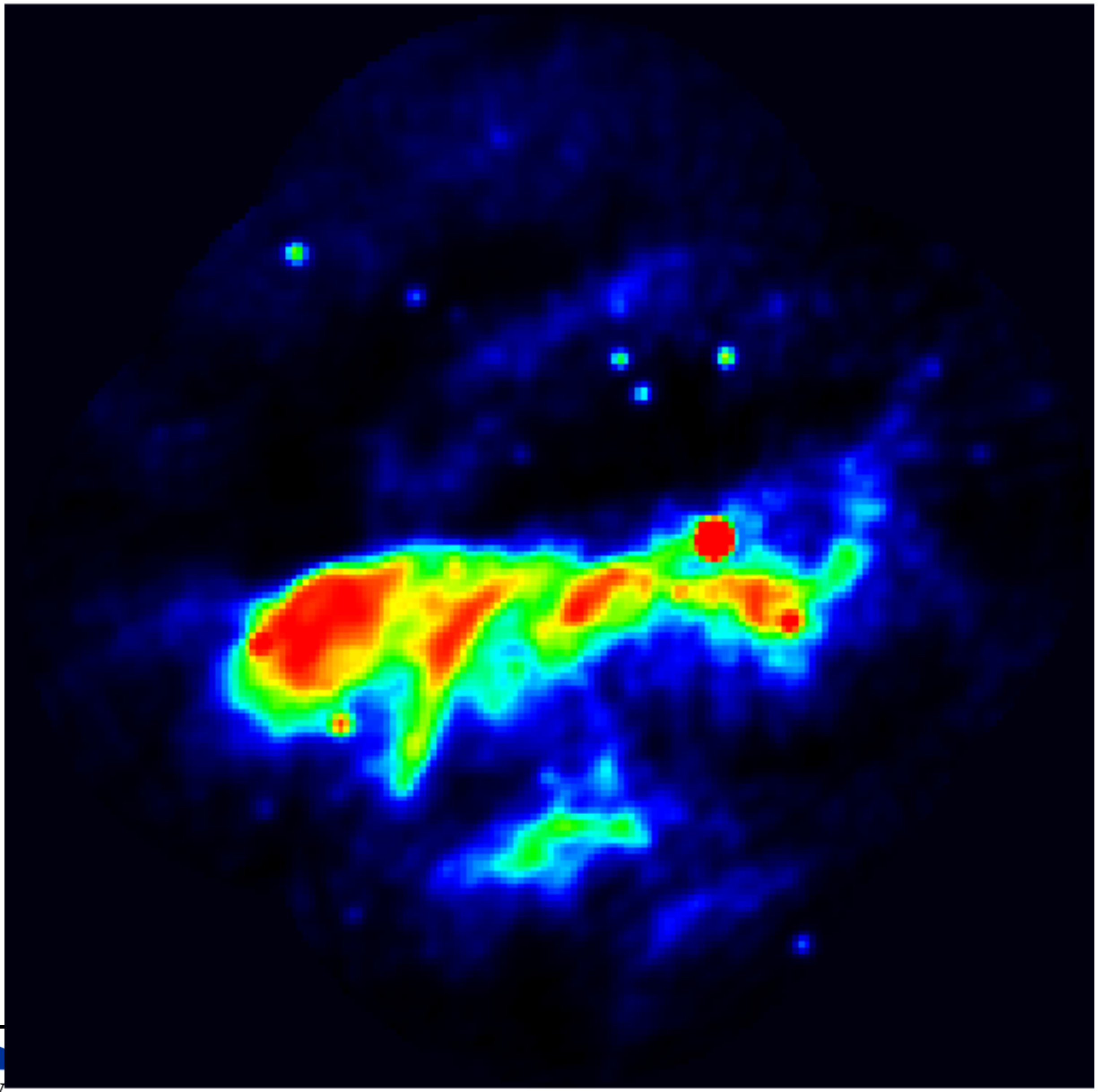


The Individual Approach

- Treat each pointing separately
- Image each pointing
- Deconvolve each pointing
- Stitch together linearly with weights for primary beam

$$I(\mathbf{x}) = \frac{\sum_p A(\mathbf{x}-\mathbf{x}_p) I_p(\mathbf{x})}{\sum_p A^2(\mathbf{x}-\mathbf{x}_p)}$$

Mosaicing: Linear Combination of Images

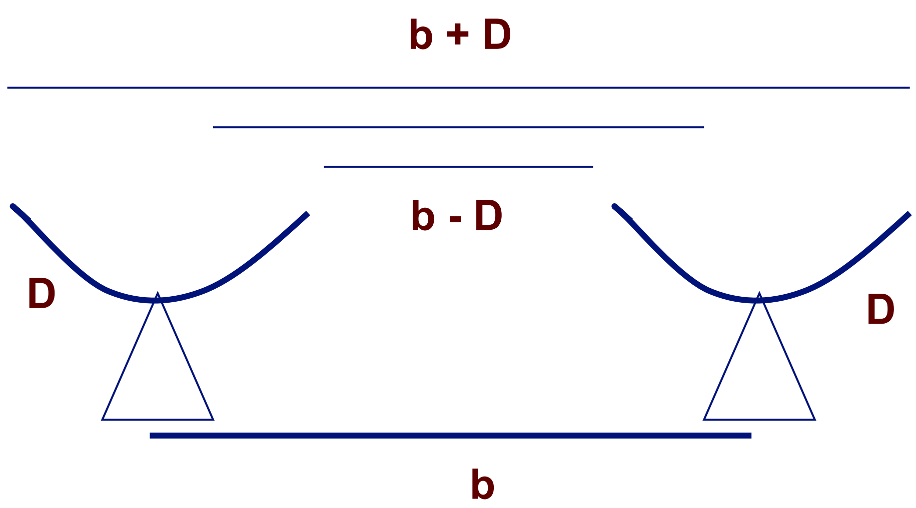


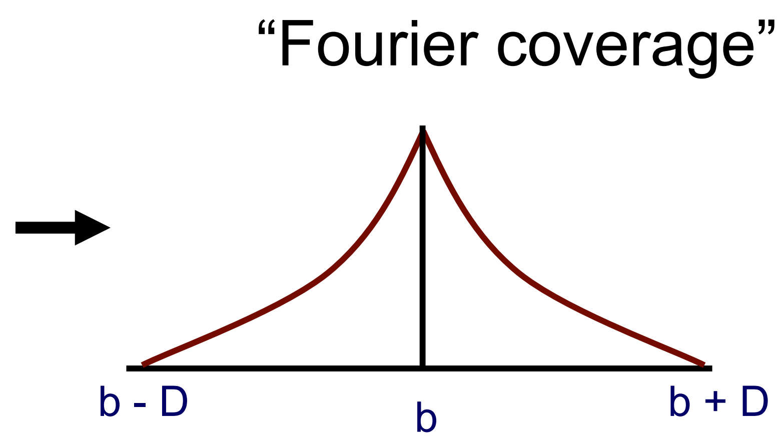
Mosaicing: Linear Combination of Images

- Most straightforward method to create a mosaic map
- But deconvolution is **non-linear**
- Artifacts, in particular at edges may creep in
- Does not use the improved uv-coverage of overlapping regions (and does not take advantage of Ekers-Rots effect)
- Might be the best solution for high-dynamic range imaging
  - It is possible to manipulate every pointing extensively (e.g. solve for off-axis gains, like ‘peeling’; or different deconvolution parameters for different pointings)
  - Depends less on the exact knowledge of primary beam shape when massively oversampled

Ekers & Rots Theorem

- An interferometer doesn't just measure angular scales $\theta = \lambda / b$ it actually measures $\lambda / (b - D) < \theta < \lambda / (b + D)$

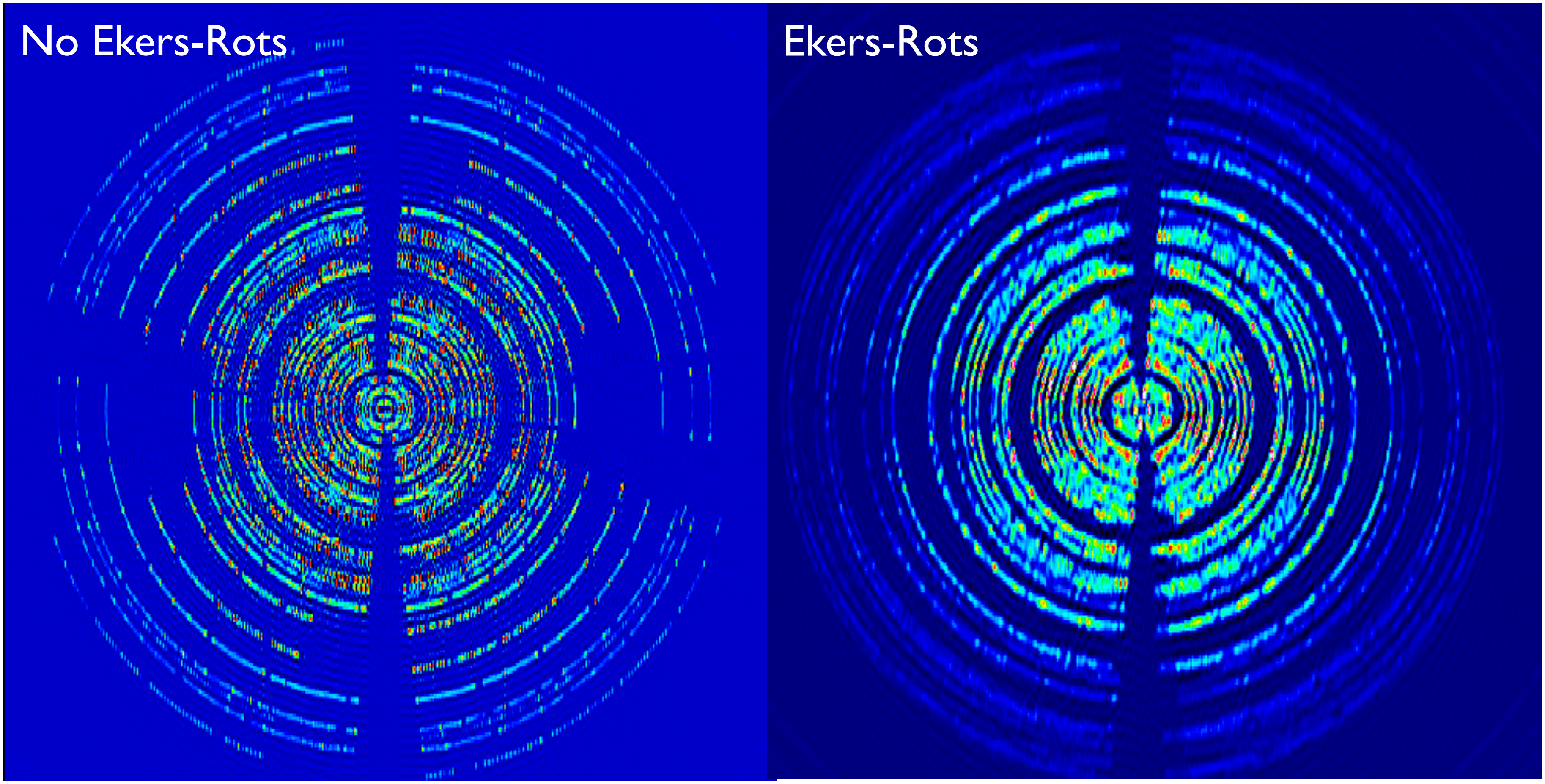




Ekers & Rots Theorem

- Sounds great but you can't get all that extra info from a single visibility
- Interferometer measures a number per baseline and integration time, not a range
- Similar to a single dish, you have to scan to get the extra "spacings"
- Mosaicing is a way to perform this scanning and unlocks the extra information
- The sampling theorem states that the maximum gain is by sampling the sky with a regular, at least Nyquist spaced grid
- **Ekers-Rots is equivalent to a convolution of the FT of the primary beam with the interferometric visibilities in the uv plane for Nyquist sampled data**

Comparison of *u-v* coverage



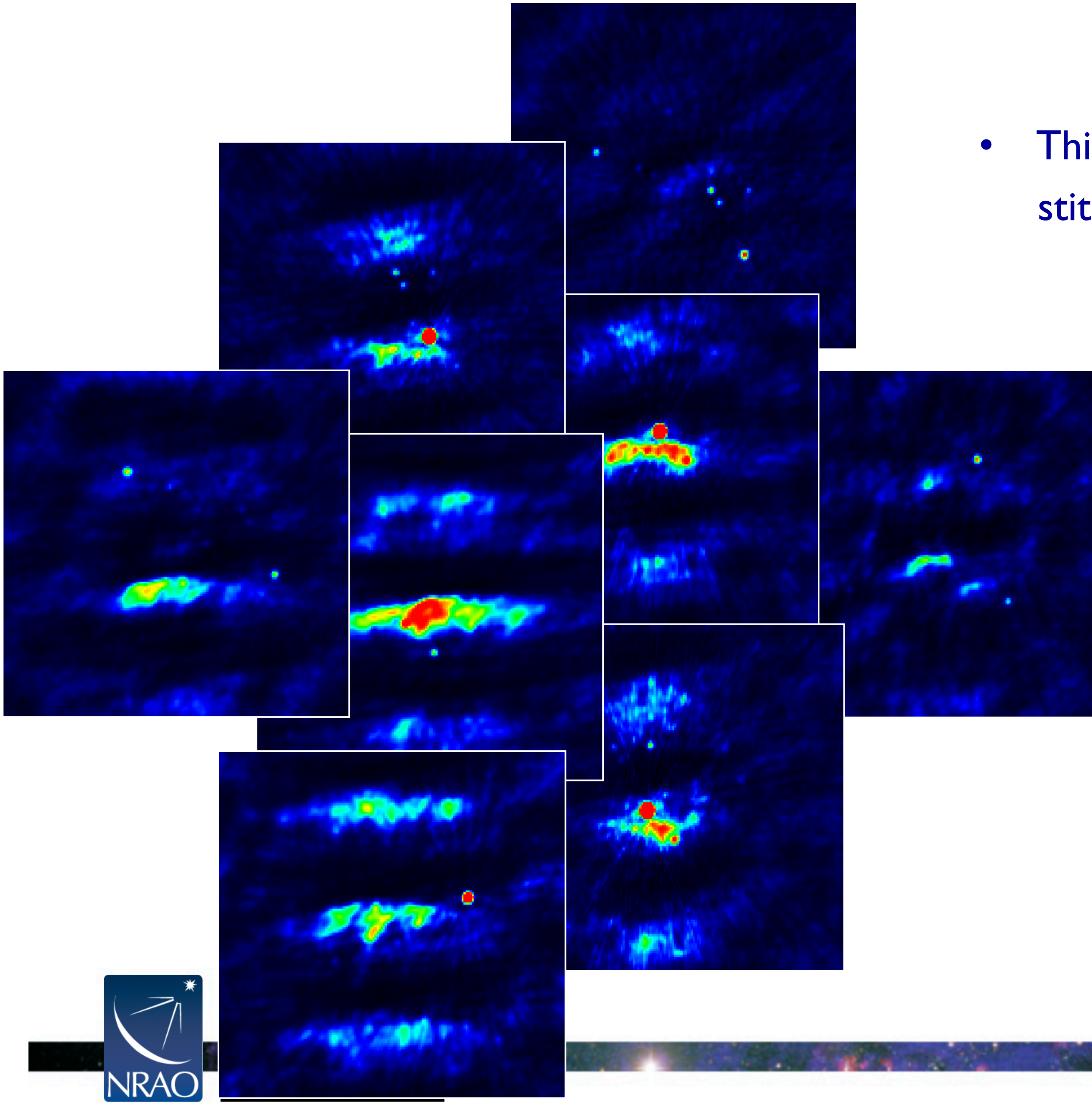
Joint Deconvolution

- Form a linear combination of the individual pointings, p on **DIRTY IMAGE**:

$$I(\mathbf{x}) = W(\mathbf{x}) \frac{\sum_p A(\mathbf{x}-\mathbf{x}_p) I_p(\mathbf{x}) / \sigma_p^2}{\sum_p A^2(\mathbf{x}-\mathbf{x}_p) / \sigma_p^2}$$

- Here σ_p is the noise variance of an individual pointing and A(**x**) is the primary response function of an antenna (primary beam)
- W(**x**) is a weighting function that suppresses noise amplification at the edge of mosaic
- correction for W-projection effects is also required

Sault, Staveley-Smith, Brouw (1996)



- This time: dirty images to be stitched together

Mosaicing: Joint Approach

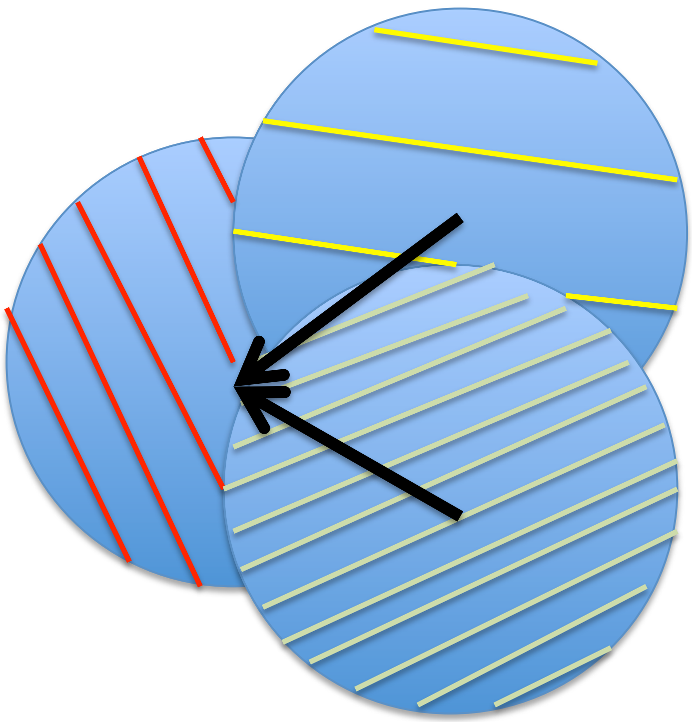
- Joint dirty beam depends on antenna primary beam, ie weight the dirty beam according to the position within the mosaiced primary beams:

$$B(\mathbf{x};\mathbf{x}_0) = W(\mathbf{x})\frac{\sum_p A(\mathbf{x}_0 - \mathbf{x}_p)B_p(\mathbf{x}-\mathbf{x}_0)/\sigma_p^2}{\sum_p A^2(\mathbf{x}-\mathbf{x}_p)/\sigma_p^2}$$

- Uses all *uv* data from all points for the beam simultaneously
  - Combined beam provides better deconvolution in overlap regions
  - Adds in Ekers & Rots spacings: more structure recovered, better beam
  - Overlapping pointings require good knowledge of PB shape further out than the half power point
  - Applies W-projection

Widefield Imaging

- Take each uv data for each pointing and regrid to a common phase reference center



Widefield Imaging

- Next Step: Perform weighting for primary beam(s)
- Multiplication in image domain = convolution in FT domain

$$I = \sum_p A_p I_p \quad \text{(+ weighting terms)}$$

$$FT\{I\} = \sum_p FT\{A_p\} \otimes FT\{I_p\}$$

- The PBs for each pointing are identical but shifted
- FT of a shift is a phase gradient
- Sum of Phase gradient for each offset pointing * single FT{A} is the weighting for each visibility to correct for primary beams
- FT to single image with a common synthesized "dirty beam"

- The widefield regridding method is also key when dealing with large number of pointings, e.g. in (future) on-the-fly interferometry

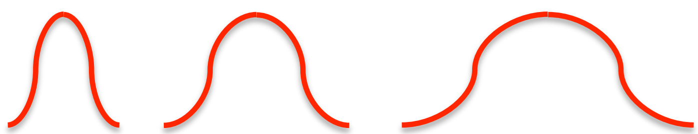
Deconvolution

- Mosaics can be lots of point like sources but typically are performed for extended emission

- **CLEAN:** subtract dirty beam (point sources) from dirty image

  (Preferably Cotton-Schwab, with small gain; FFT of major cycle will reduce sidelobes)

- **Multiscale clean**: Use a number of kernel sizes for different scales



- **Maximum entropy**: iterate on minimizing χ^2 between

  data and a model, fit for maximum smoothness

Mosaicing in CASA

- Most of the tricky techniques are performed under the hood for your convenience
- Calibrate as you would do for a single pointing
- Use the **clean** task with your favorite parameters
- In *imagermode* use *'mosaic'*
- Use *ftmachine='ft'* for joint deconvolution, *'mosaic'* for the widefield imaging
- Use *psfmode='clark'* for Cotton-Schwab Algorithm
- Fill in *'multiscale'* parameters (scales) for MS Clean
- Maximum Entropy and linear mosaicing of cleaned images are only available from the CASA toolkit at this moment

- Contributed tasks for mosaicing setups – also check ALMA OT/JVLA OPT

Practical Considerations

- What grids to use?
- How often to come back to a individual pointing
- Slew time of Antennas
- Change of atmospheric conditions

Practical Consideration: Sensitivity/Primary Beam/Pointing

- Pointings overlap which increases the sensitivity per position. For a fixed time observation the total noise is

$$\sigma_t \sim \sigma_p \sqrt{n} / 1.4$$

where n is the number of pointings

- Mosaicing requires a good model of the primary beam in particular for less-dense packings

- For wide band mosaicking best to use the highest frequency for your pointing positions but some compromise is possible; for less dense packings W-projection becomes more important

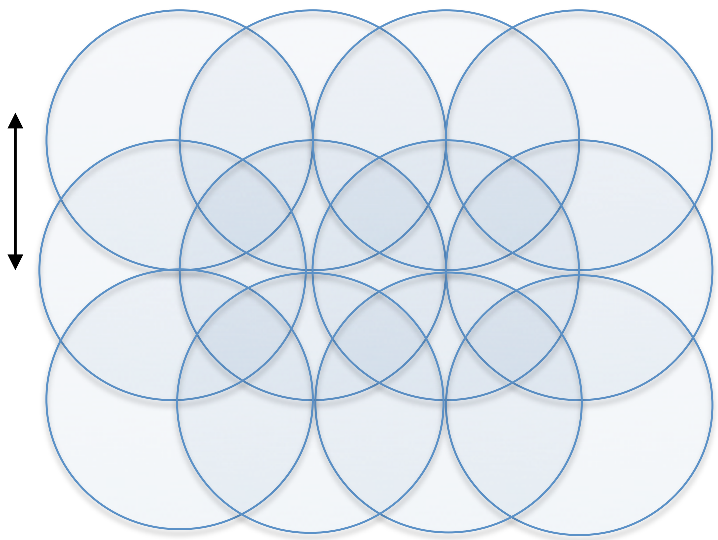
- Pointing errors are first order in mosaics (only second order in single pointing observations of sources smaller than primary beam)

Practical Consideration: Choice of Grid

- Different ways to layout the grid on the sky:
- Nyquist sampling:

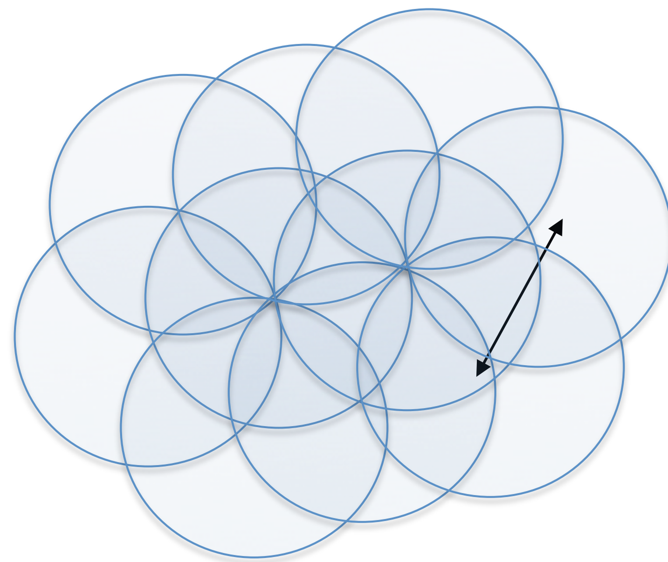
Rectangular grid

$$\frac{\lambda}{2D}$$



Minimum Nyquist for structure information recovery

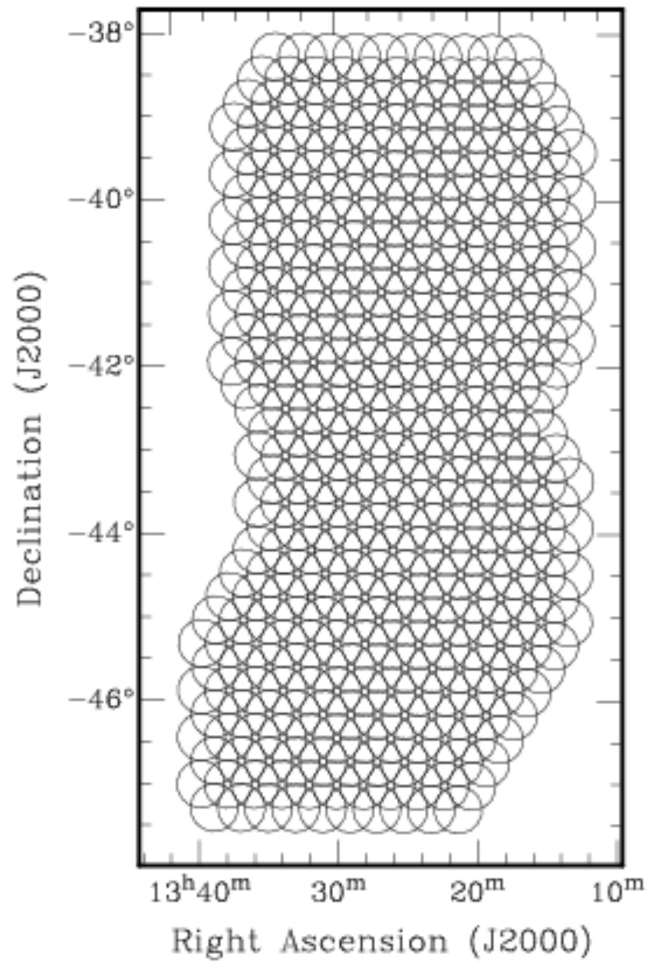
Hexagonal grid

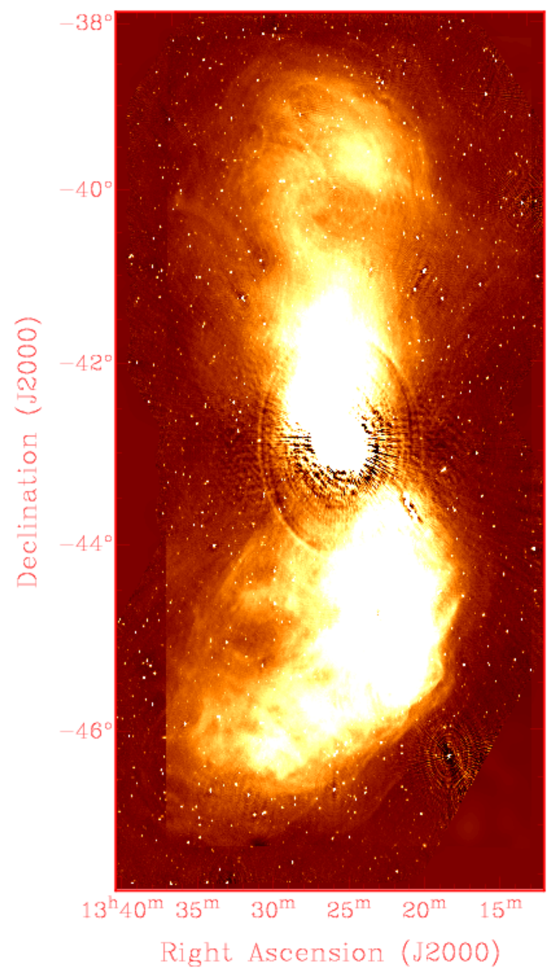


$$\frac{2}{\sqrt{3}}\frac{\lambda}{2D}$$

Oversampled but every position at least covered three times, best for uniform noise, but many pointings are needed

Centaurus A: 406 pointings



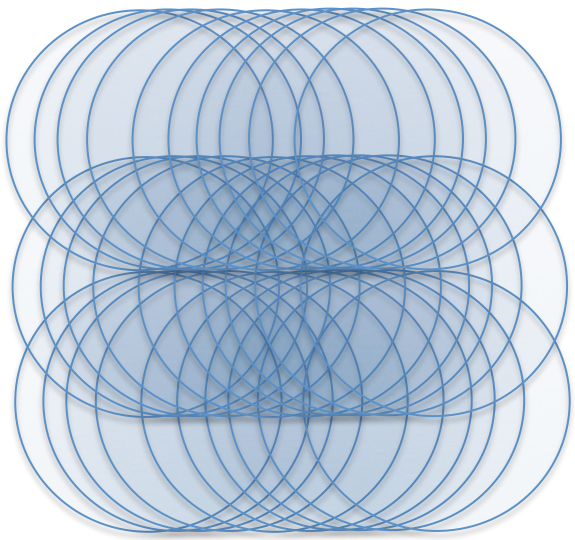


Feain et al. 2011

Practical Consideration: Choice of Grid

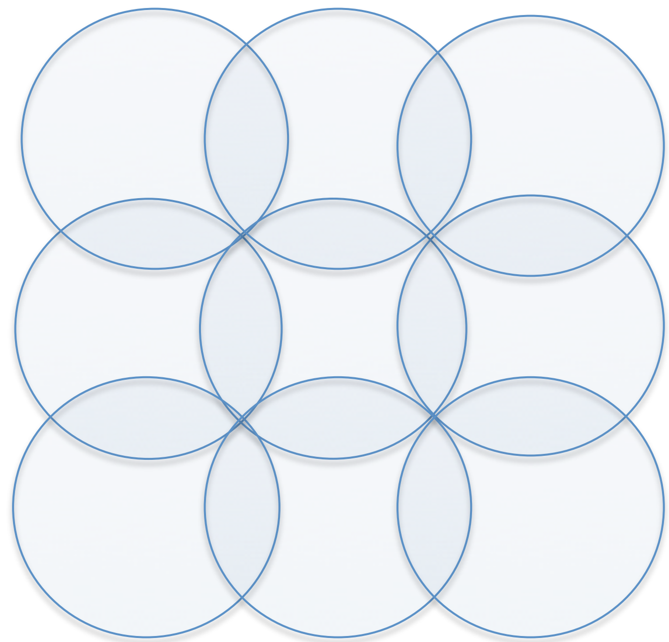
- On-The-Fly Interferometry
- Non-Nyquist sampling

OTF



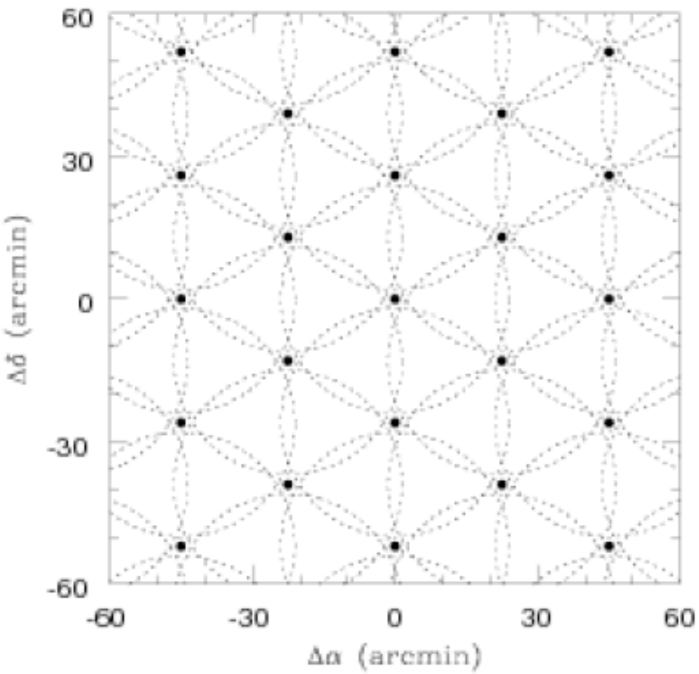
Scan does not stop, fast dumping of data, influences the primary beam shape, produces lots of data but reduces overhead

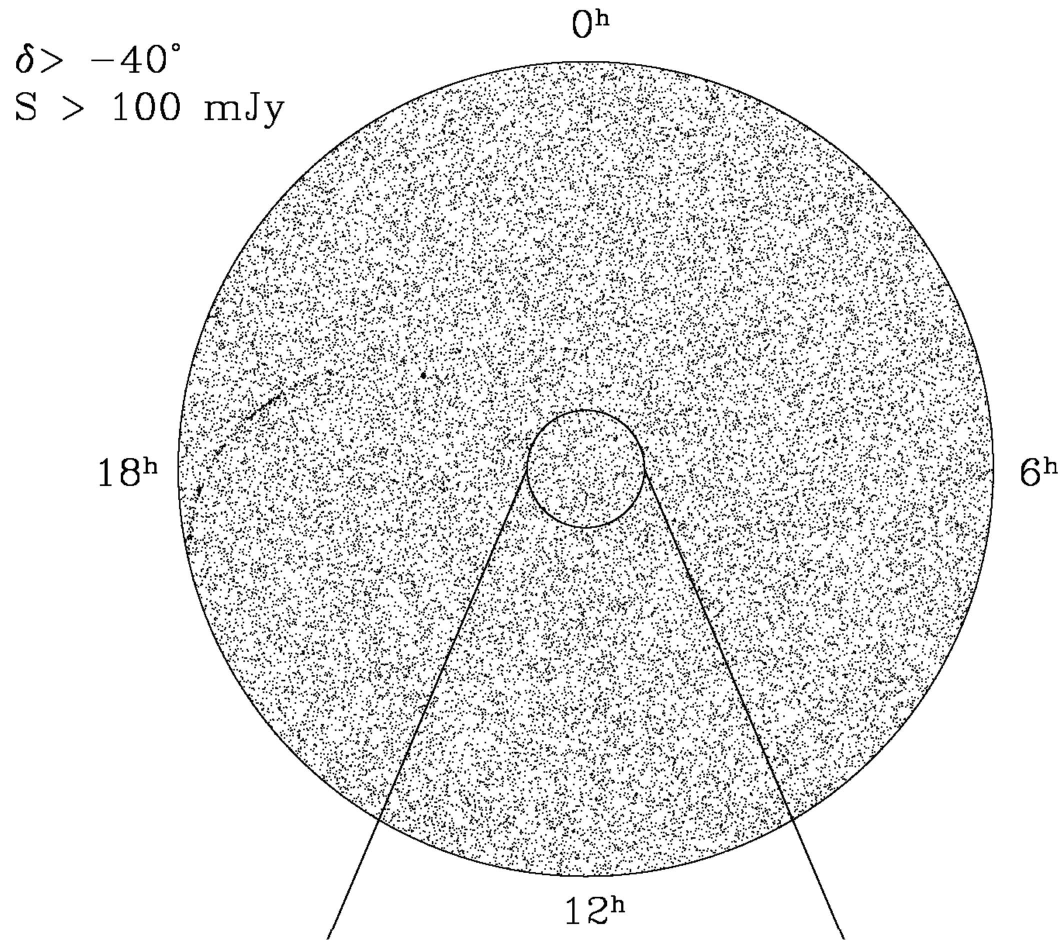
Non-Nyquist



Basic, fast sky coverage

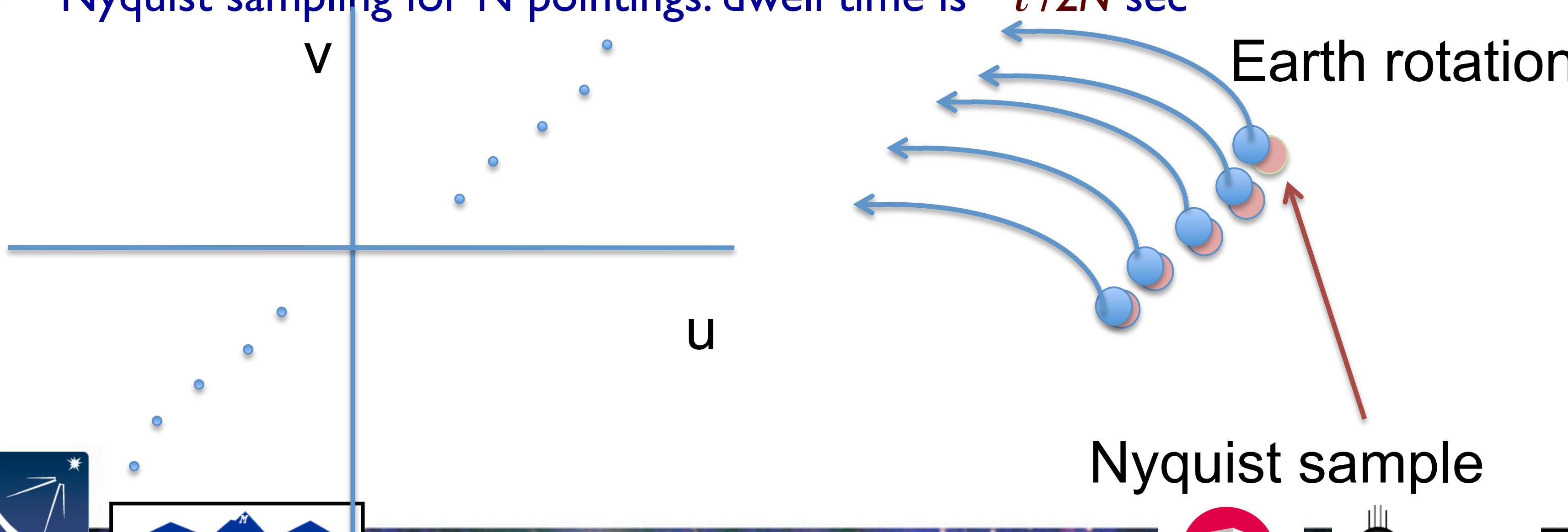
NVSS: 217,446 pointings





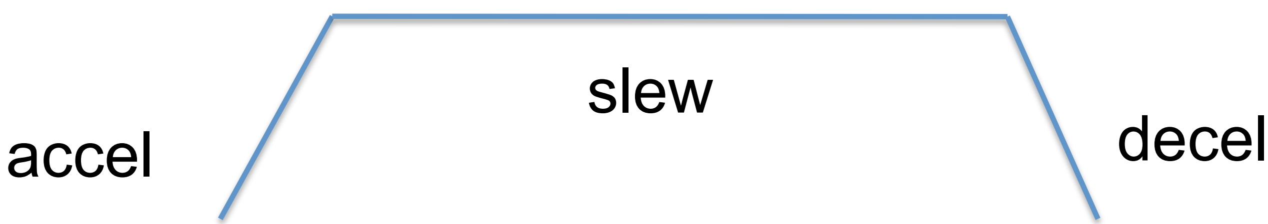
Practical Considerations: Nyquist *uv* structure sampling

- How often to visit a pointing?
- One baseline measures region in *uv* plane with size 2*D*
- Want adjacent samples to be completely independent
- At transit, the time between independent points is $\tau = (86400 / 2\pi)(2D / L)$ sec, where D = antenna diameter, L = longest baseline
- Nyquist sampling for N pointings: dwell time is $\tau / 2N$ sec

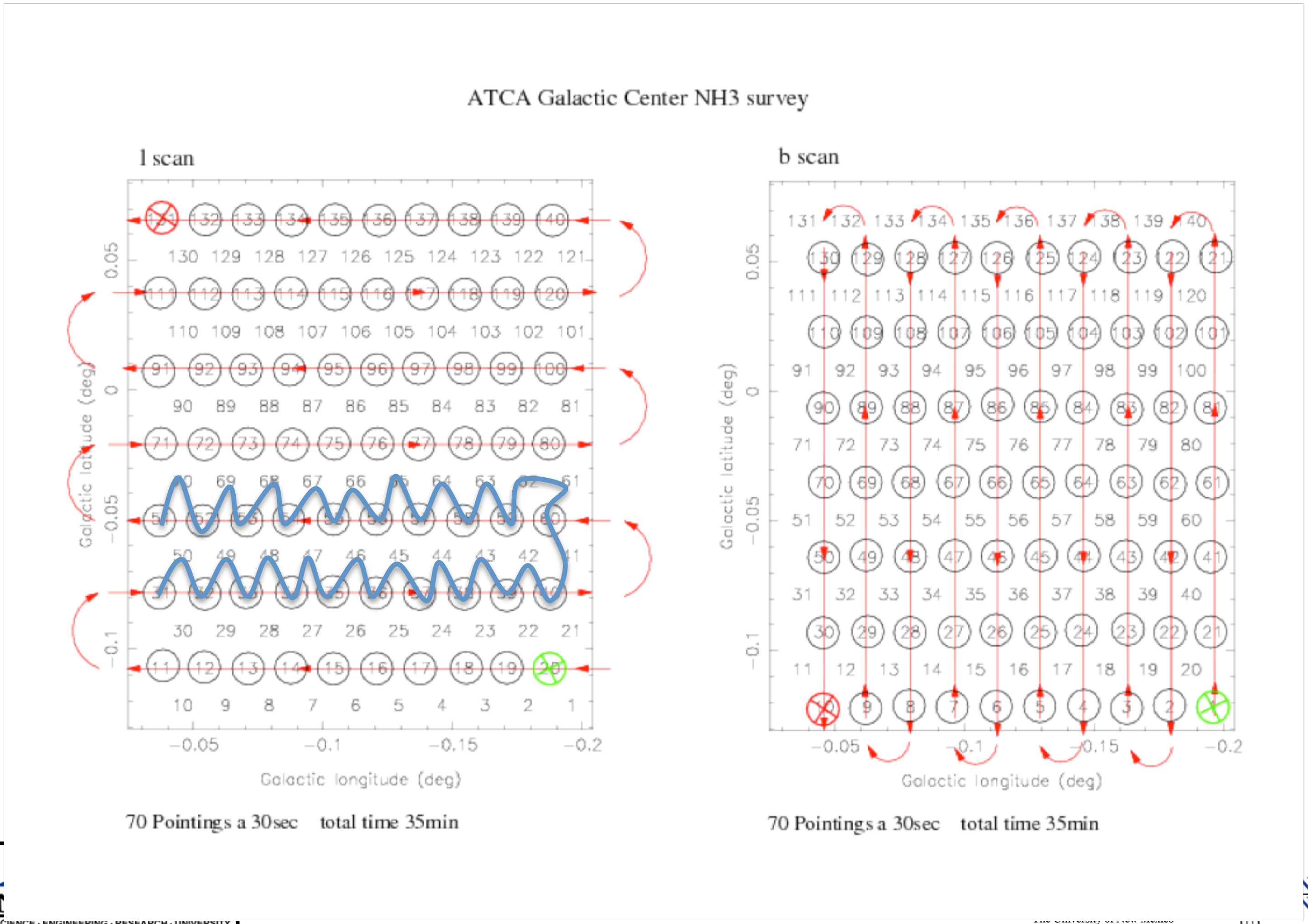


Practical Consideration: Slew Time

- Telescope slew times are calculated by:
- Acceleration
- Constant Slew velocity
- Deceleration
- Settling time
- Some telescopes may have variations in Az and El
- JVLA: acceleration: 2.2 deg s^{-2}, slew rate: 20 deg min^{-1} in El, 40 in Az
- Settling time: ~1-3s shorter in El, longer in Az
- ALMA: acceleration: 24 deg s^{-2}, slew rate 180 deg min^{-1} in El, 360 in Az



Practical Considerations: Slew time vz uv-coverage



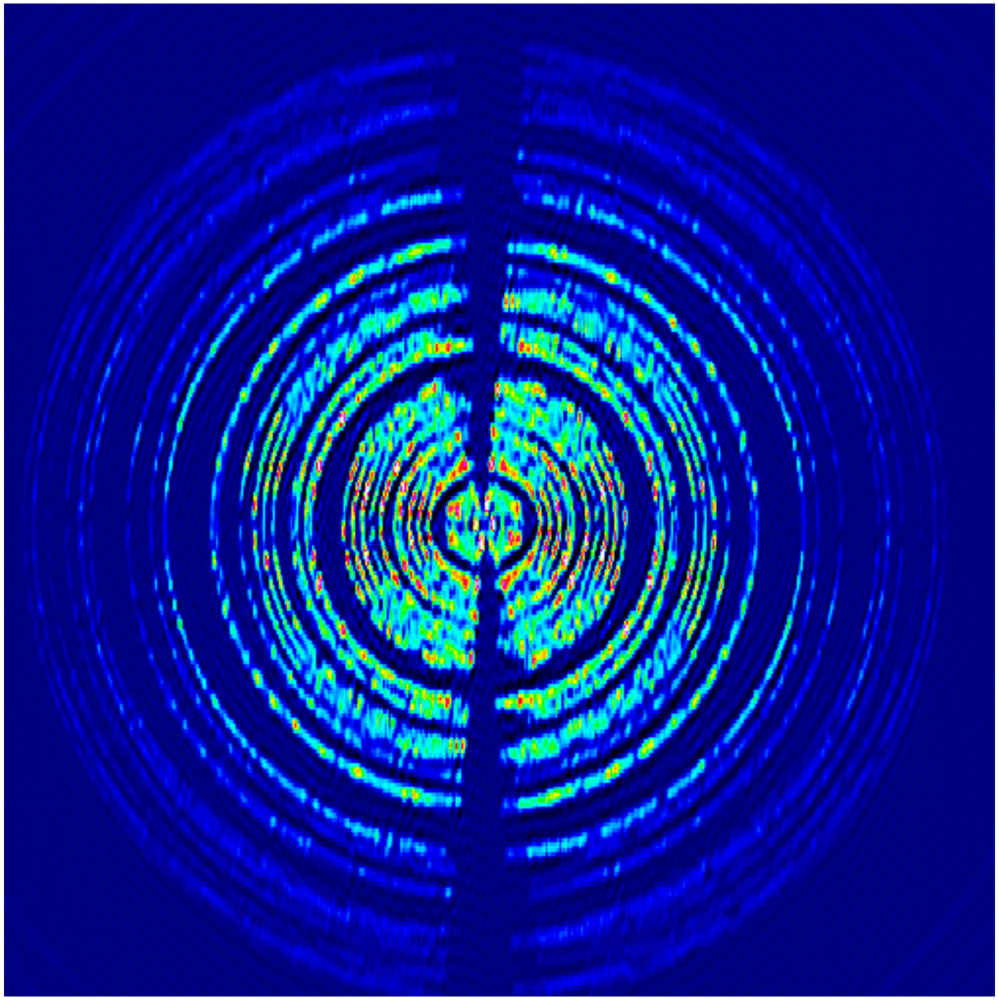
Practical Consideration: Changing Atmosphere (Ionosphere)

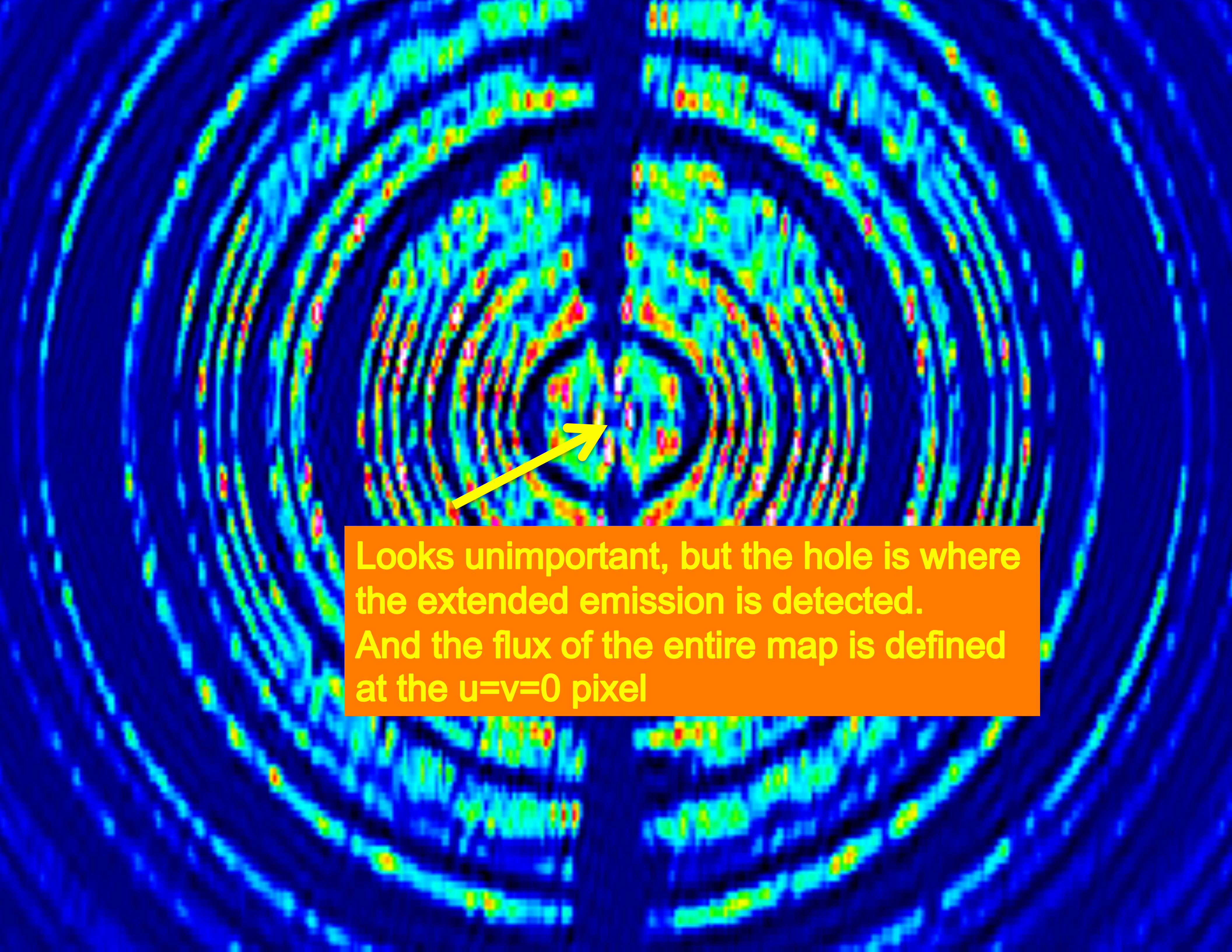
- The water vapor content of the atmosphere can change on small timescales
- In particular there can be large variations in individual cells
  - Changing sky brightness
  - Changing opacity
  - Increased phase noise

- Delay variations due to ionosphere are possible at low frequencies

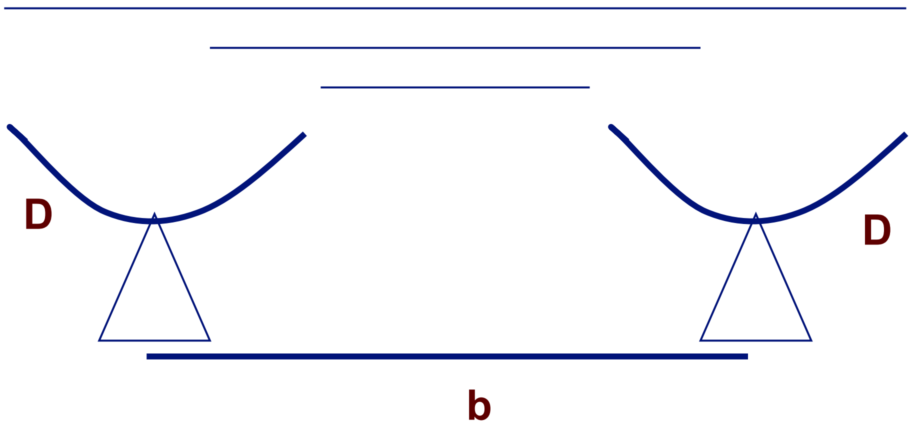
→ Try to cover the full mosaic fast but more frequently
This will make the map more uniform, but it increases your overhead

Short and Zero Spacing Correction




Looks unimportant, but the hole is where
the extended emission is detected.
And the flux of the entire map is defined
at the u=v=0 pixel

uv-coverage



- What's the problem with the hole?
- It's the short baselines
- They define the largest spatial frequencies, or the largest angular scales that an interferometer is sensitive to
- The field of view is given by the beam of a single antenna
- The largest angular scale is given by the shortest distance between 2 antennas $\theta<2\lambda/d_{min}$
- Single antenna diameter < shortest distance
- Field of view > largest sensitive scale
- Extreme: full flux in field of view is given by the central pixel in the uv-coverage

Short spacings & zero spacings problem

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

JVLA
largest angular scale

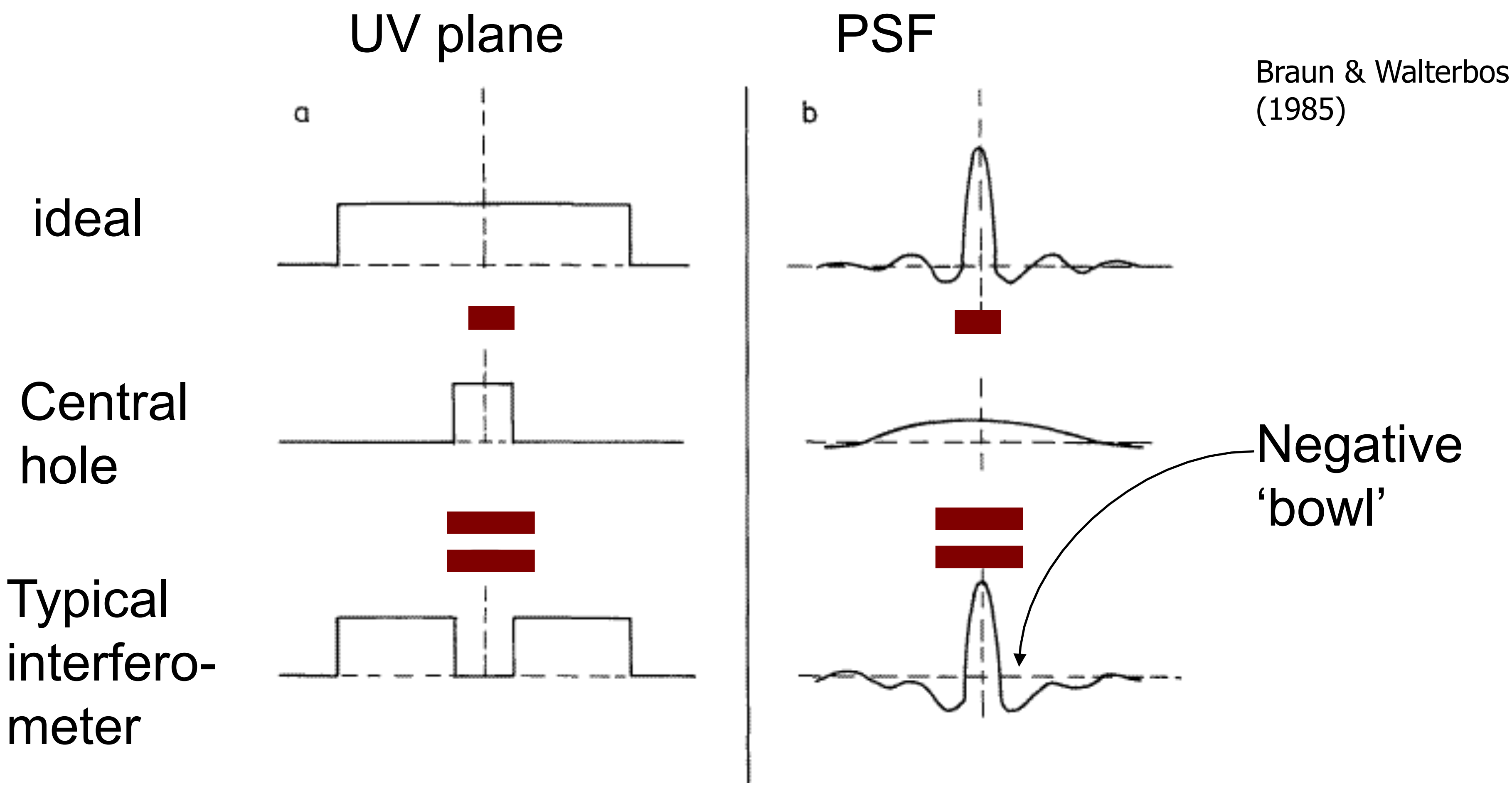
- Primary beam / field of view: 45'/ν (GHz)

Largest angular scale < field of view

Table 5: Configuration Properties

	Configuration	A	B	C	D
	B_{max} (km[1])	36.4	11.1	3.4	1.03
	B_{min} (km[1])	0.68	0.21	0.035[5]	0.035
		Synthesized Beamwidth θ_{HPBW}(arcsec)[1,2,3]			
	74 MHz (4 band)	24	80	260	850
	1.5 GHz (L)	1.3	4.3	14	46
	3.0 GHz (S)[6]	0.65	2.1	7.0	23
	6.0 GHz (C)	0.33	1.0	3.5	12
	8.5 GHz (X)[7]	0.23	0.73	2.5	8.1
	15 GHz (Ku)[6]	0.13	0.42	1.4	4.6
	22 GHz (K)	0.089	0.28	0.95	3.1
	33 GHz (Ka)	0.059	0.19	0.63	2.1
	45 GHz (Q)	0.043	0.14	0.47	1.5
		Largest Angular Scale θ_{LAS}(arcsec)[1,4]			
36500	74 MHz (4 band)	800	2200	20000	20000
1800	1.5 GHz (L)	36	120	970	970
900	3.0 GHz (S)[6]	18	58	490	490
450	6.0 GHz (C)	8.9	29	240	240
320	8.5 GHz (X)[7]	6.3	20	170	170
180	15 GHz (Ku)[6]	3.6	12	97	97
120	22 GHz (K)	2.4	7.9	66	66
80	33 GHz (Ka)	1.6	5.3	44	44
60	45 GHz (Q)	1.2	3.9	32	32

The short(zero)-spacing problem

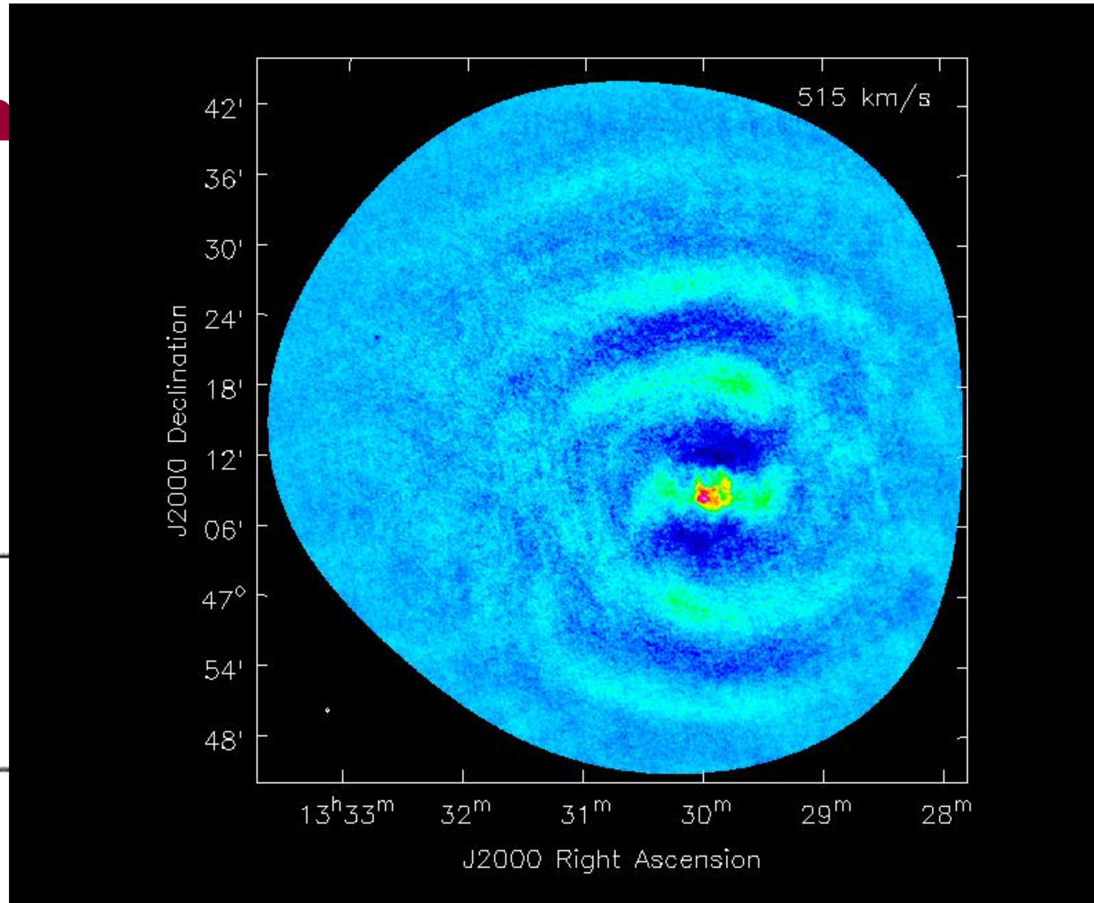


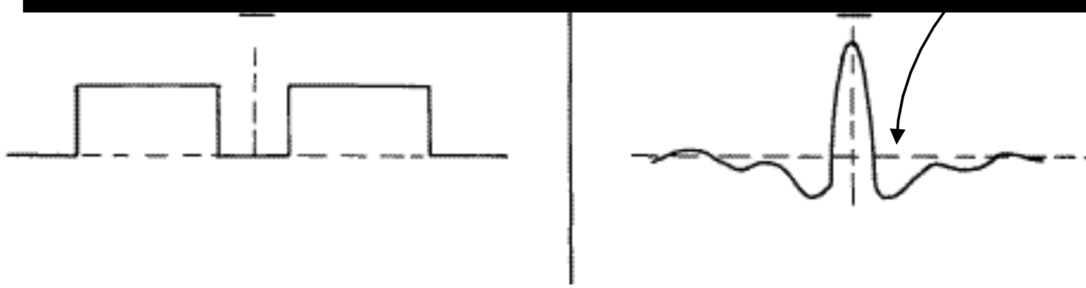
The sh

ideal

Central hole

Typical interfero-meter



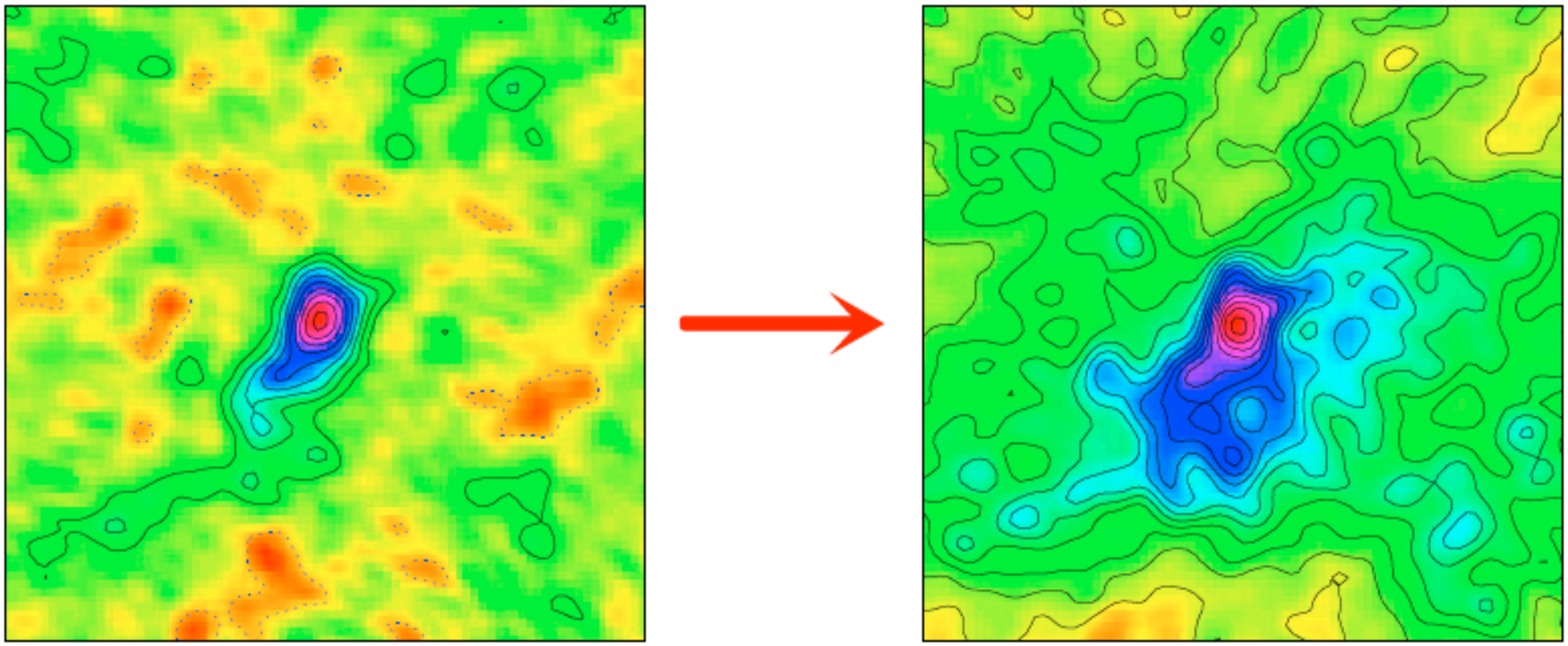


Braun & Walterbos (1985)

Negative 'bowl'

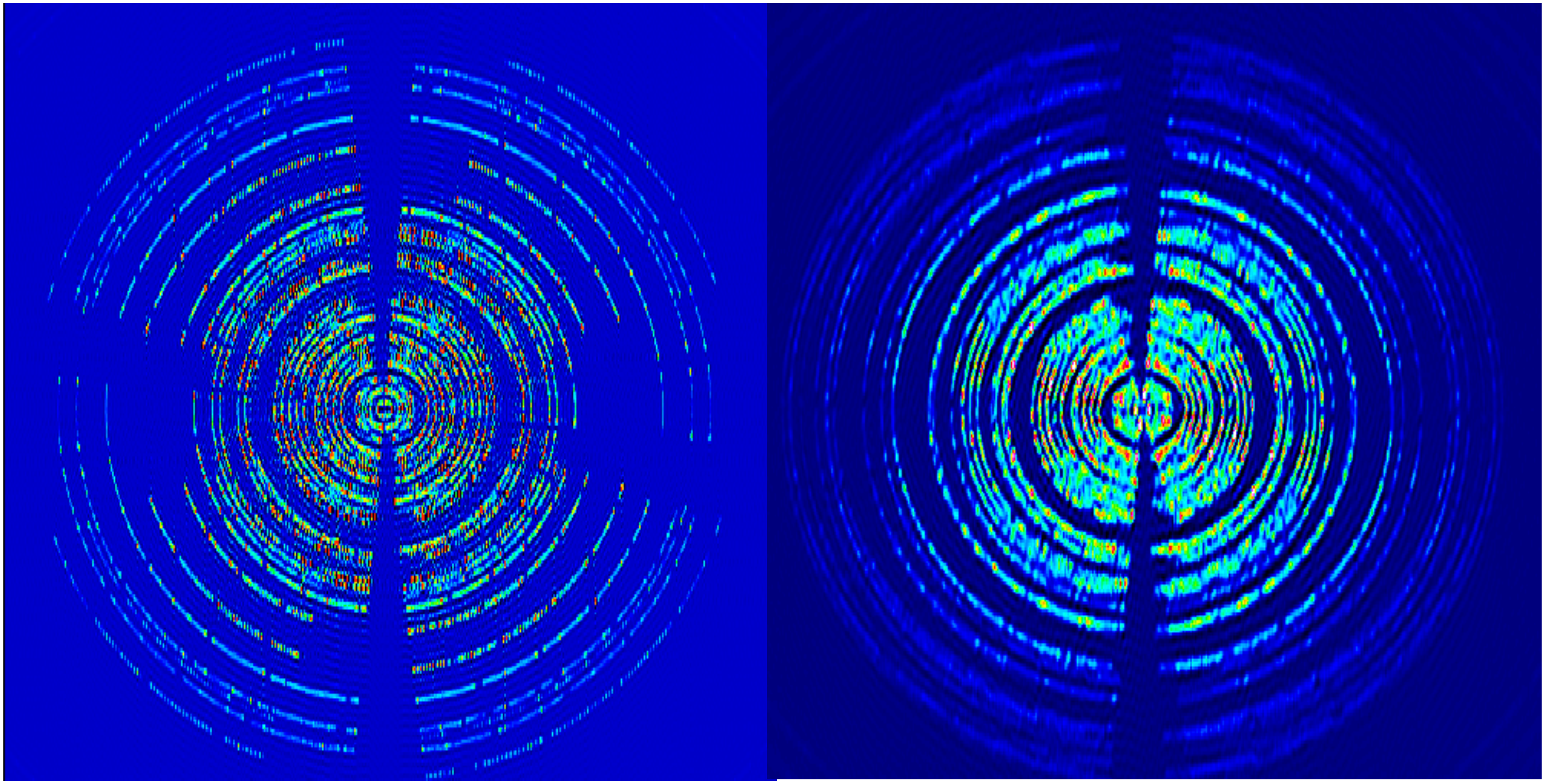
Short Spacings
Example

Without short spacings — With short spacings



^{13}CO (1–0) in the L 1157 protostar (Gueth et al. 1997)

Ekers-Rots helps



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

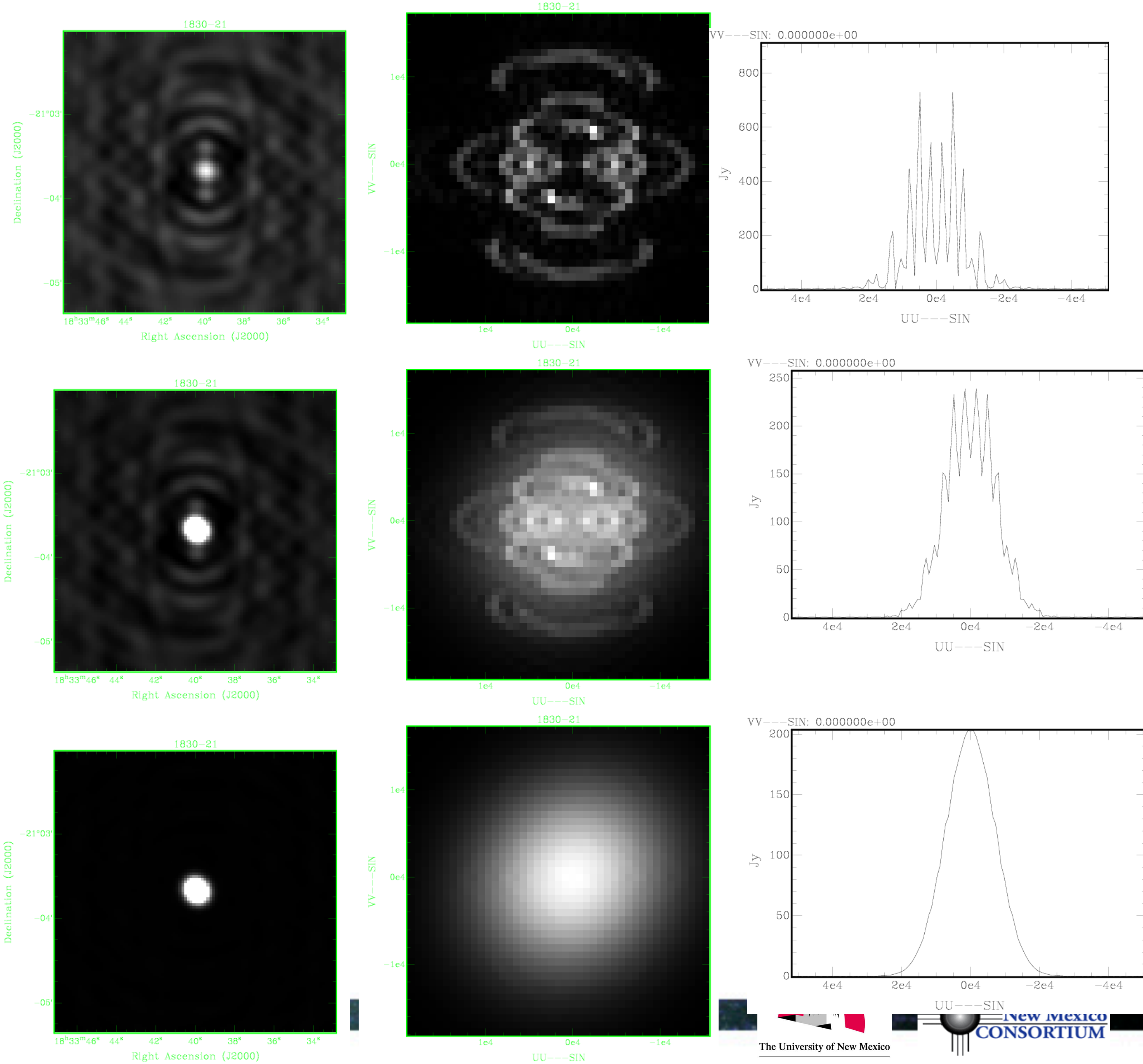
CLEAN extrapolates short spacings

FFT | FFT-slice

dirty

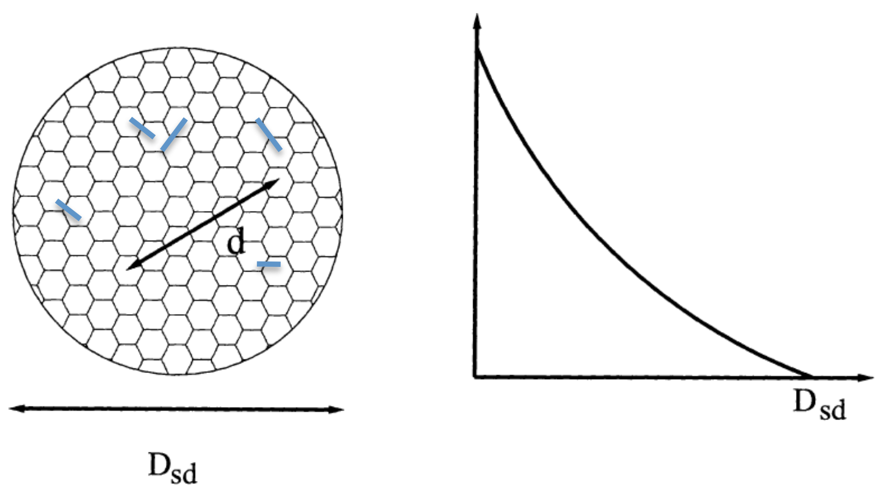
Few iterations

more iterations



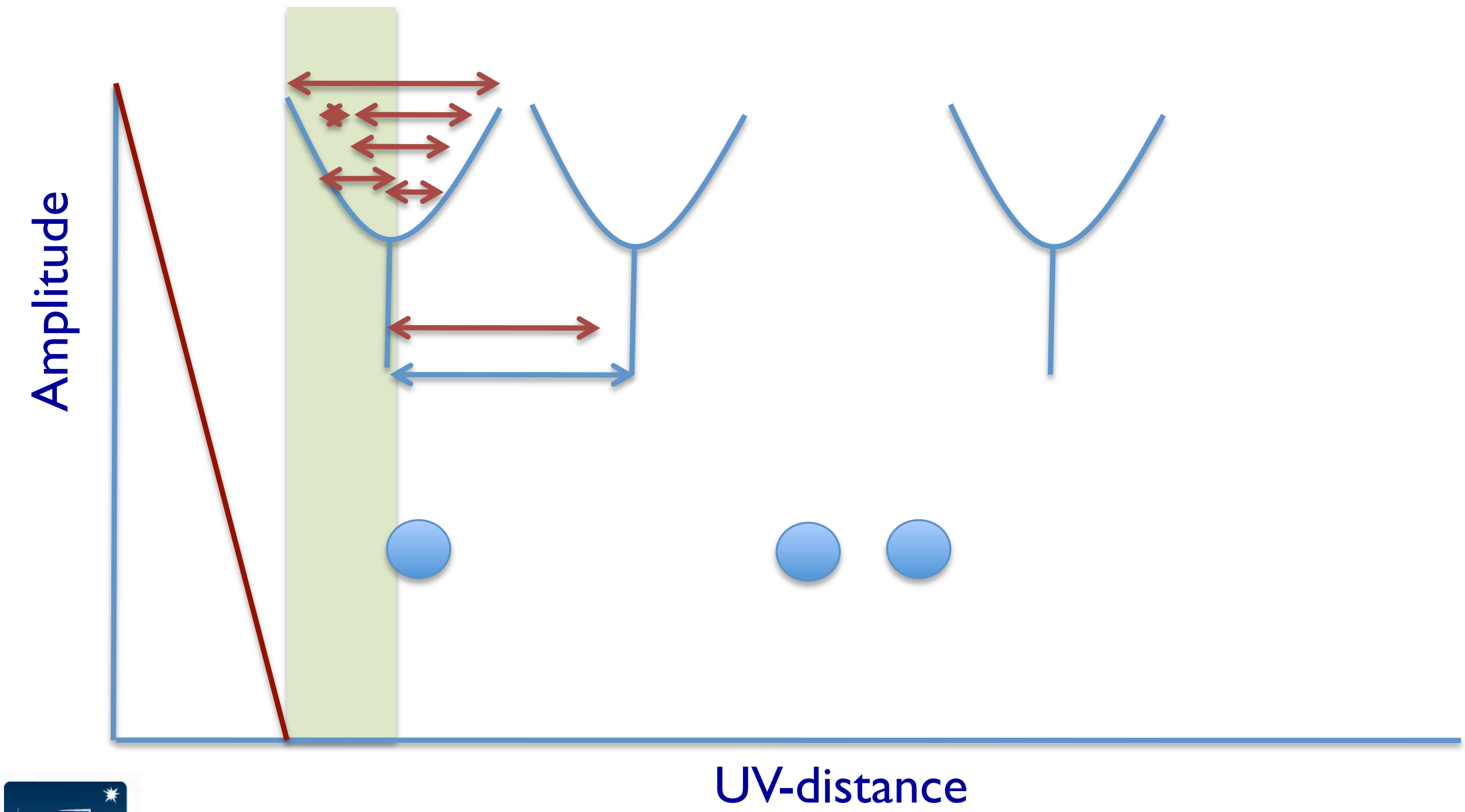
uv-coverage

- CLEAN **EXTRAPOLATES** to the central short and zero spacings
- But we would like to measure those; can we do that?
- Yes!   Use the Fourier magic of a Single Dish



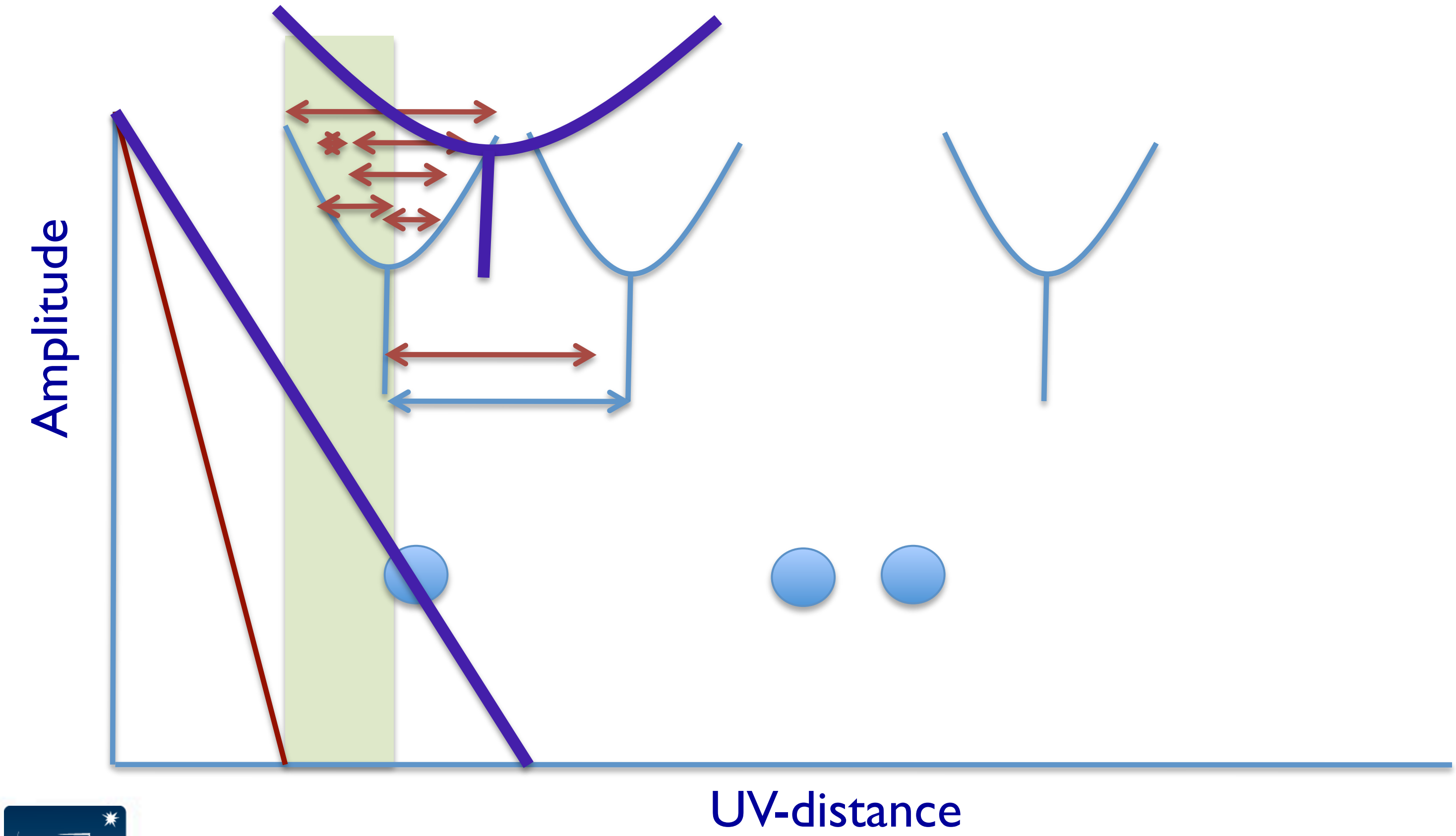
Zero spacing correction

- Get an interferometric observation
- Go to a single dish and map the same region, use a SD with a diameter larger than the shortest baseline of your interferometric map



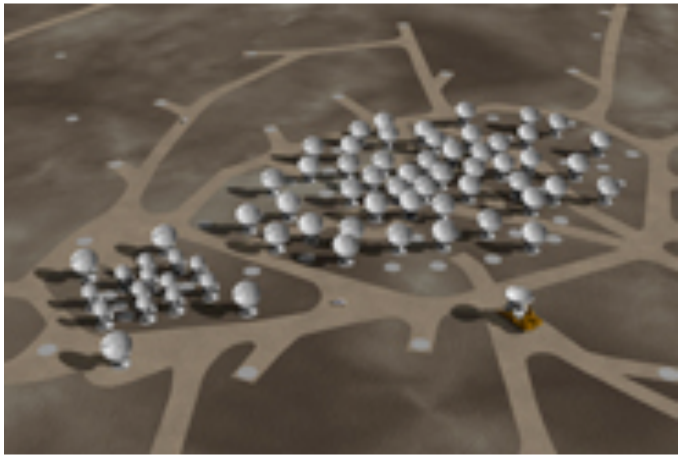
Zero spacing correction

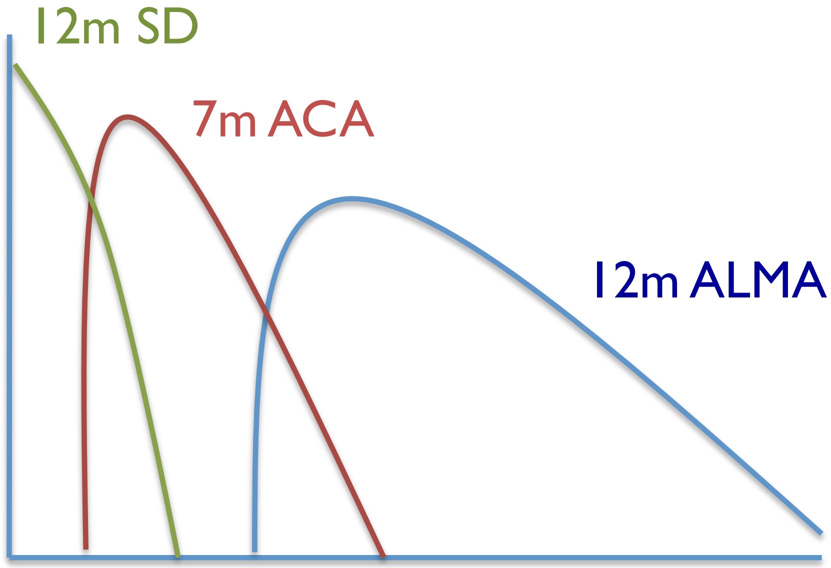
- Get an interferometric observation
- Go to a single dish and map the same region, use a SD with a diameter larger than the shortest baseline of your interferometric map



Zero/Short Spacings @ ALMA

- ALMA approach:
- 12m antennas + 7m antennas +
- 12m antennas that operate as SD





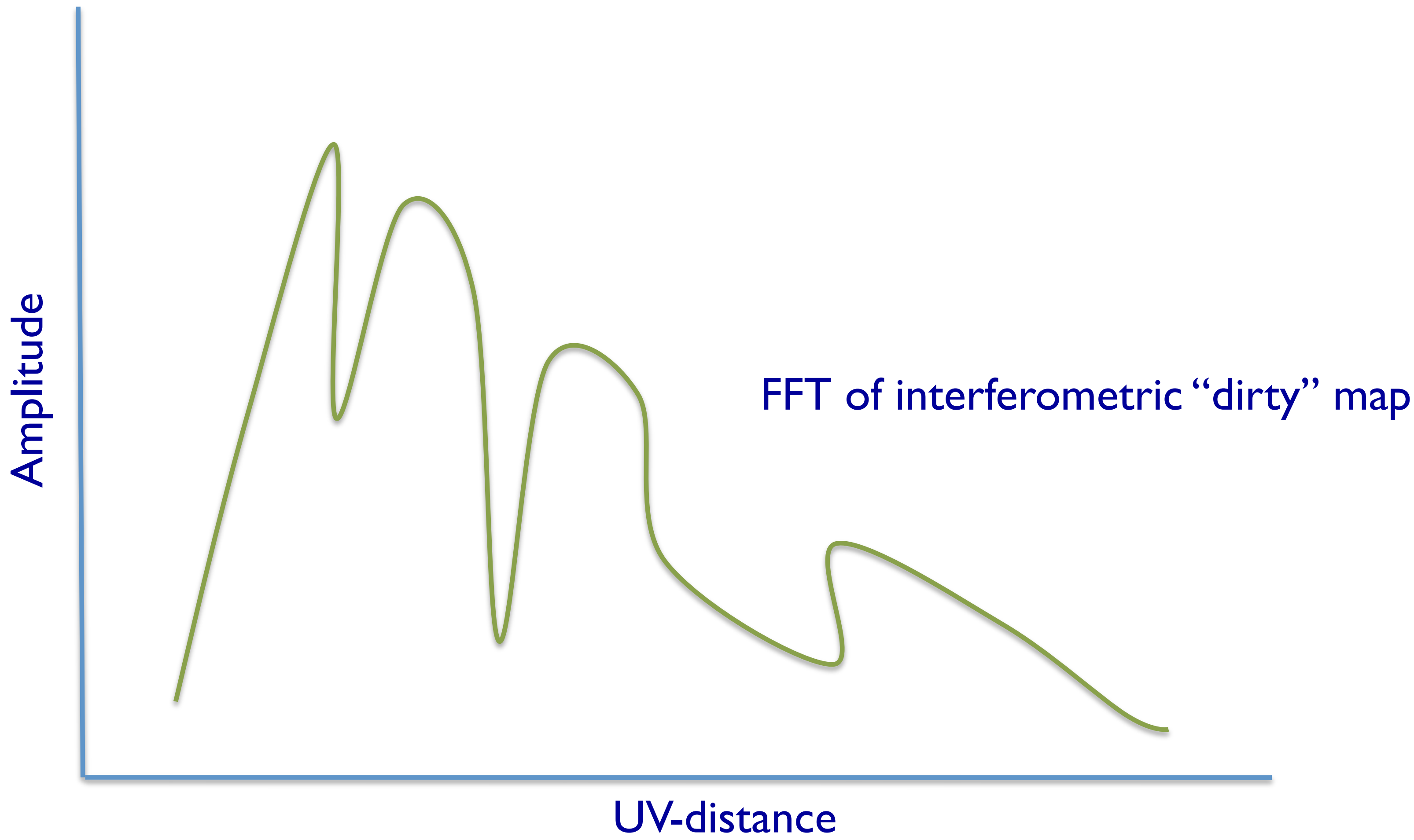
Zero spacing correction

- Get an interferometric observation
- Go to a single dish and map the same region, use a SD with a diameter larger than the shortest baseline of your interferometric map

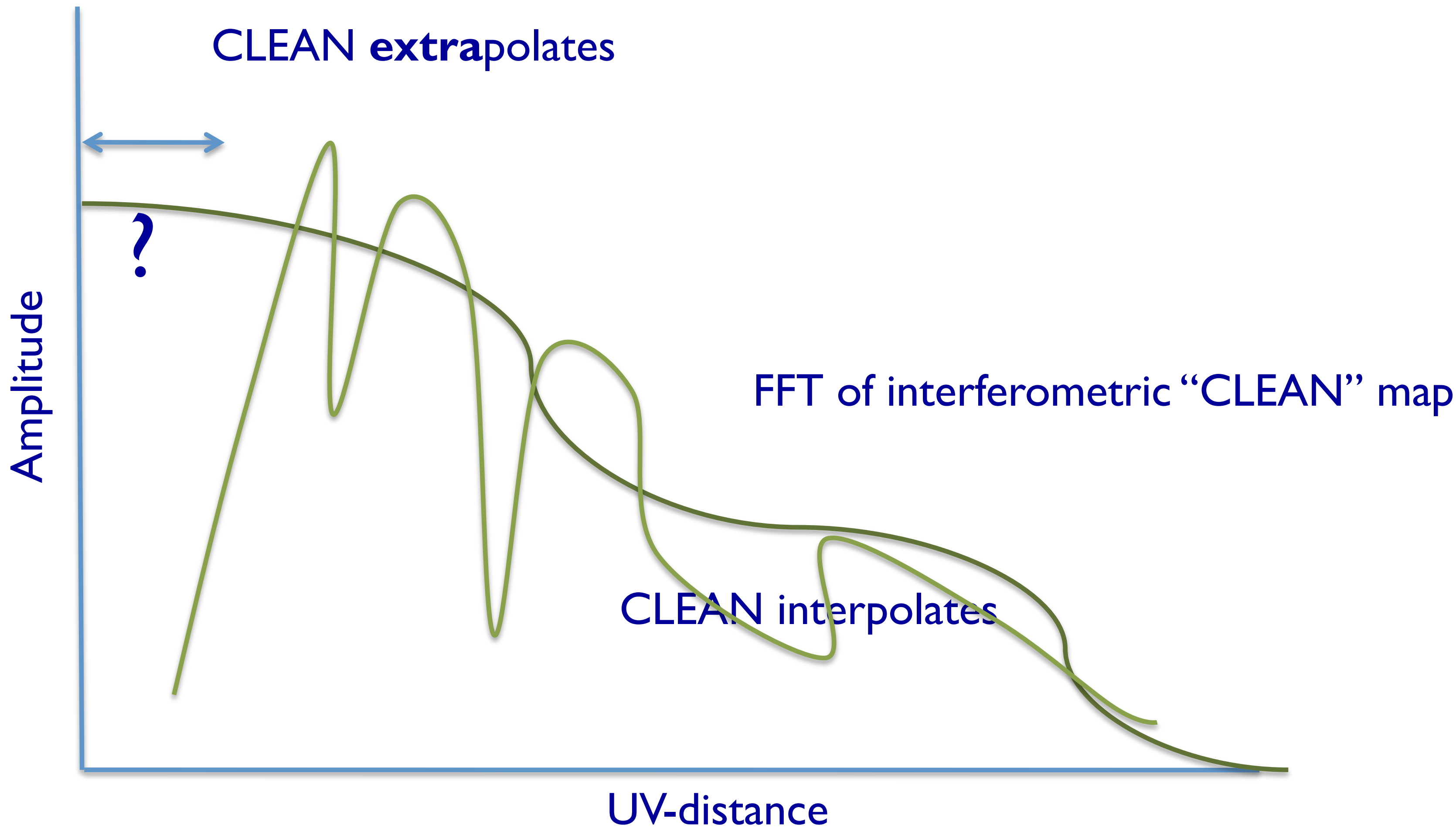
- Aim for same surface brightness sensitivity at shortest BL and SD
- Calibrate, calibrate, calibrate!

- Feathering: FT SD map → FT cleaned, interferometric map → combine both with weighting in uv-space → FT back to combined image
- Use the SD map as a model for deconvolution with (multi-scale-)clean
- Minimize Maximum Entropy χ^2 for both the SD and the interferometric map simultaneously
- Linear Combination in image domain

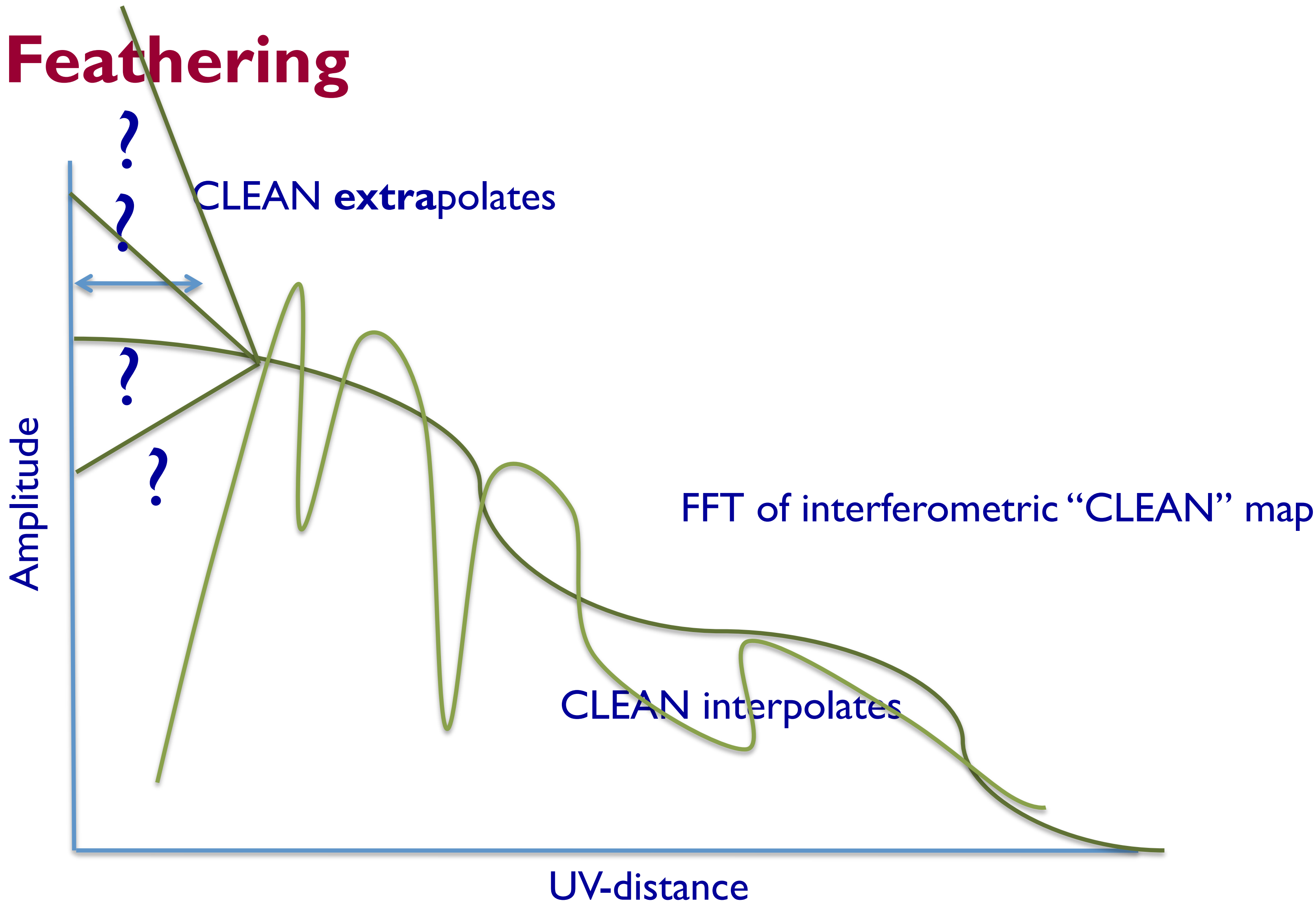
Feathering



Feathering

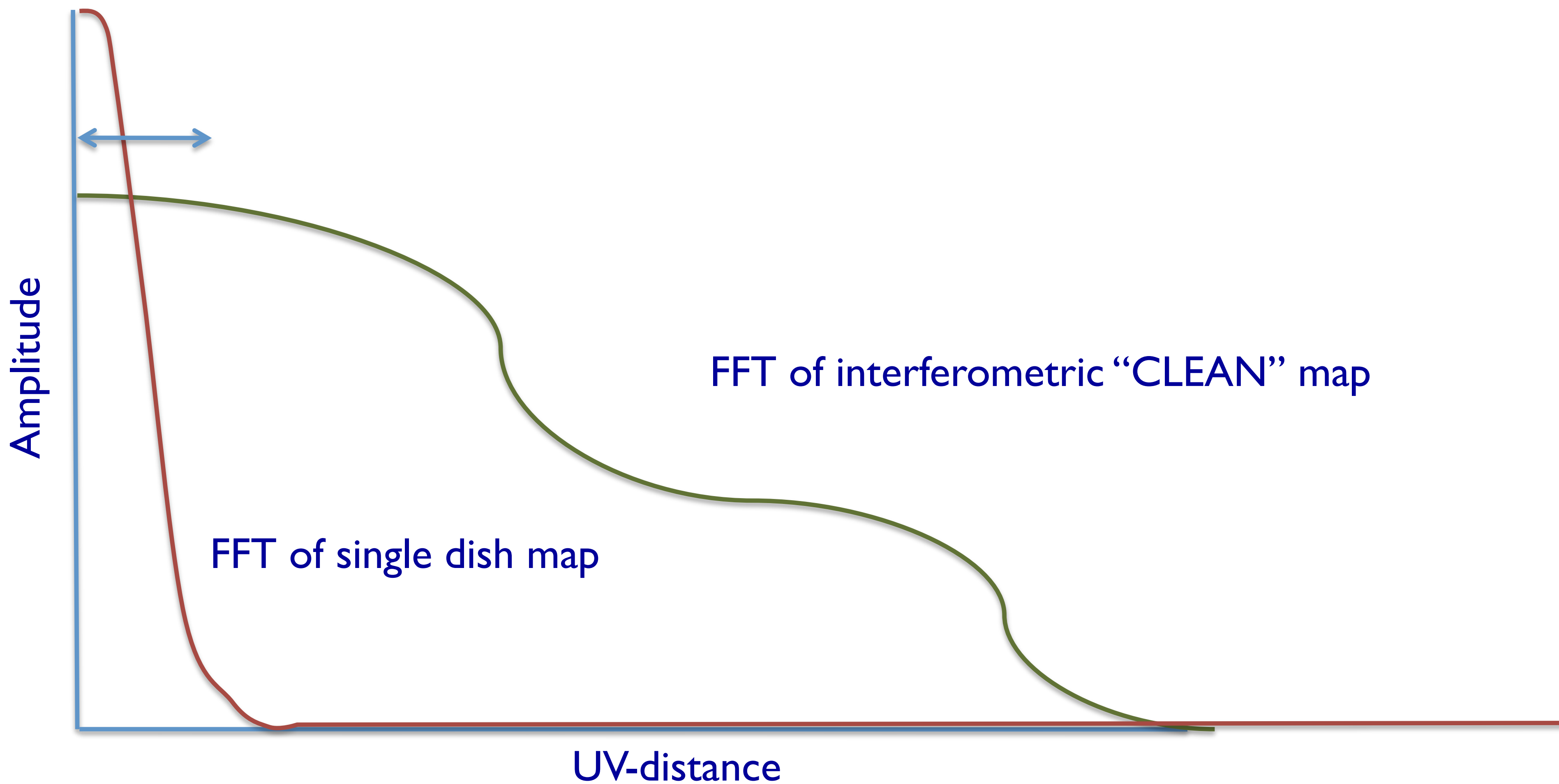

CLEAN **extra**polates
?
Amplitude
FFT of interferometric "CLEAN" map
CLEAN interpolates
UV-distance

Feathering

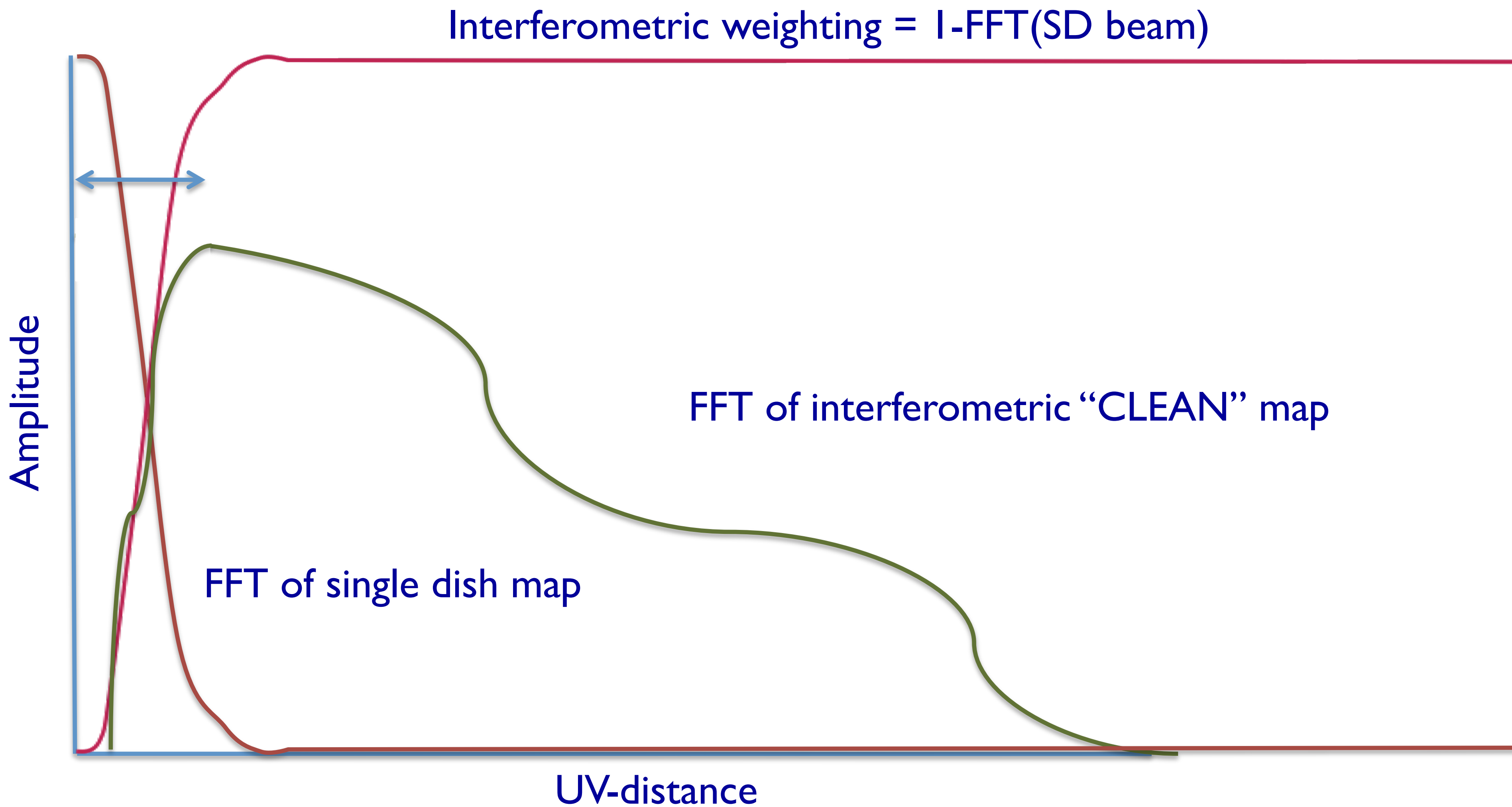


?
?
CLEAN **extra**polates
?
?
FFT of interferometric "CLEAN" map
CLEAN interpolates
Amplitude
UV-distance

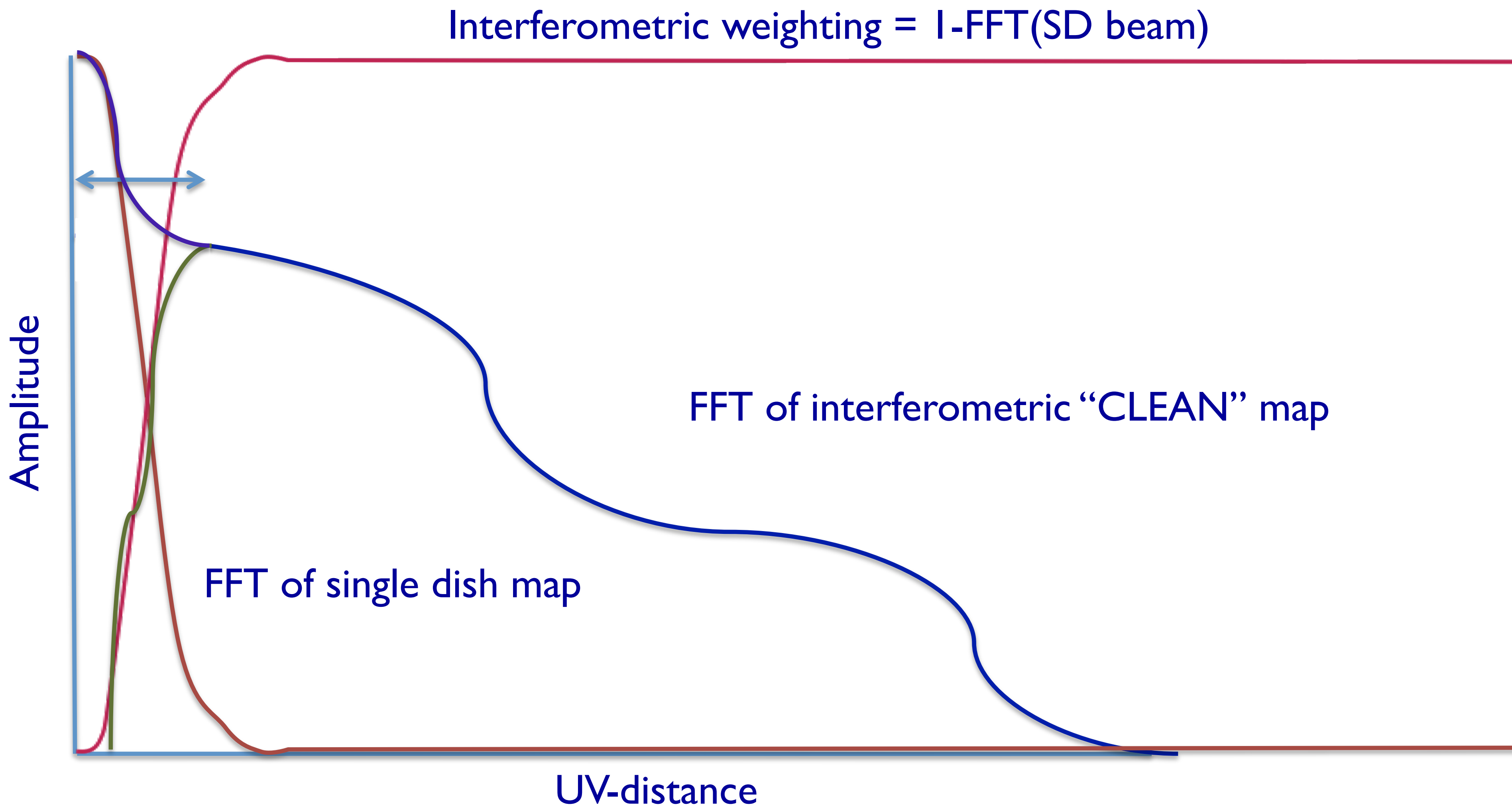
Feathering


Amplitude
FFT of interferometric "CLEAN" map
FFT of single dish map
UV-distance

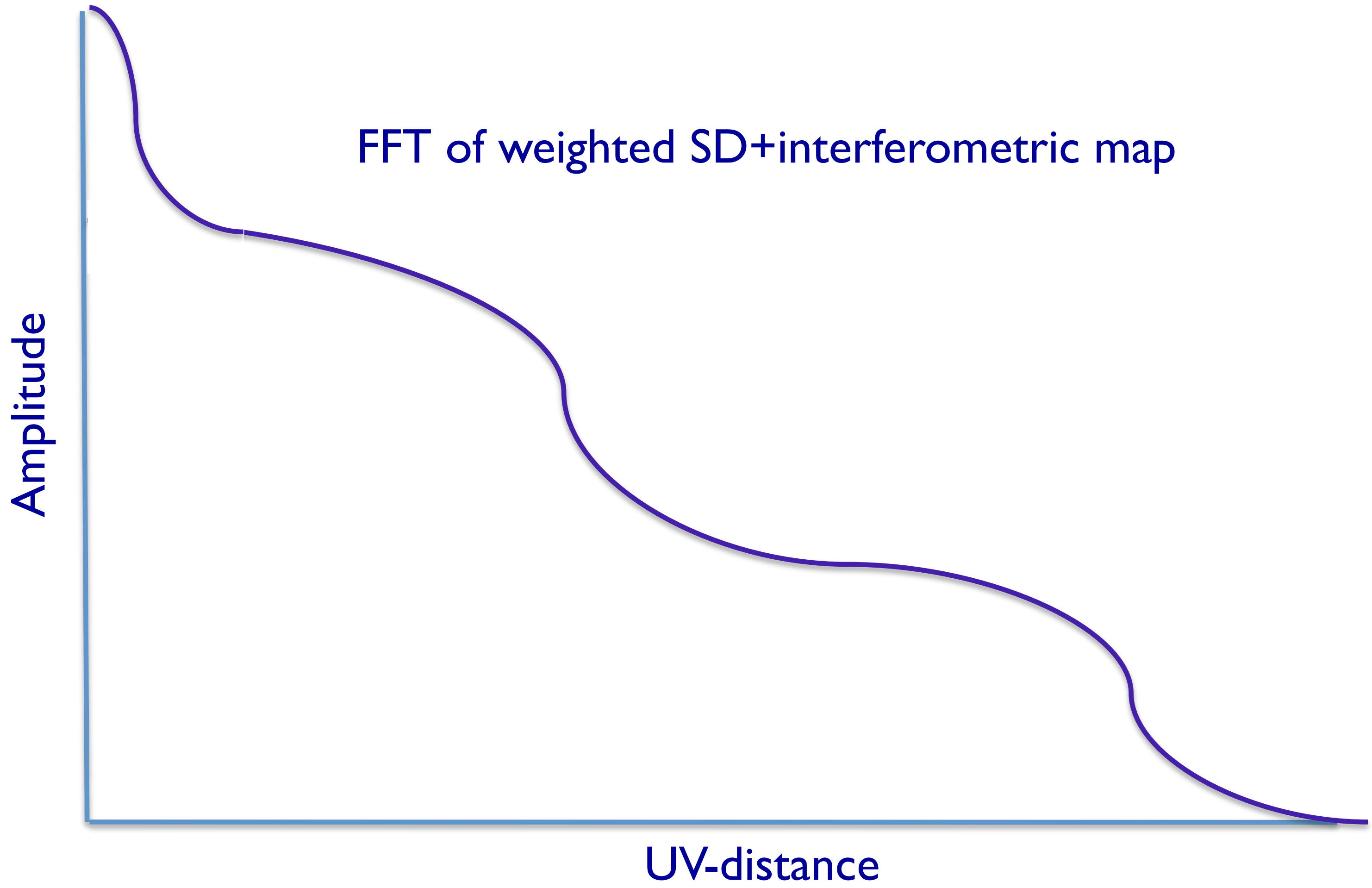
Feathering

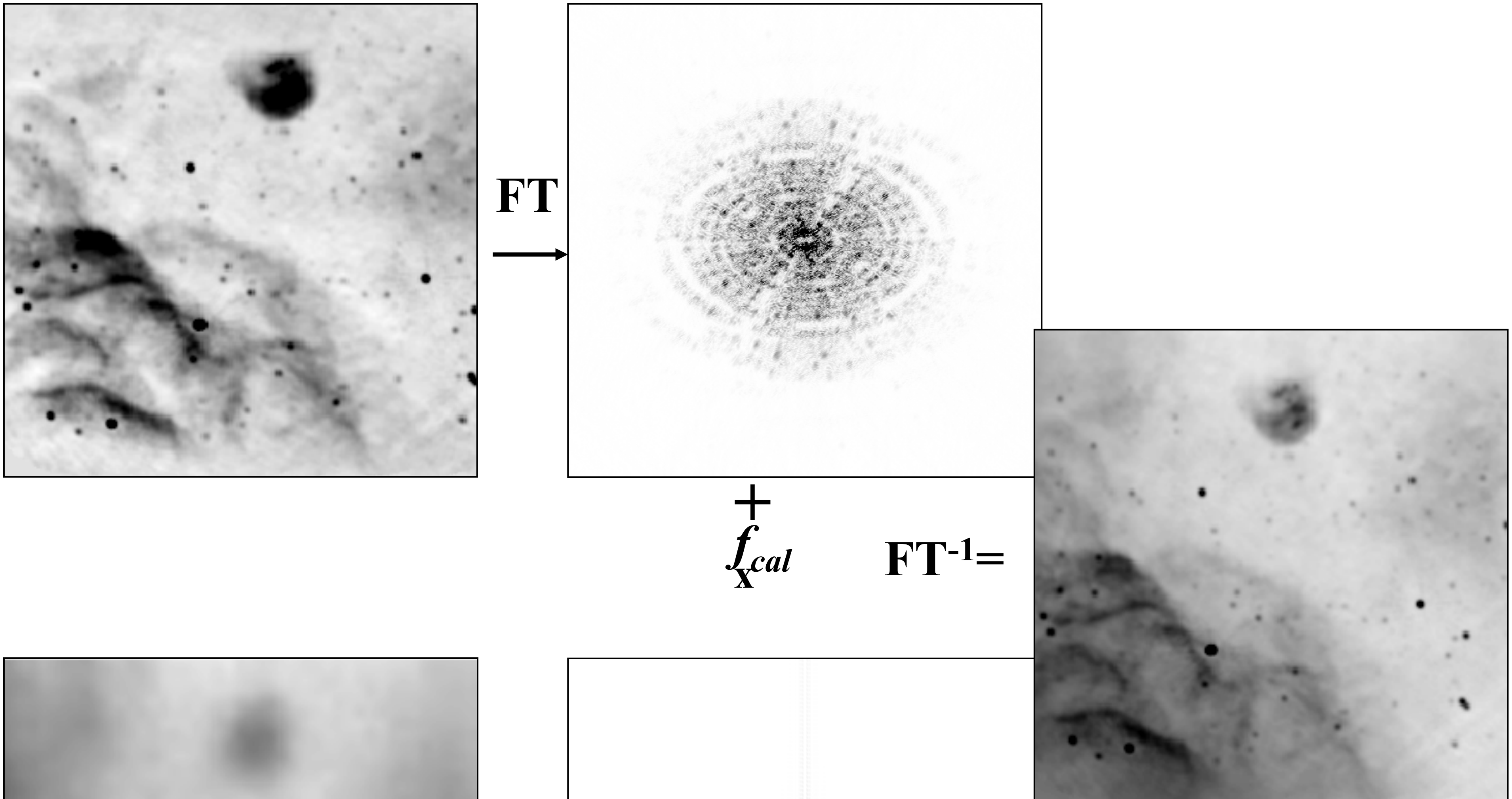

Interferometric weighting = 1-FFT(SD beam)
Amplitude
FFT of interferometric "CLEAN" map
FFT of single dish map
UV-distance

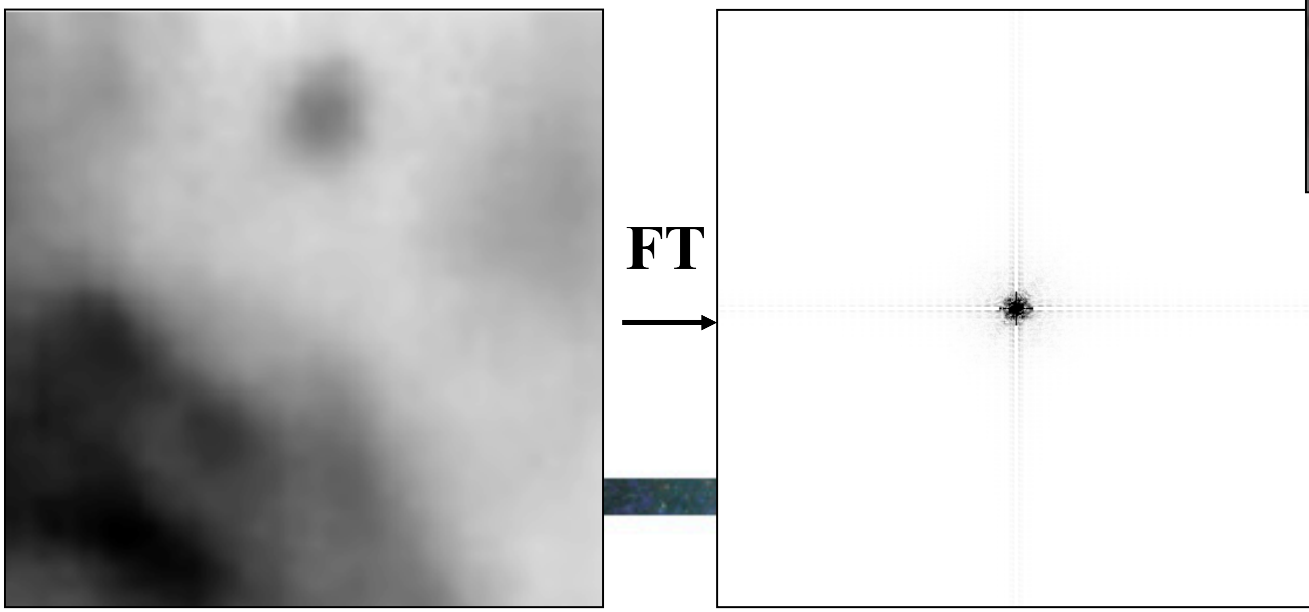
Feathering


Interferometric weighting = 1-FFT(SD beam)
Amplitude
FFT of interferometric "CLEAN" map
FFT of single dish map
UV-distance

Feathering


FFT of weighted SD+interferometric map
Amplitude
UV-distance





Example of Fourier plane combination: McClure-Griffiths et al.

Methods – Joint Deconvolution with Maximum Entropy

- Deconvolve both images simultaneously, with 2 conditions for improving the quality (entropy)

$$\aleph = -\sum_i I_i \ln\left(\frac{I_i}{M_i e}\right)$$

Subject to (1)

$$\sum_i \left\{ I_{\text{int}}^D - B_{\text{int}} * I \right\}_i^2 < N\sigma_{\text{int}}^2$$

(2)

$$\sum_i \left\{ I_{sd}^D - \frac{B_{sd} * I}{f_{sd}} \right\}_i^2 < M\sigma_{sd}^2$$

Methods – ‘Linear’ Combination in Image Domain

Combined image:

$$I_{tot} = w_{int} I_{int} + w_{sd} f_{sd} I_{sd}$$

Weights:

$$w_{int} = \frac{\Omega_{sd}}{\Omega_{int} + \Omega_{sd}}$$

$$w_{sd} = \frac{\Omega_{int}}{\Omega_{int} + \Omega_{sd}}$$

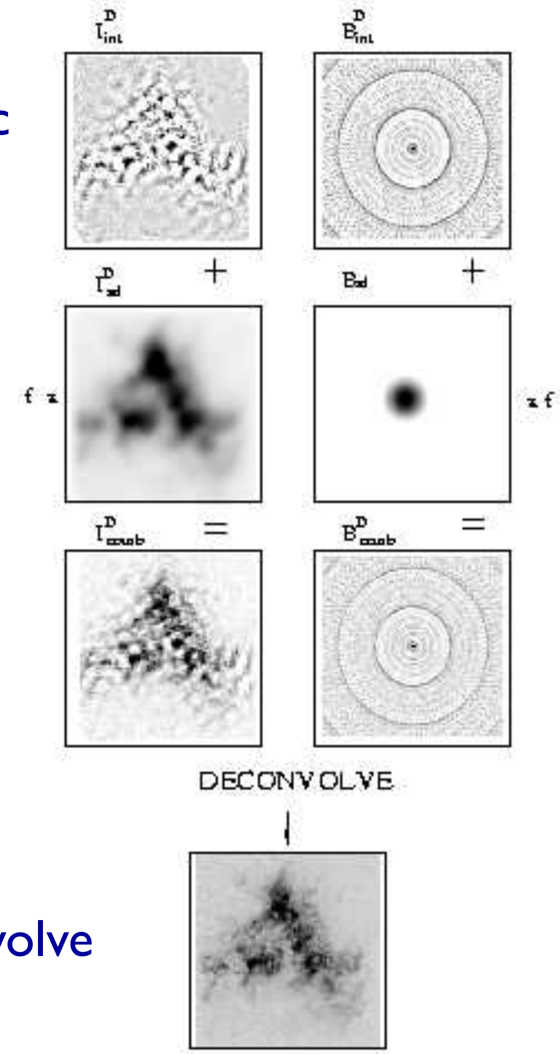
Stanimirovic et al.

Interferomteric Dirty image and beam

SD image and beam

Combinination

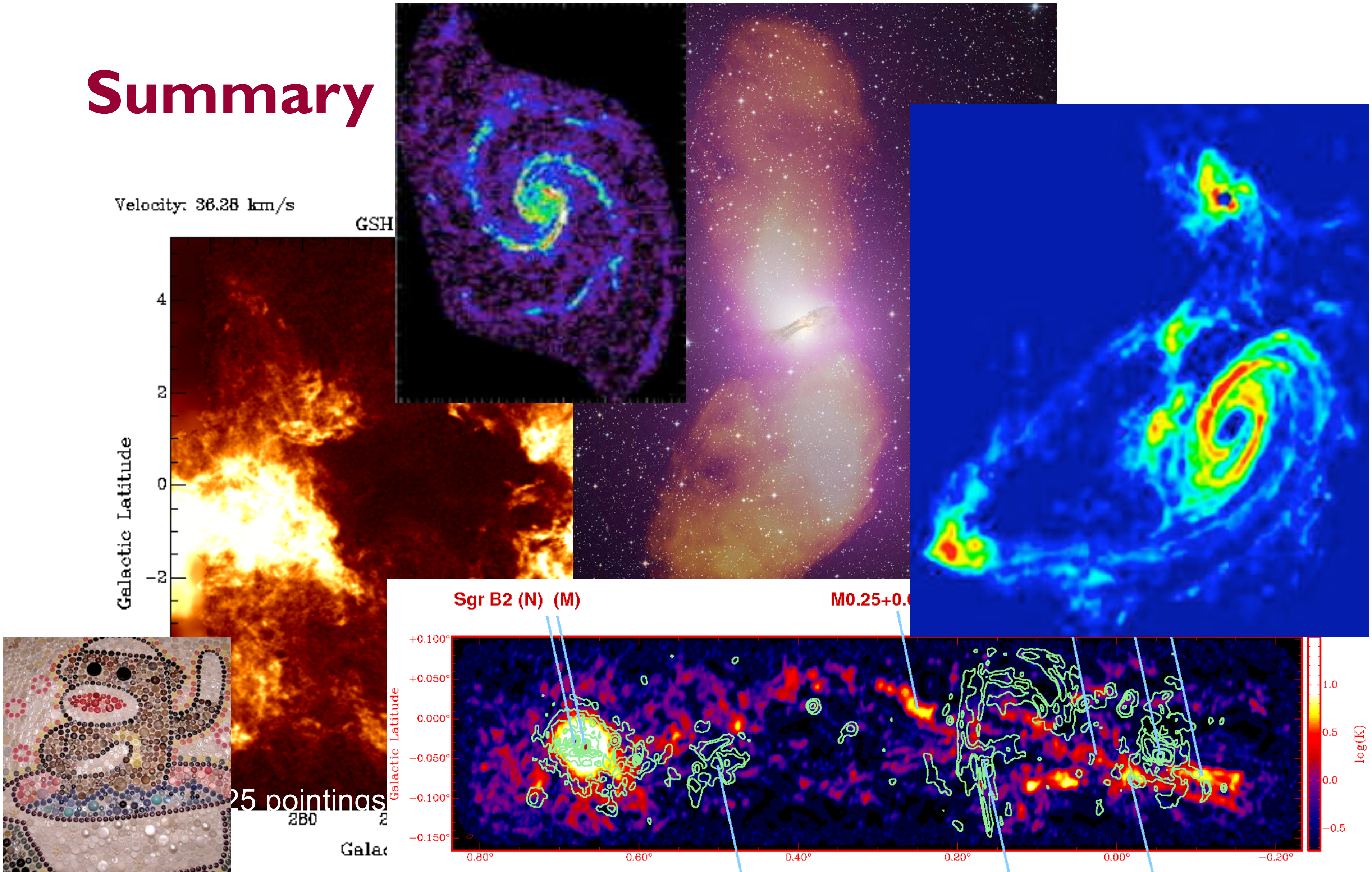
deconvolve



Summary

- Mosaicing is a technique to image objects much larger than the primary beam
- Unlocks additional uv spacings by using dish extends for additional uv-coverage (convolution of visibilities with FT[PB])
- Needs a bit of care and thought to setup
- Mosaicing techniques will be used very commonly in the future:
  - ALMA features a small PB from 1' @3mm to 10" @600GHz
  - SKA/pathfinders sport large beam but aim for large (all-sky) sky coverage
- Zero/short spacing reconstruction may be required
  - Use a large SD telescope and carefully apply calibration
- Choose the best mosaicing and zero/short spacing correction method for your problem
- Fun to reduce and beautiful images!

Summary



Velocity: 36.28 km/s
GSH
Galactic Latitude
25 pointings
Sgr B2 (N) (M)
M0.25+0.
Galactic Latitude
Galactic Longitude
log(K)
Sgr B1
Pistol
M−0.02−0.07