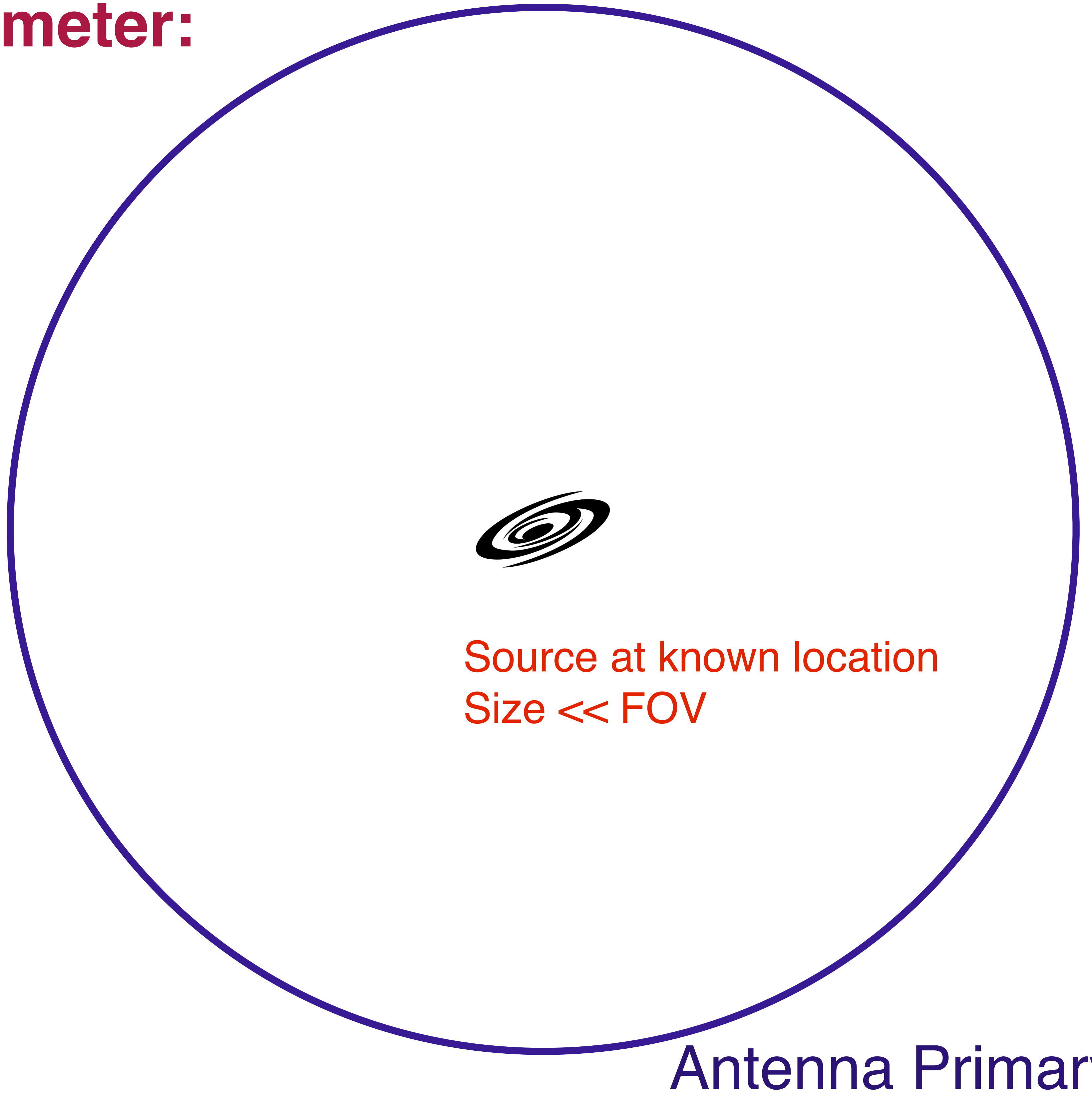
Mosaicking

Brian Mason (NRAO)

Sixteenth Synthesis Imaging Workshop

16-23 May 2018

The simplest observing scenario for an interferometer:

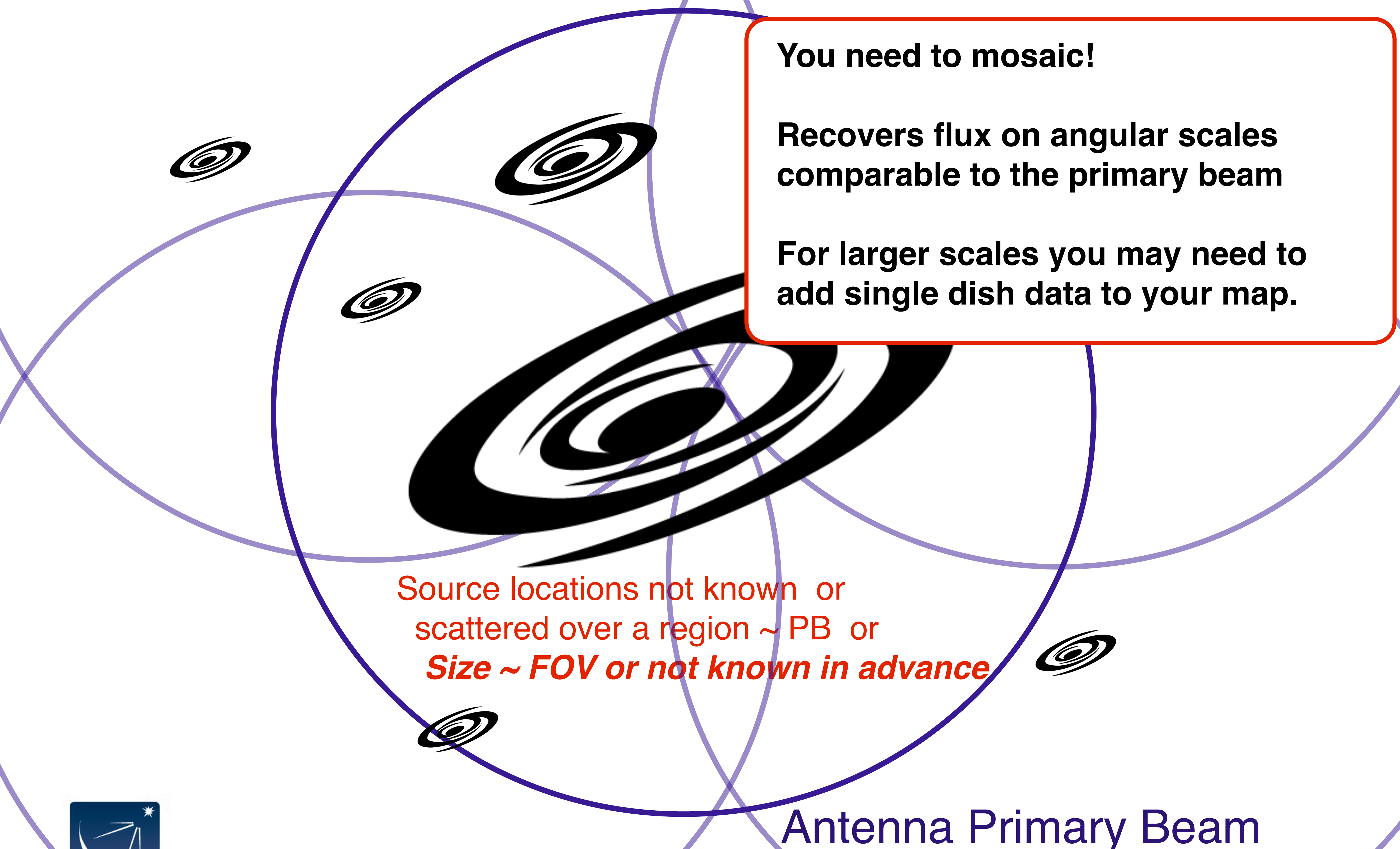


But that’s often not the case...

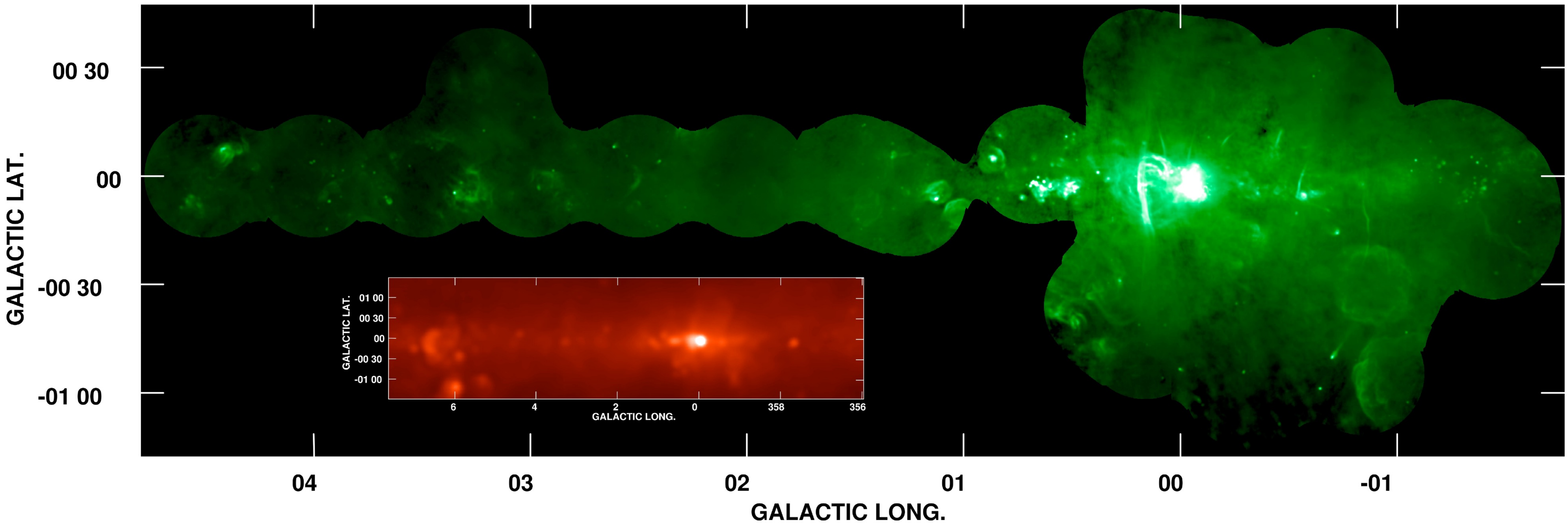
You need to mosaic!

Recovers flux on angular scales comparable to the primary beam

For larger scales you may need to add single dish data to your map.



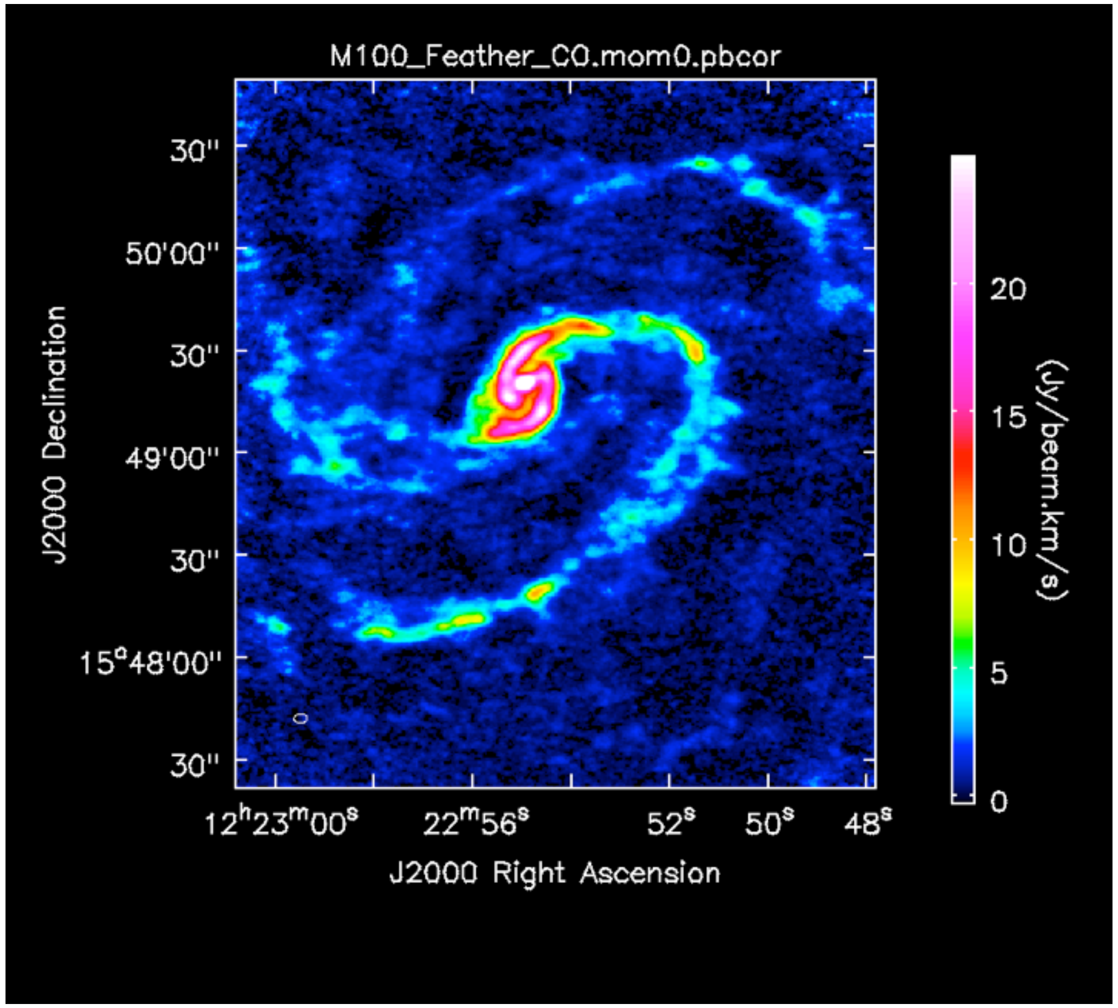
NRAO

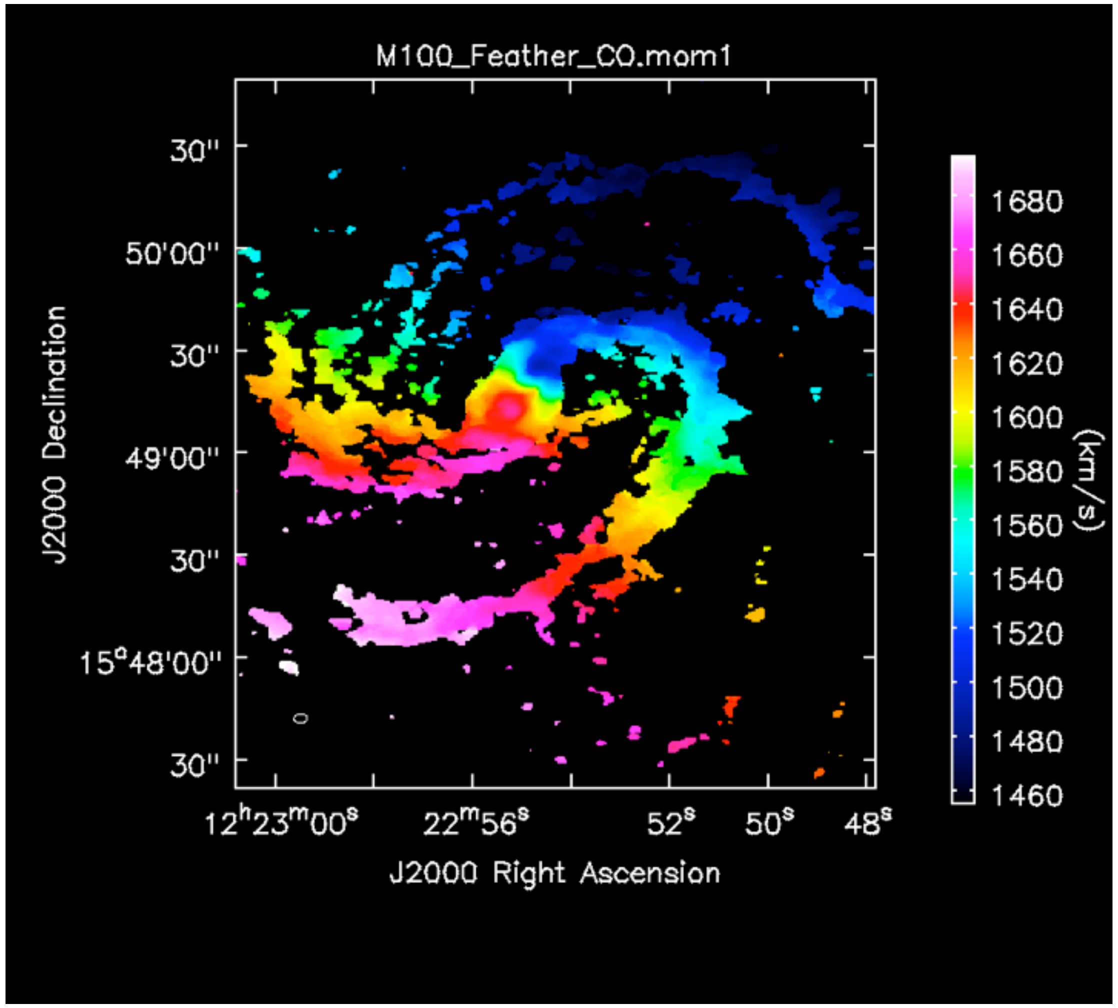


20cm VLA Mosaic+GBT Single Dish (green) (red inset :GBT only)

Law, Yusef-Zadeh, & Cotton (2008)

ALMA Science Verification: M100

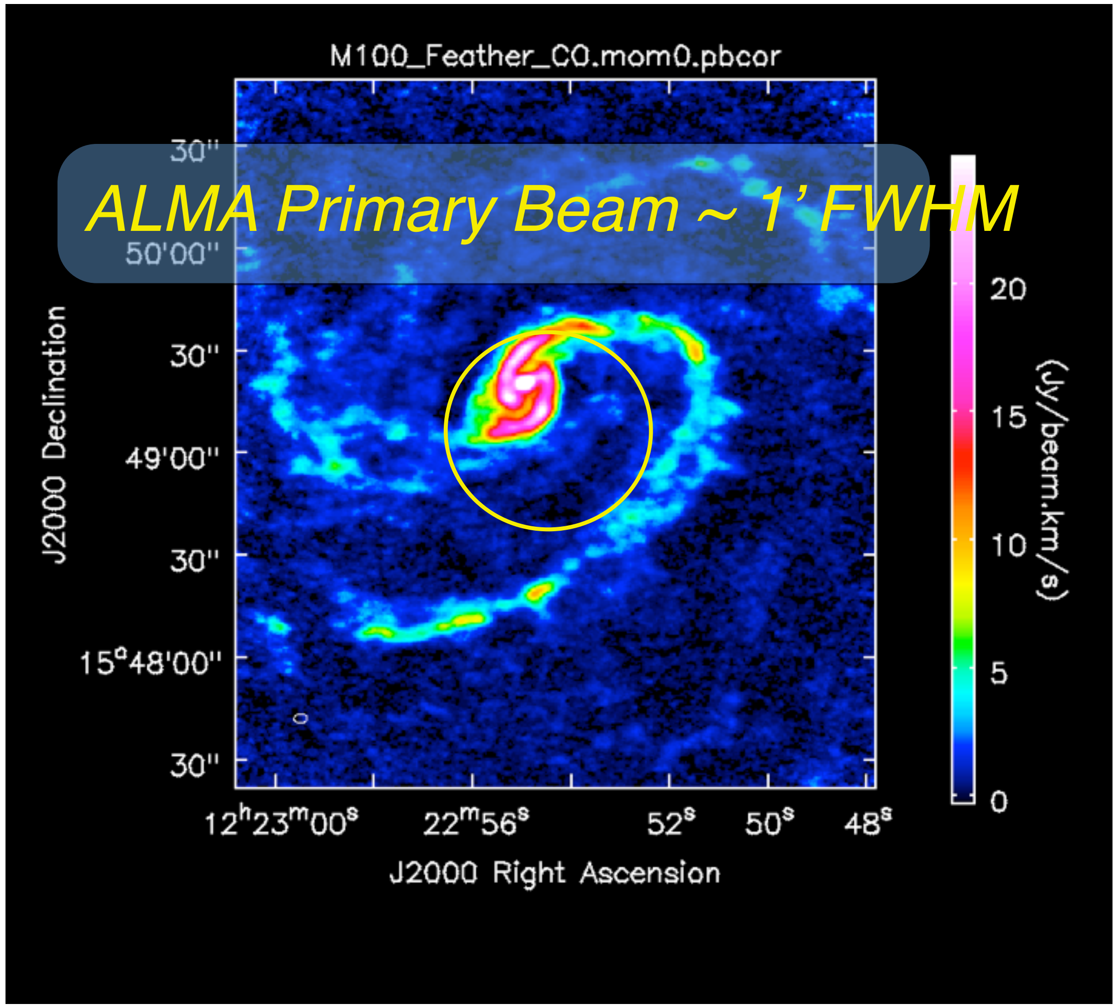


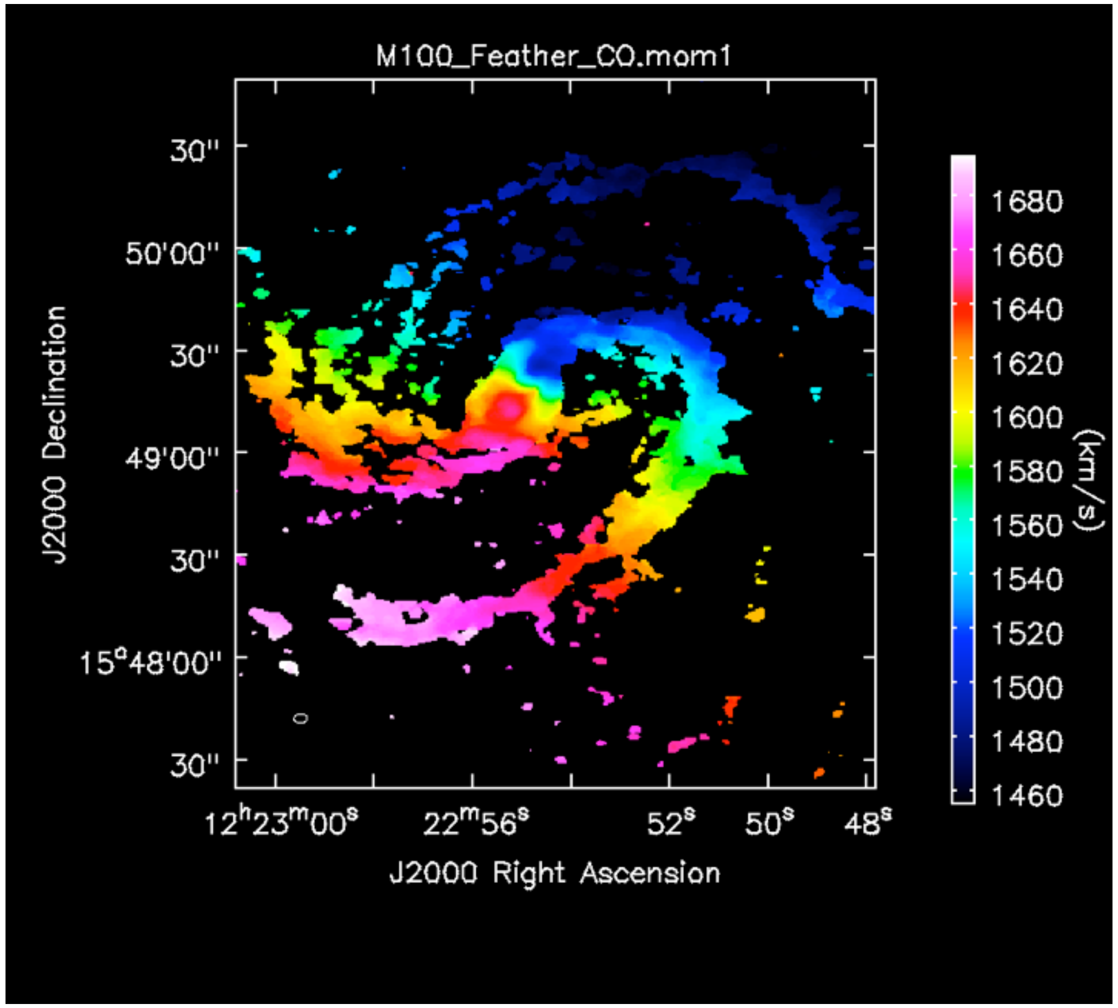


Integrated CO line intensity line)
Band 3 (115 GHz, ~2.6mm)

1st moment map (velocity field of CO)

ALMA Science Verification: M100

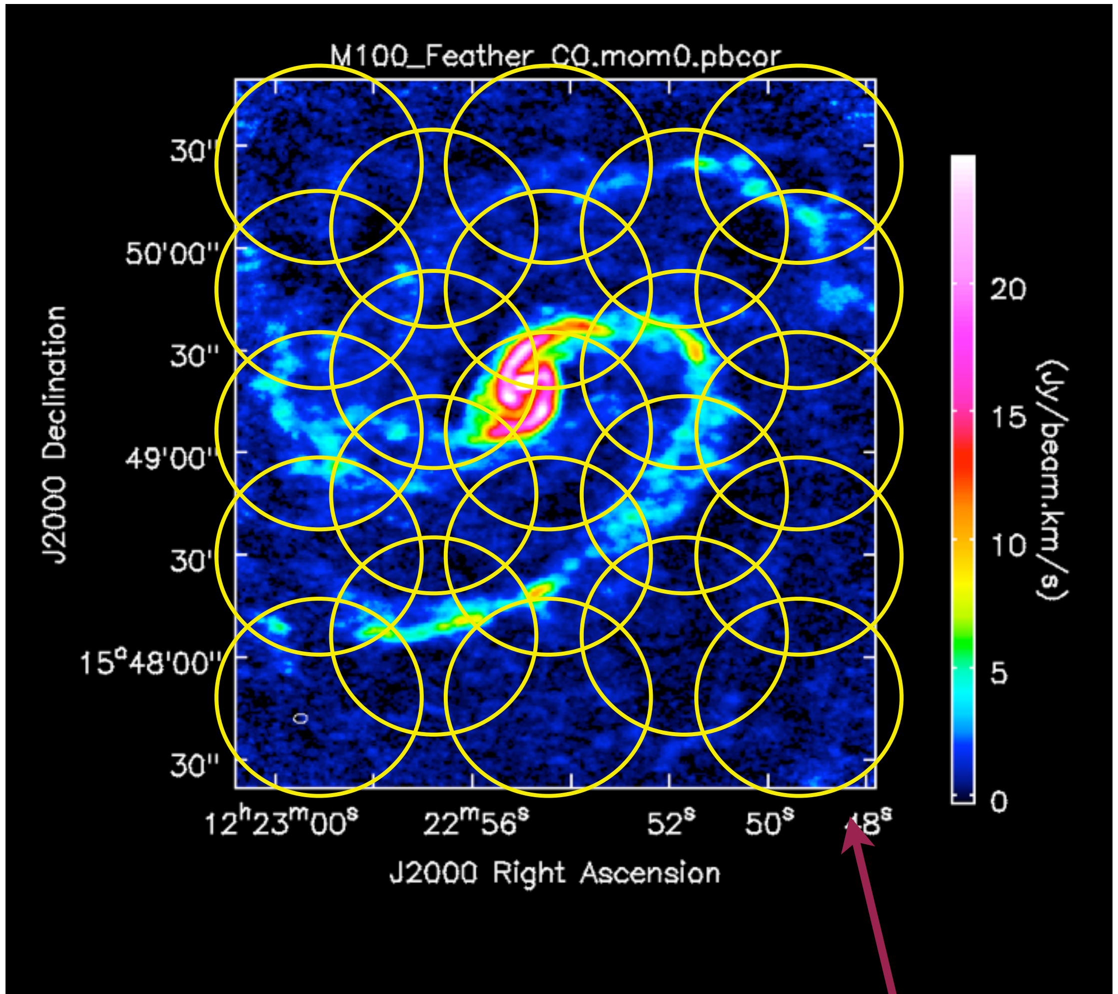


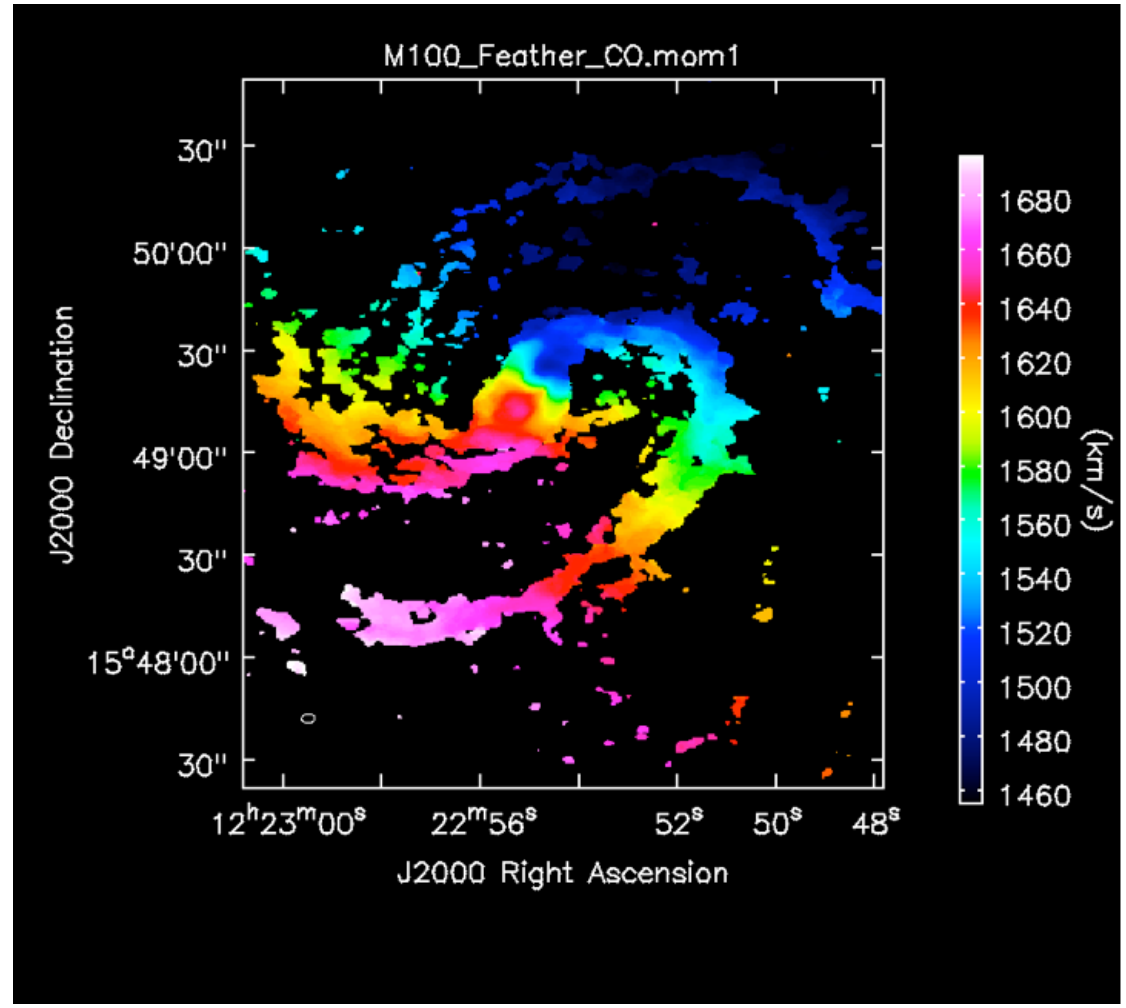


Integrated CO line intensity line)
Band 3 (115 GHz, ~2.6mm)

1st moment map (velocity field of CO)

ALMA Science Verification: M100





Integrated CO line intensity line)
Band 3 (115 GHz, ~2.6mm)

1st moment map (velocity field of CO)

At short wavelengths, mosaicking is very commonly required

Limiting Angular Scales for an Interferometer

$$\theta_{PB} = (1.03 \rightarrow 1.2) \times \frac{\lambda}{D}$$

~ the diameter of the area imaged by one pointing of the interferometer (instantaneous field of view)

$$\theta_{LAS} = \frac{1}{2}\left(\frac{\lambda}{b_{min}}\right)$$

The "Spatial Period" of the largest angular scale Fourier component of the sky brightness measured by the interferometer

In practice, you only measure things *half* that big (say) very well. (even that might be optimistic)

Exercise: you can quantify the LAS yourself using the "Gaussian Flux Loss" rule of thumb (D.Wilner lecture on deconvolution)

Limiting Angular Scales for an Interferometer

$$\theta_{PB} = (1.03 \rightarrow 1.2) \times \frac{\lambda}{D}$$

~ the diameter of the area imaged by one pointing of the interferometer (instantaneous field of view)

$$\theta_{LAS} = \frac{1}{2} \frac{\lambda}{b_{min}}$$

The "Spatial Period" of the largest angular scale Fourier component of the sky brightness measured by the interferometer

CAVEAT: a single short baseline doesn't do a lot of good — b_{min} should be taken to be the shortest spacing at which there is good uv-coverage

In practice, you only measure things *half* that big (say) very well. (even that might be optimistic)

Exercise: you can quantify the LAS yourself using the "Gaussian Flux Loss" rule of thumb (D.Wilner lecture on deconvolution)

NRAO

Limiting Angular Scales for an Interferometer

$$\theta_{PB} = (1.03 \rightarrow 1.2) \times \frac{\lambda}{D}$$

VLA: L-band (20cm) = 30' Q-band (7mm) = 1'

ALMA(12m): Band3 (3mm) = 1' Band9 (0.44mm) = 9"

$$\theta_{LAS} = \frac{1}{2} \frac{\lambda}{b_{min}}$$

VLA: L-band (20cm), D-array = 16' *30m* Q-band (7mm), A-array = 1.2" *537m*

ALMA(12m): Band3 (3mm), C43-1= 28" *15m* Band6 (1.3mm), C43-10= 0.2" *43m*

(based on currently advertised capabilities)

NRAO

Limiting Angular Scales for an Interferometer

If your region of interest is larger than this, you need to mosaic together many interferometer pointings.

$$\theta_{PB} = (1.03 \rightarrow 1.2) \times \frac{\lambda}{D}$$

VLA: L-band (20cm) = 30' Q-band (7mm) = 1'

ALMA(12m): Band3 (3mm) = 1' Band9 (0.44mm) = 9"

$$\theta_{LAS} = \frac{1}{2} \frac{\lambda}{b_{min}}$$

If the structures you are interested in are larger than this, you need to mosaic and/or get data from a more compact configuration of the interferometer or single dish.

VLA: L-band (20cm), D-array = 16' Q-band (7mm), A-array = 1.2"

30m *537m*

ALMA(12m): Band3 (3mm), C43-1= 28" Band6 (1.3mm), C43-10= 0.2"

15m *43m*

(based on currently advertised capabilities)

Limiting Angular Scales for an Interferometer

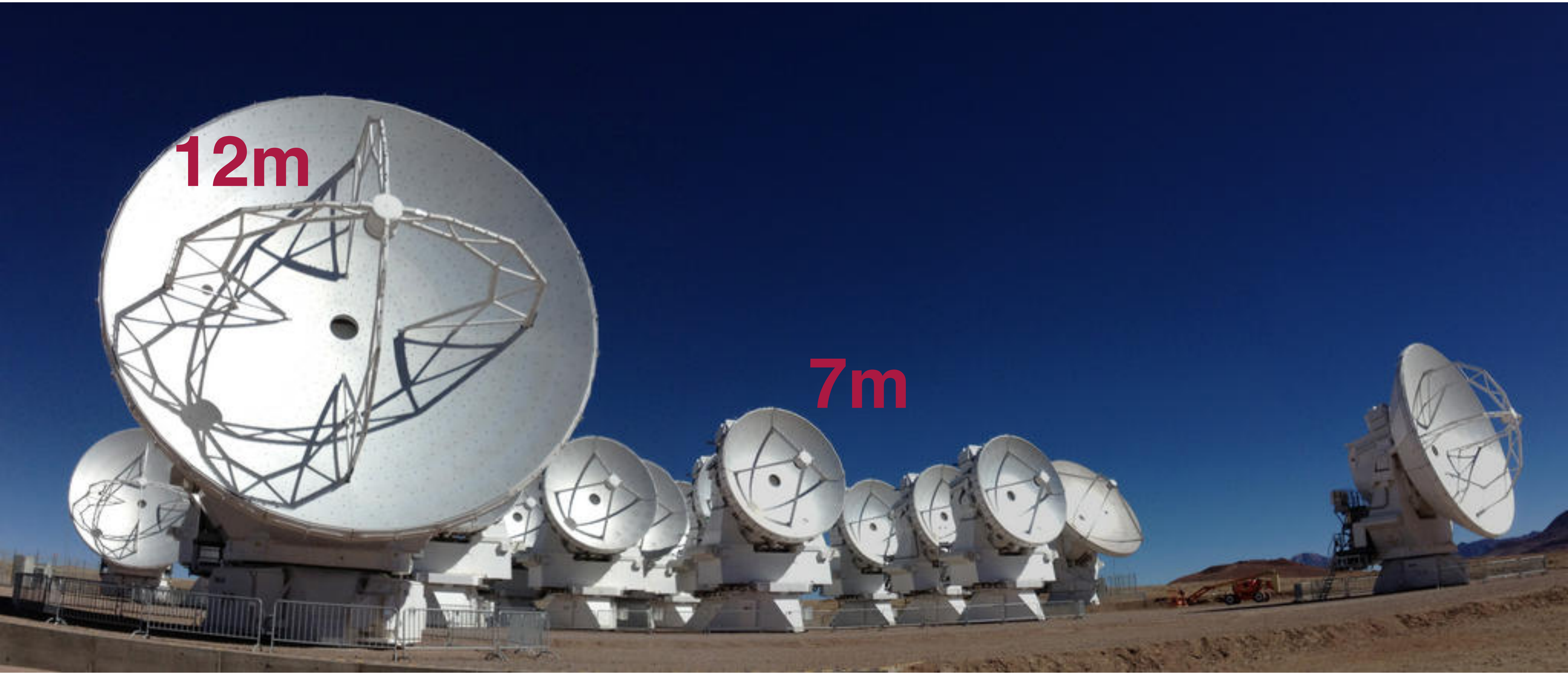
$$\theta_{PB} = (1.03 \rightarrow 1.2) \times \frac{\lambda}{D}$$

There is a limit to how compact a given interferometer can get

$$\theta_{LAS} = \frac{1}{2}\frac{\lambda}{b_{min}} \qquad \leq \frac{1}{2}\frac{\lambda}{D}$$

For angular scales much bigger than that you need smaller dishes (or data from a single dish telescope).

The ALMA Compact Array (ACA)



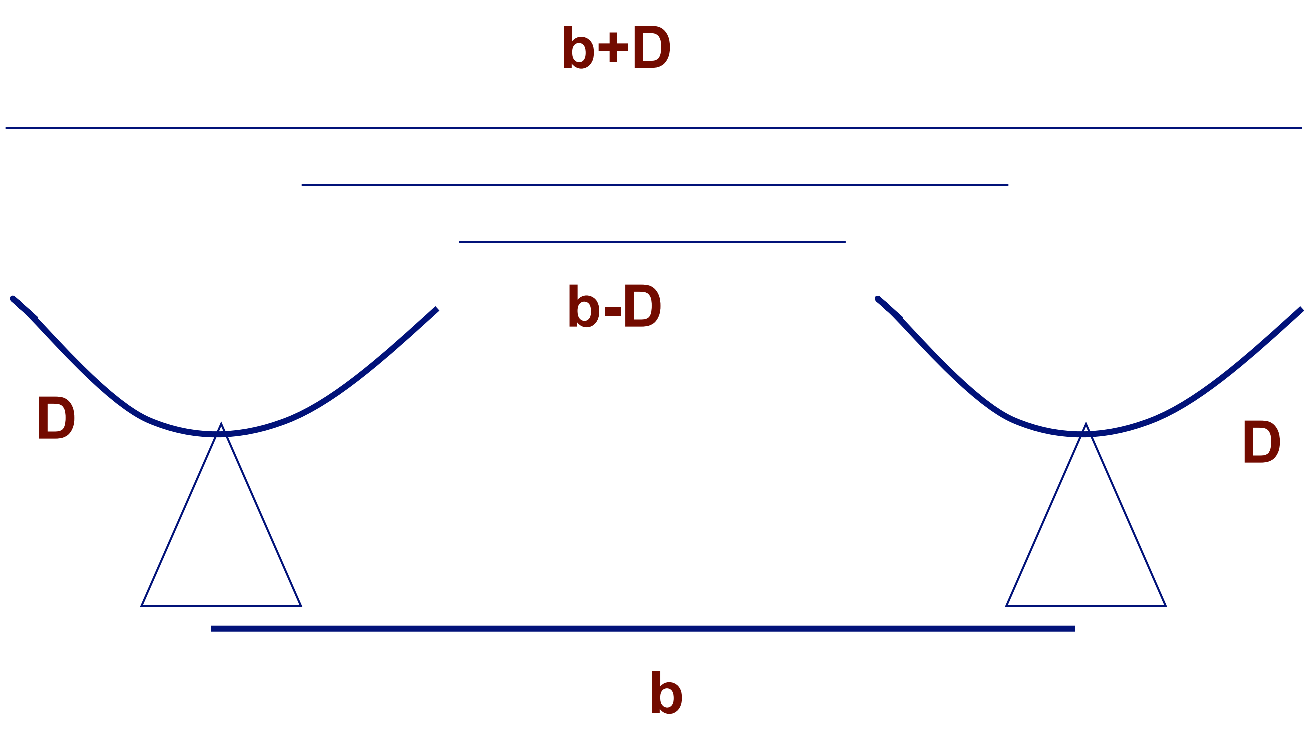
The ALMA Compact Array (ACA)

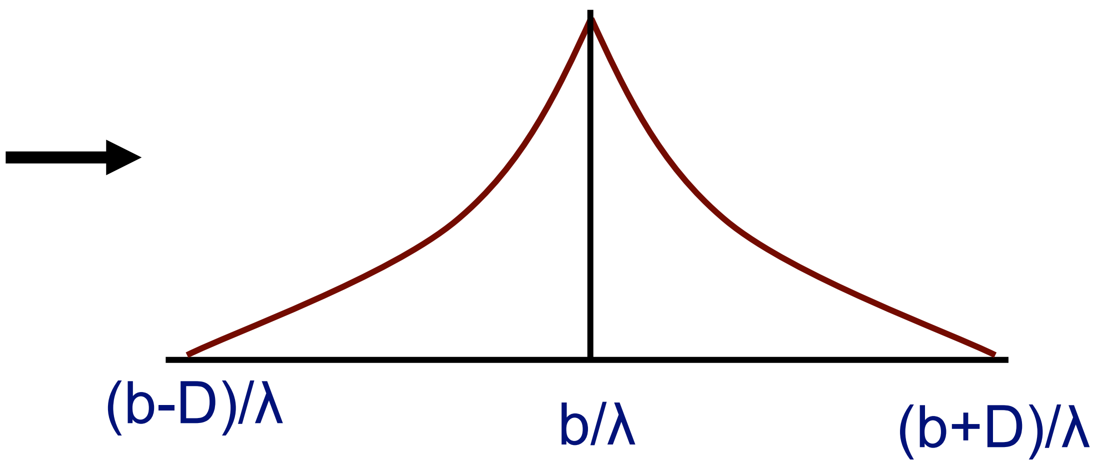
The ALMA Compact Array (ACA)

but there’s a trick…

Theory of Mosaicking: Ekers & Rots Theorem

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

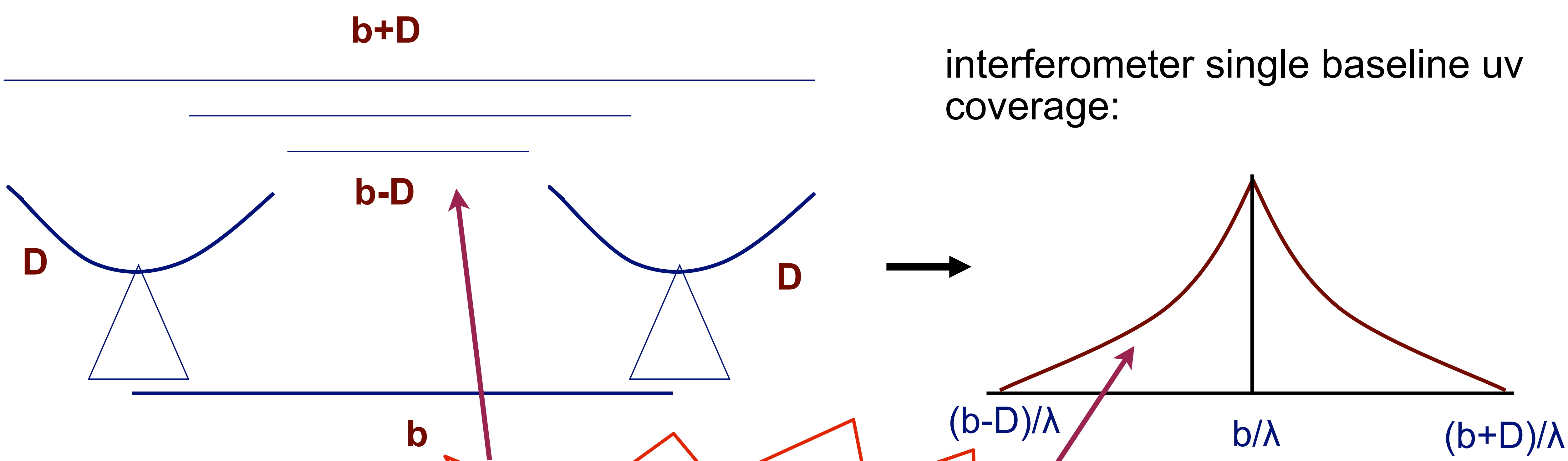




Ekers & Rots (1979)

Theory of Mosaicking: Ekers & Rots Theorem

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



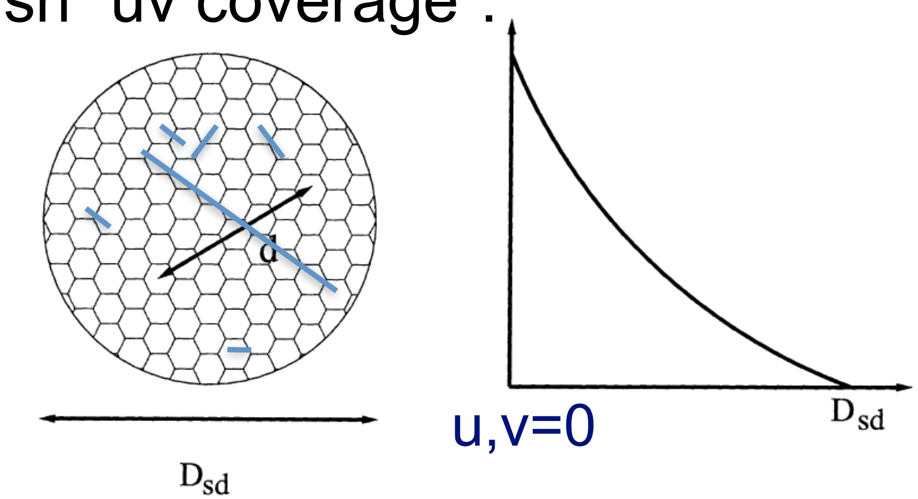
Information on scales larger than the shortest baseline

Ekers & Rots (1979)

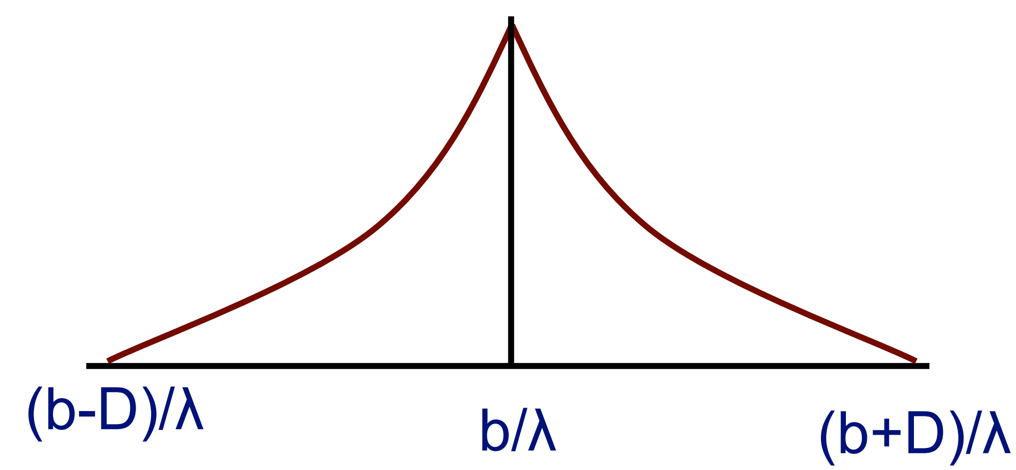
Theory of Mosaicking: Ekers & Rots Theorem

Similarly: a single dish measures a range of baselines from spatial frequencies of *zero* (the mean level of the sky) up to (the dish diameter)/λ

single dish "uv coverage":



interferometer single baseline uv coverage:



Ekers & Rots (1979)

Theory of Mosaicking: Ekers & Rots Theorem

"An interferometer measures $\lambda/(b-D) < \theta < \lambda/(b+D)$"

Motivation/Derivation:

$$\begin{aligned} V(u,v) = \int\int d\ell\, dm\, A(\ell,m) I(\ell,m)\, e^{-2\pi(u\ell+vm)} &= FT[A(\ell,m) I(\ell,m)] \\ &= FT[A(\ell,m)] \otimes FT[I(\ell,m)] \end{aligned}$$

Theory of Mosaicking: Ekers & Rots Theorem

"An interferometer measures $\lambda/(b-D) < \theta < \lambda/(b+D)$"

Motivation/Derivation:

$$
\begin{aligned}
V(u,v) = \int\int d\ell\, dm\, A(\ell,m) I(\ell,m)\, e^{-2\pi(u\ell+vm)} &= FT[A(\ell,m) I(\ell,m)] \\
&= FT[A(\ell,m)] \otimes FT[I(\ell,m)]
\end{aligned}
$$

$$
\begin{aligned}
A(\ell,m) &= \left| \int\int_{aperture} du\, dv\, E(u,v)\, e^{-2\pi(u\ell+vm)} \right|^2 \\
&= FT[E(u,v)]\, FT[E(u,v)] \\
&= FT[E(u,v) \otimes E(u,v)]
\end{aligned}
$$

$\xleftrightarrow{\text{FT}}$ $FT[A(\ell,m)] = E(u,v) \otimes E(u,v)$

NRAO

Theory of Mosaicking: Ekers & Rots Theorem

"An interferometer measures $\lambda/(b-D) < \theta < \lambda/(b+D)$"

Motivation/Derivation:

$$V(u,v) = \int\int d\ell\, dm\, A(\ell,m) I(\ell,m)\, e^{-2\pi(u\ell+vm)} = FT[A(\ell,m)I(\ell,m)]$$

$$= FT[A(\ell,m)] \otimes FT[I(\ell,m)]$$

$$= [E(u,v) \otimes E(u,v)] \otimes FT[I(\ell,m)]$$

$$A(\ell,m) = \left| \int\int_{aperture} du\, dv\, E(u,v)\, e^{-2\pi(u\ell+vm)} \right|^2$$

$$= FT[E(u,v)]\, FT[E(u,v)]$$

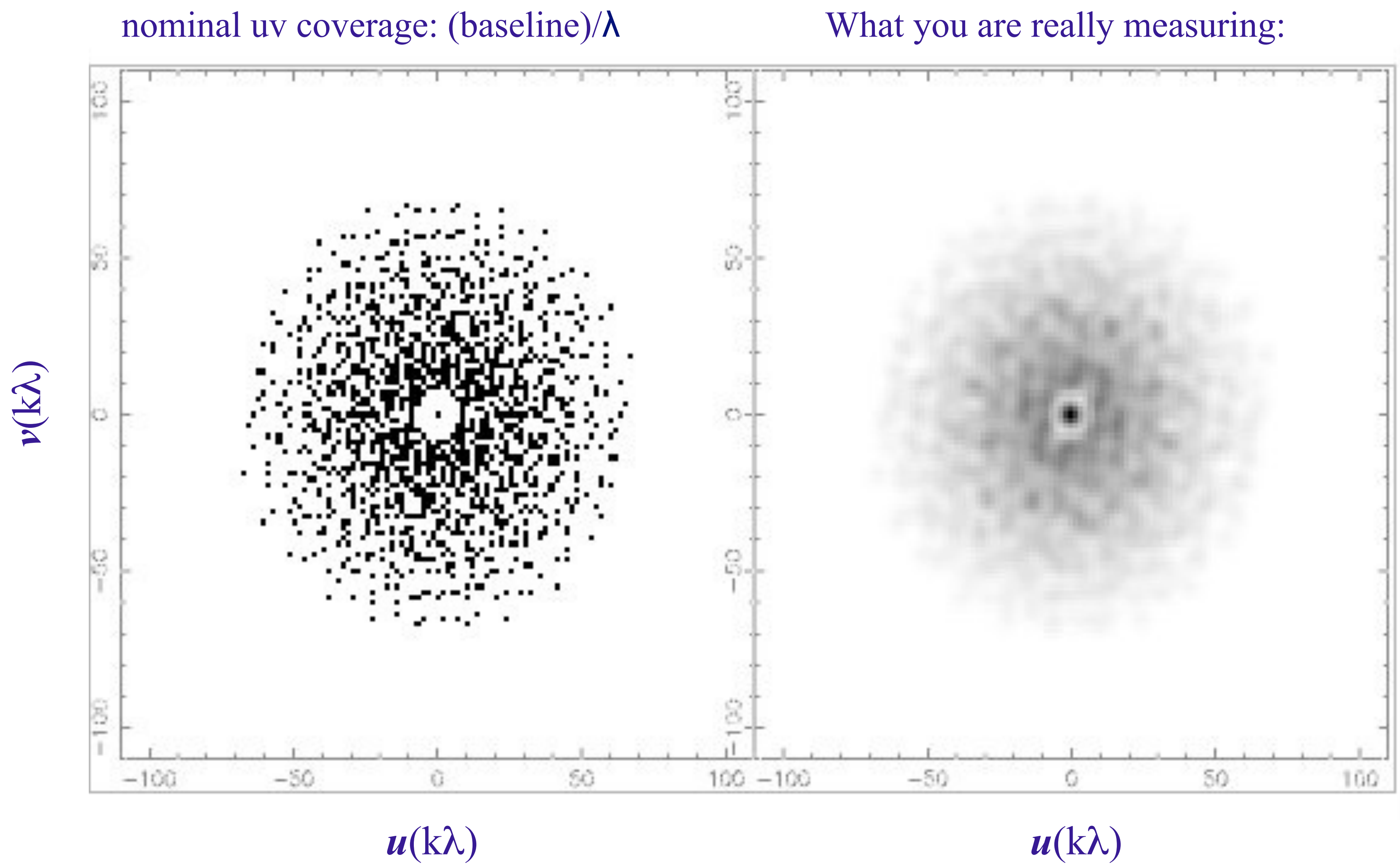
$$= FT[E(u,v) \otimes E(u,v)]$$

Auto-correlation of aperture plane illumination function; support within r=(0,+D)

$$FT[A(\ell,m)] = E(u,v) \otimes E(u,v)$$

NRAO

Theory of Mosaicking: Ekers & Rots

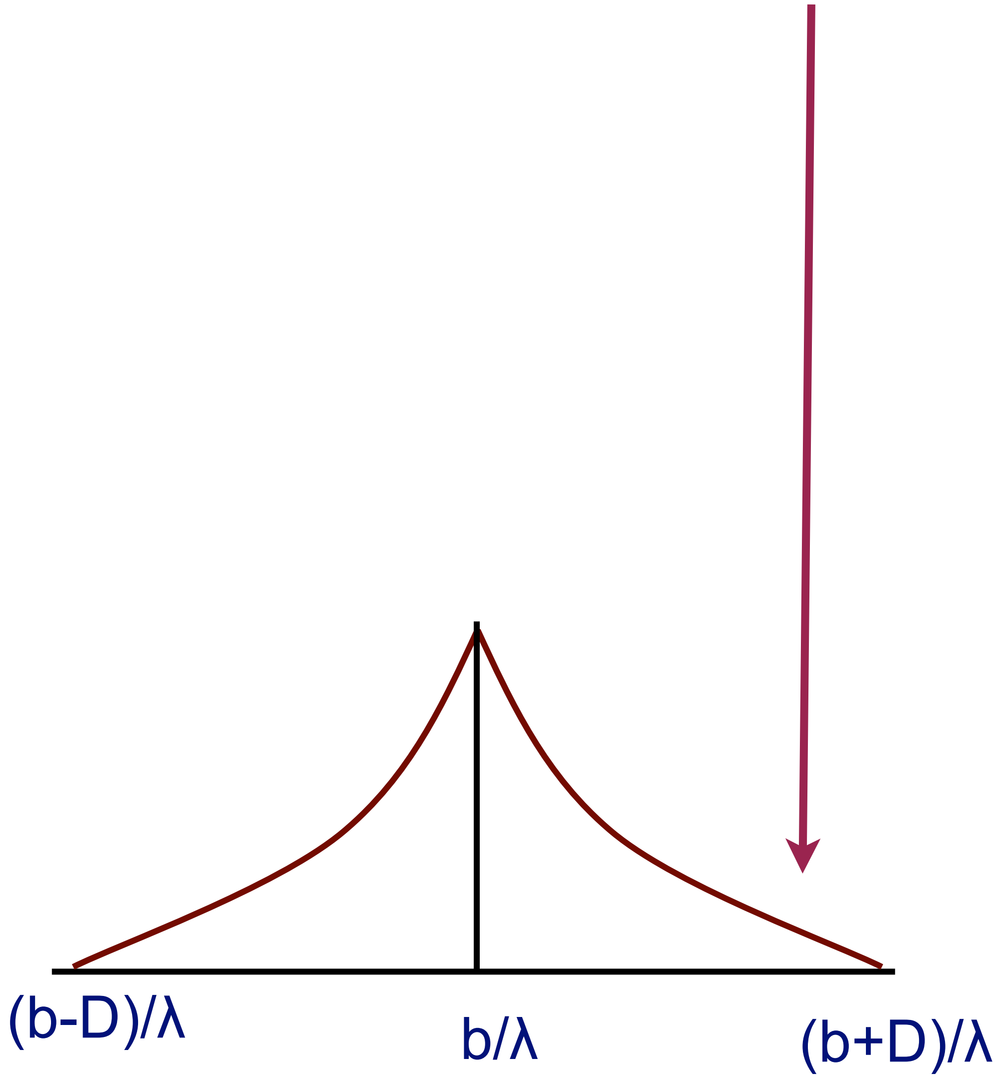


Interferometer + Single Dish

The problem:

You want to separately estimate many Fourier component amplitudes between (b-D)/λ and (b+D)/λ, but you have measured only a single complex visibility!

(a single dish has the same problem)



The problem:

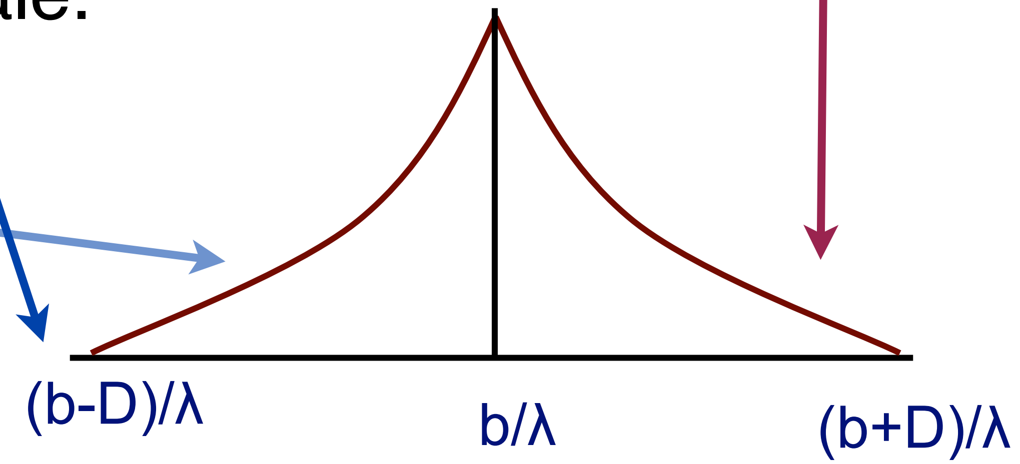
You want to separately estimate many Fourier component amplitudes between (b-D)/λ and (b+D)/λ, but you have measured only a single complex visibility!

Solution: scan the telescope over the sky and measure the visibility (V) multiple times.

i.e. - make a mosaic!

This allows you to separate out the the Fourier modes each measurement contains, increasing the maps' Fourier resolution & Largest (useful) Angular Scale.

Caveat: *signals away from b are attenuated so not measured as well.*

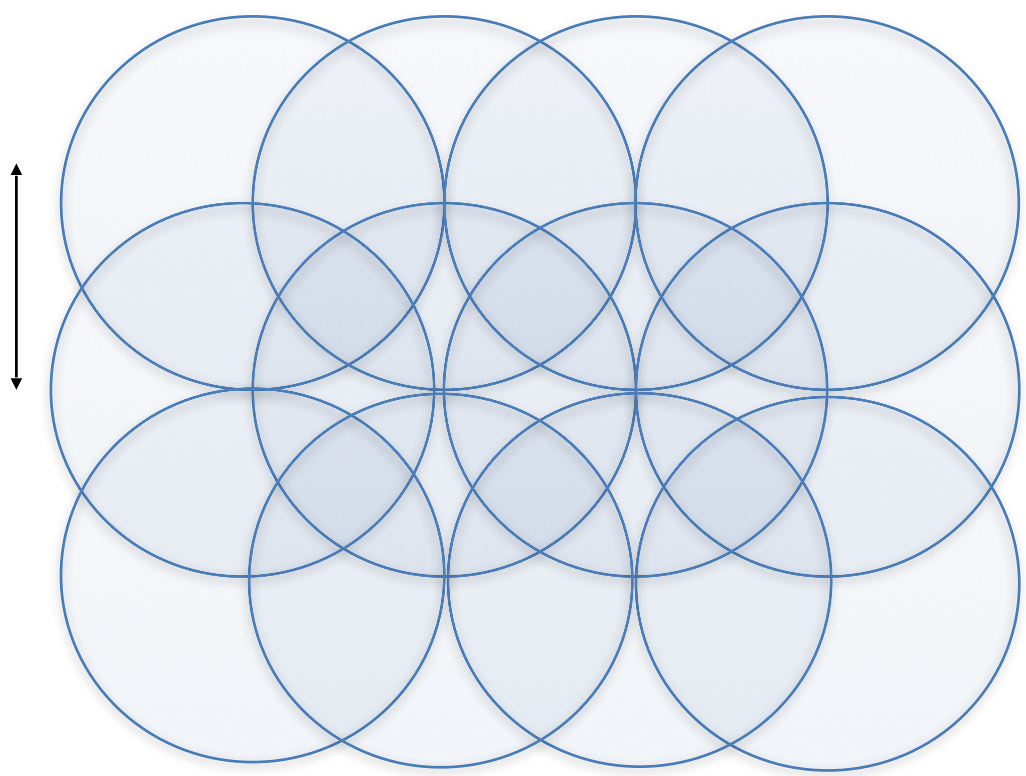


Choice of Pointings

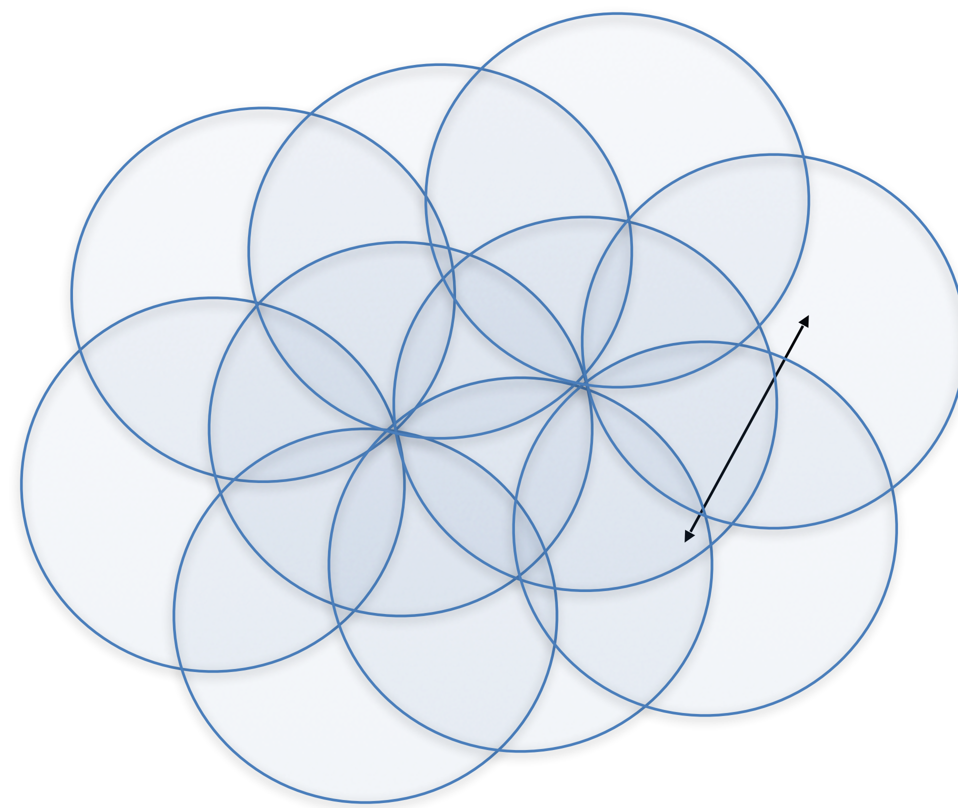
- Different ways to layout the grid on the sky:
- Theoretically optimal sampling (Cornwell 1988):

Rectangular grid

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



Hexagonal grid



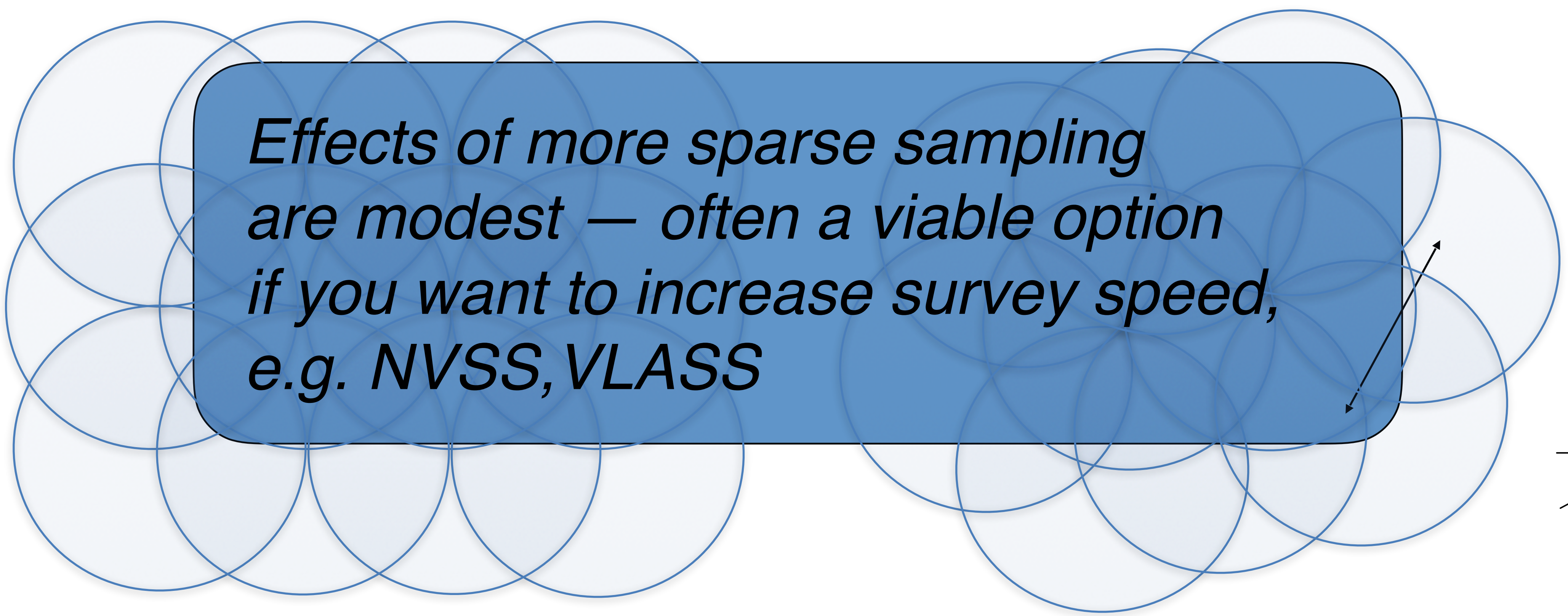
$$\frac{2}{\sqrt{3}}\frac{\lambda}{2D}$$

Preferred - very uniform image domain noise

Choice of Pointings

- Different ways to layout the grid on the sky:
- Theoretically optimal sampling (Cornwell 1988):

Rectangular grid | Hexagonal grid



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

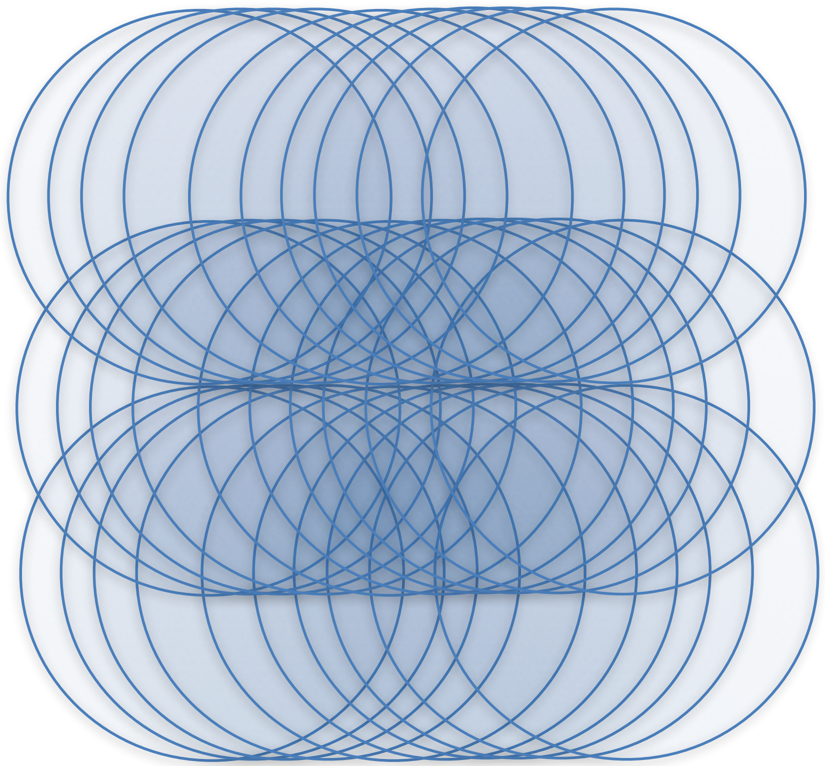
$$\frac{2}{\sqrt{3}}\frac{\lambda}{2D}$$

Effects of more sparse sampling are modest — often a viable option if you want to increase survey speed, e.g. NVSS,VLASS

Preferred - very uniform image domain noise

Choice of Pointings

- On-The-Fly Interferometry - analogous to single dish "On-the-fly Mapping"



Scan continuously, dumping correlations & all antenna positions rapidly; high data rate, low overhead.

VLA Sky Survey ; ALMA (future)

Stitching the Interferometer Maps together: Mosaic Imaging Algorithms in Practice

Widely-used methods for mosaic image reconstruction:

- **Linear combination**

  Make individual ptg dirty maps → deconvolve individually → combine deconv'd maps

- **Joint deconvolution**

  Make individual ptg dirty maps → combine into one dirty map → deconvolve together

  (w/spatially varying PSF)

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

  Combine visibilities from all pointings in uv-space → single dirty map → deconvolve

Stitching the Interferometer Maps together: Mosaic Imaging Algorithms in Practice

Widely-used methods for mosaic image reconstruction:

- **Linear combination**

  Make individual ptg dirty maps → deconvolve individually → combine deconv'd maps

- **Joint deconvolution**

  Make individual ptg dirty maps → combine into one dirty map → deconvolve together

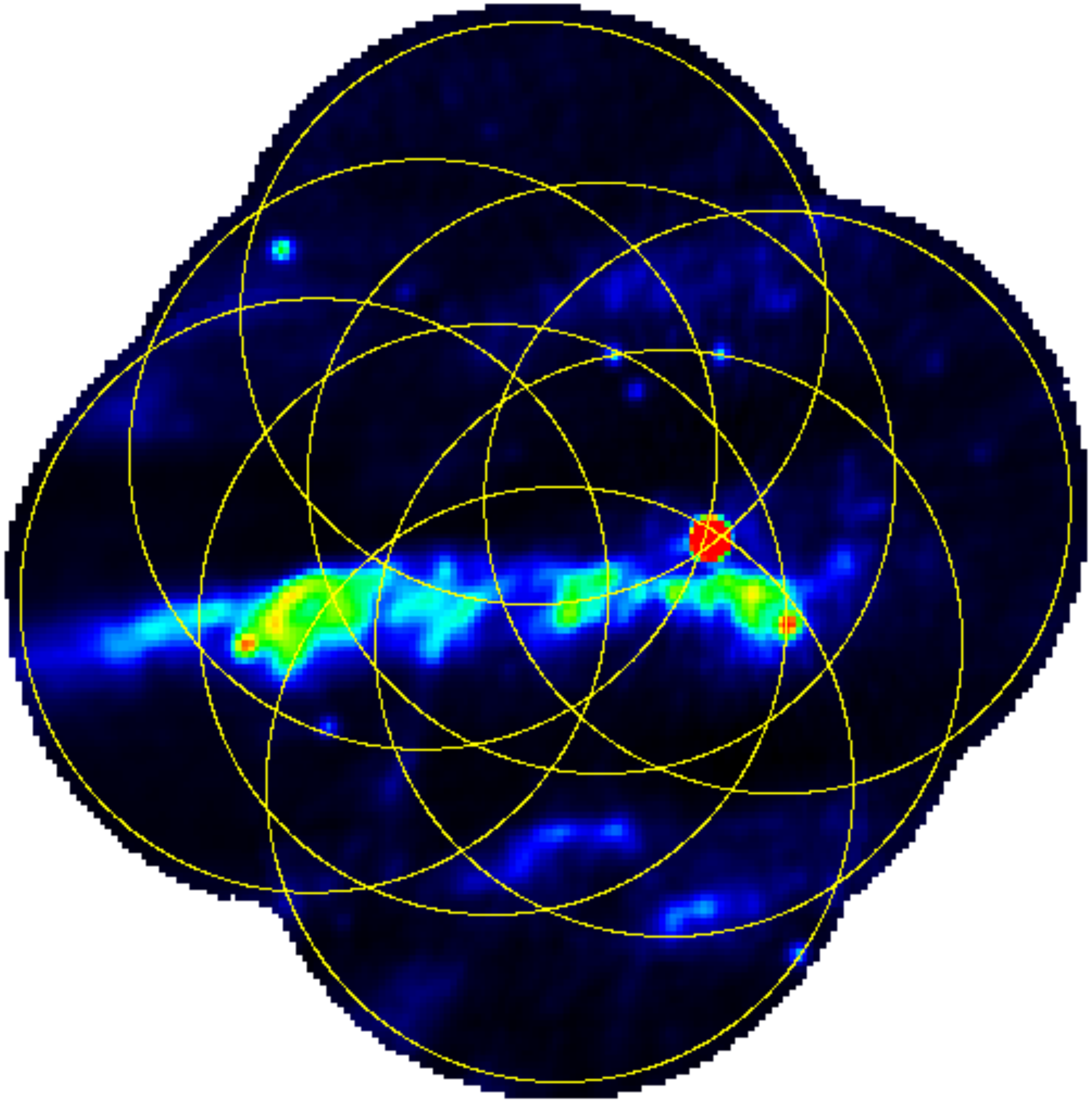
  (w/spatially varying PSF)

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

  Combine visibilities from all pointings in uv-space → single dirty map → deconvolve

U.Rao will discuss advanced algorithms Monday
(e.g. A-projection, dealing with non-coplanar baselines)

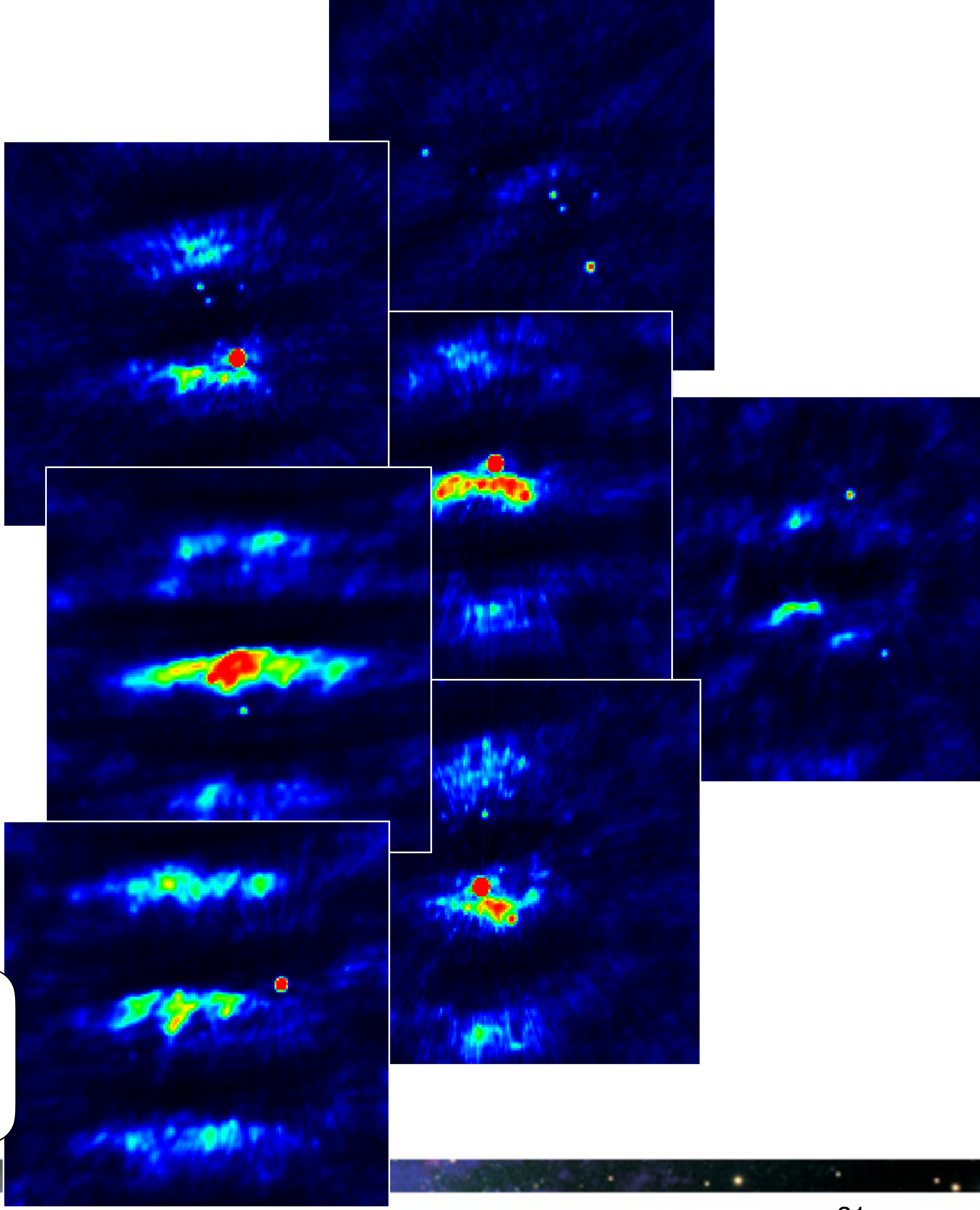
Linear Mosaic – observe pointings



Linear Mosaic – individual images

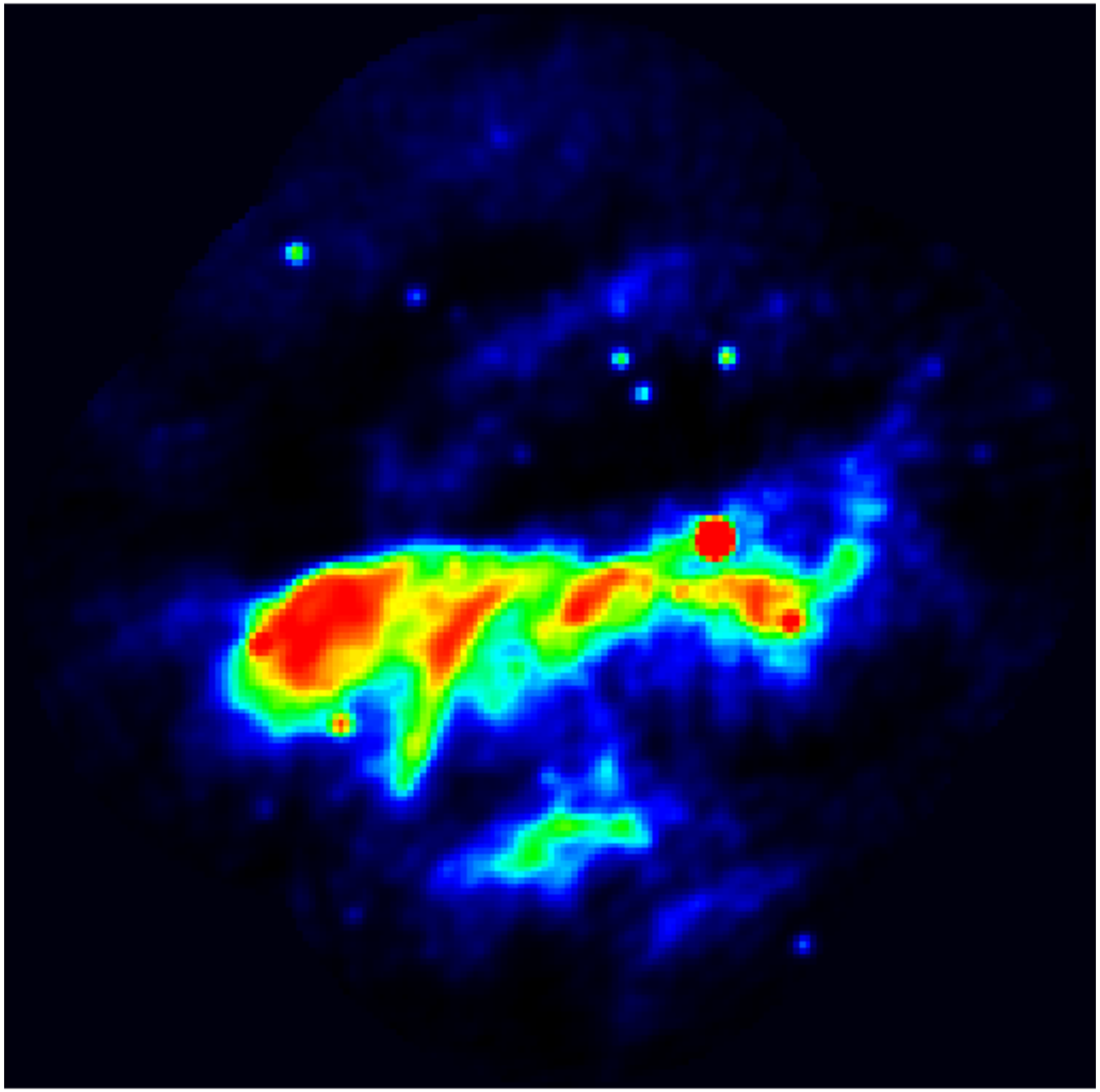
- Treat each pointing separately
- Image & deconvolve each pointing
- Stitch together linearly with optimal pointing weights from noise and primary beam

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

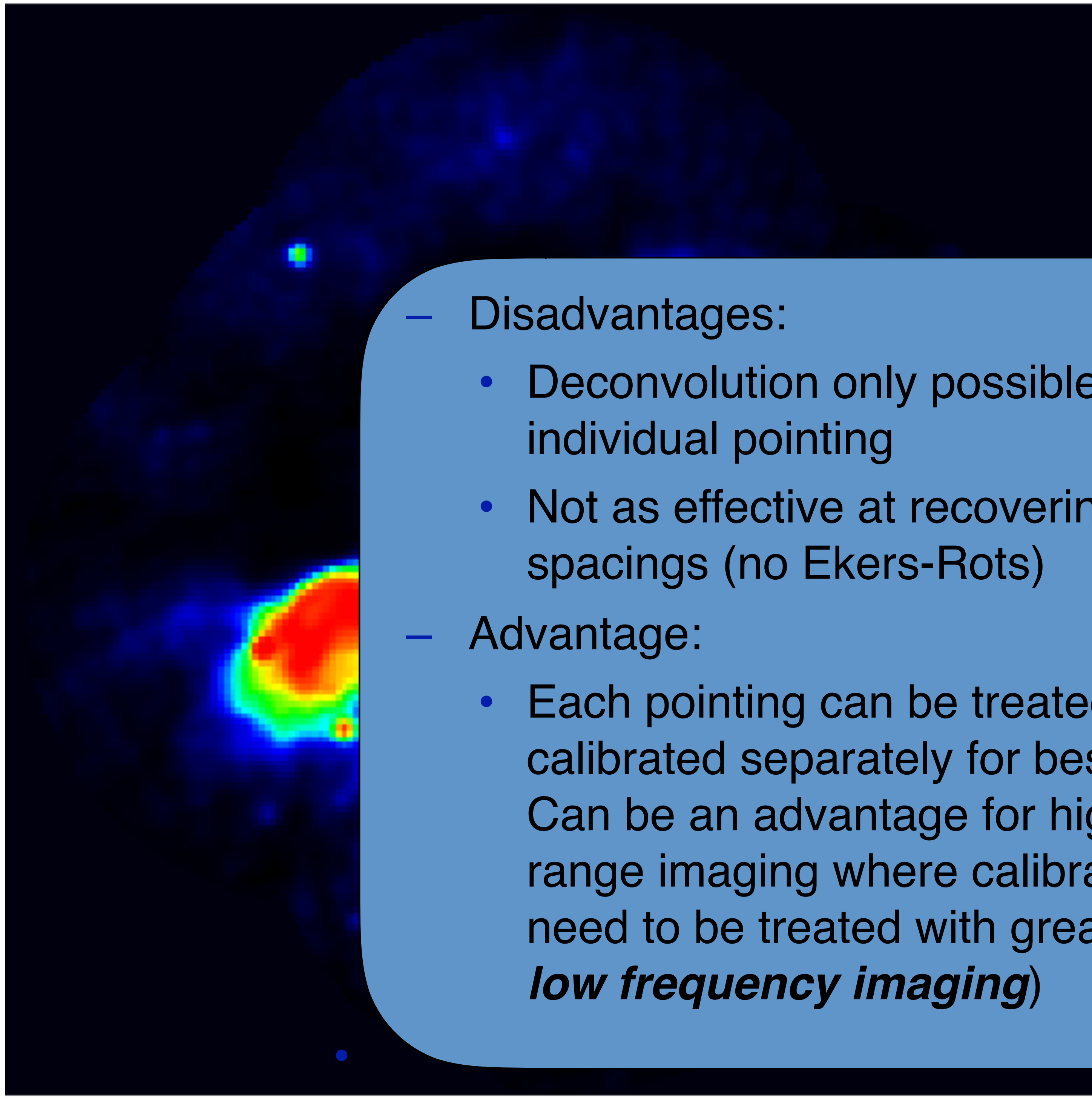


NRAO

Linear Mosaic – combine pointings



Linear Mosaic – combine pointings

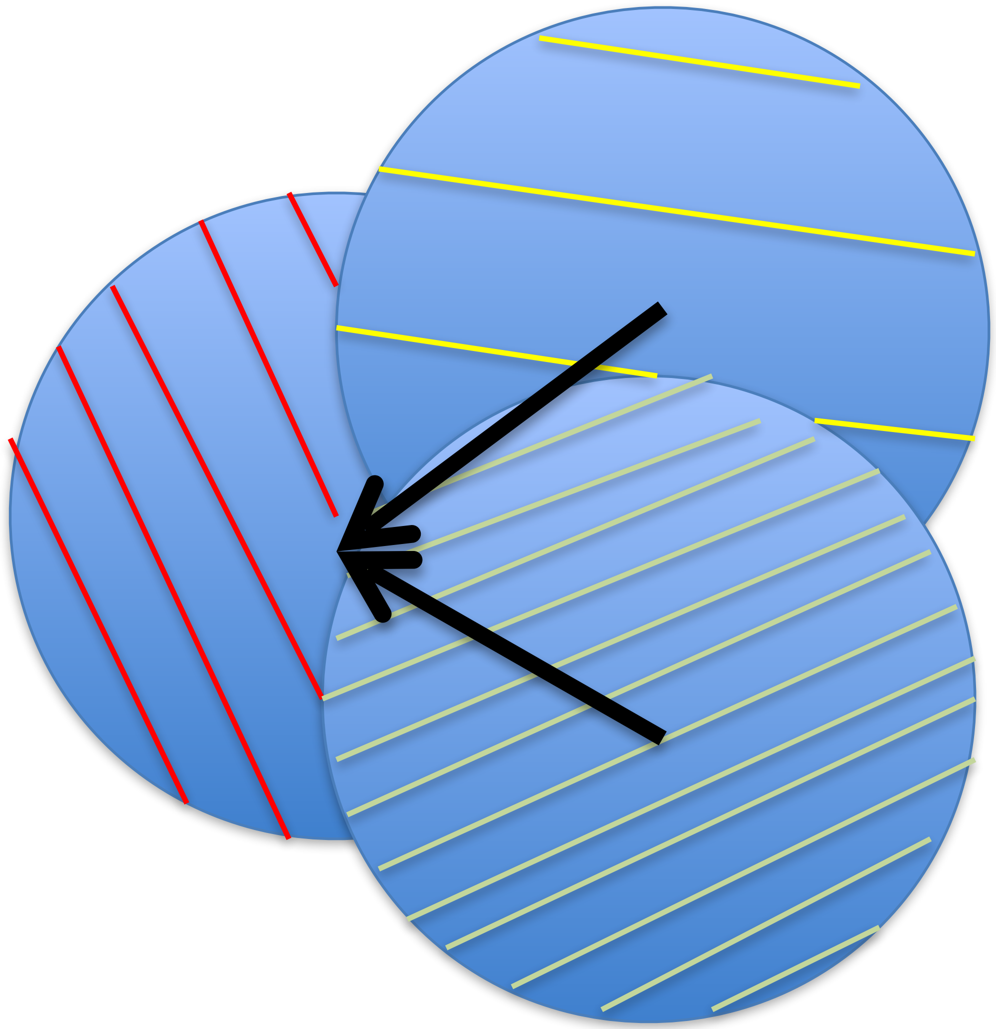


- Disadvantages:
  - Deconvolution only possible to depth of individual pointing
  - Not as effective at recovering shorter spacings (no Ekers-Rots)
- Advantage:
  - Each pointing can be treated and calibrated separately for best results. Can be an advantage for high-dynamic range imaging where calibration effects need to be treated with great care (e.g., ***low frequency imaging***)

Widefield Imaging

Combine data from different pointings in uv domain, then deconvolve

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



Widefield Imaging

Combine data from different pointings in uv domain, then deconvolve

- Take each uv data for each pointing and shift to a common phase reference center.
- re-grid all visibilities to a common UV plane (PB kernel).
- FT to a single “dirty image” with a common PSF
  » *Deconvolve*

ADVANTAGES

- Uses all uv info per overlap → better beam, deeper clean
- deconv. has all the (Ekers-Rots) information at every point in the sky: more large-scale structure recovered
- Works well with on-the-fly interferometry data (many, many pointing centers)
- Naturally works well with heterogeneous arrays (different sized antennas)

Cost: you need to know your PB well

Mosaicking in CASA (simple use case)

- Calibrate as you would do for a single pointing (e.g. pipeline)
- Use the **tclean** task with your favorite parameters
  - (current "clean" is deprecated and will go away)
- in **tclean** parameter *gridder* use *'mosaic'* for joint, wide-field imaging (preferred)
  - Uses Cotton-Schwab (major/minor cycle) algorithm
  - Use *deconvolver='hogbom'* (default, best for poor psf) *or 'clark' (faster)*
    - *deconvolver='mtmfs'* (wide bandwidth continuum)
    - *deconvolver='multiscale'*
  - currently only ALMA supported as Heterog. Array
- Linear mosaicking of cleaned images only available at present from the CASA toolkit (`im.linearmosaic`). [AIPS FLATN]

Interferometric Mosaicking Issues

- Pointings are in a time sequence:
  - Each pointing has a different uv-coverage
  - Atmospheric water vapor/Ionospheric variations from pointing to pointing
- Pointing is more critical than for non-mosaicked observation with an isolated source in the beam center

Deconvolution

Mosaicking is often done for ***extended*** sources.

Deconvolution in this case is tricky.

Deconvolution

Mosaicking is often done for ***extended*** sources.

Deconvolution in this case is tricky.

You need to clean deeply (~1σ) for extended emission.

Justification: in general the "CLEAN model" is not your best estimate of the sky; the reconvolved CLEAN model+residuals is.

- BUT **Do not do this** if you are going to self-cal using the CLEAN model! (consider multi-scale)
- helps to have good uv coverage, a judiciously chosen clean box, & careful monitoring (interactive)
- may take a long time for a spectral line cube

NRAO

Deconvolution

Mosaicking is often done for ***extended*** sources.

Deconvolution in this case is tricky.

CLEAN: Issues to be aware of

- "CLEAN Bias": constructive interference of synthesized beam sidelobes can make them appear higher than the main lobe of the synth. beam.
  - Reduces the apparent source fluxes recovered
  - most severe for extended sources
  - *mitigated by good UV coverage (lower sidelobes), good masking.*
  - see Condon et al. (1998) [NVSS survey paper]

- Mismatch of Clean & Dirty Beams: beam areas differ within relevant apertures, biasing integrated flux density values upward.
  - mitigated by deeper cleaning, correction factor
  - see Jorsater & VanMoorsel (1995) and Walter et al. (2008)

NRAO

Deconvolution

Mosaicking is often done for ***extended*** sources.

Deconvolution in this case is tricky.

CLEAN: Issues to be aware of

tclean automatic clean masking algorithm can be very useful (mask='auto-multithresh')

- "CLEAN Bias": make them app...
  - Reduces the apparent source fluxes recovered
  - most severe for extended sources
  - *mitigated by good UV coverage (lower sidelobes),* ***good masking****.*
  - see Condon et al. (1998) [NVSS survey paper]

- Mismatch of Clean & Dirty Beams: beam areas differ within relevant apertures, biasing integrated flux density values upward.
  - mitigated by **deeper cleaning,** correction factor
  - see Jorsater & VanMoorsel (1995) and Walter et al. (2008)

NRAO

Deconvolution

Mosaicking is often done for ***extended*** sources.

Deconvolution in this case is tricky.

Multi-Scale CLEAN

✳Generalize CLEAN to allow components of multiple sizes

✳Obviously better suited to extended emission!

✳Fully supported in CASA **tclean**() task

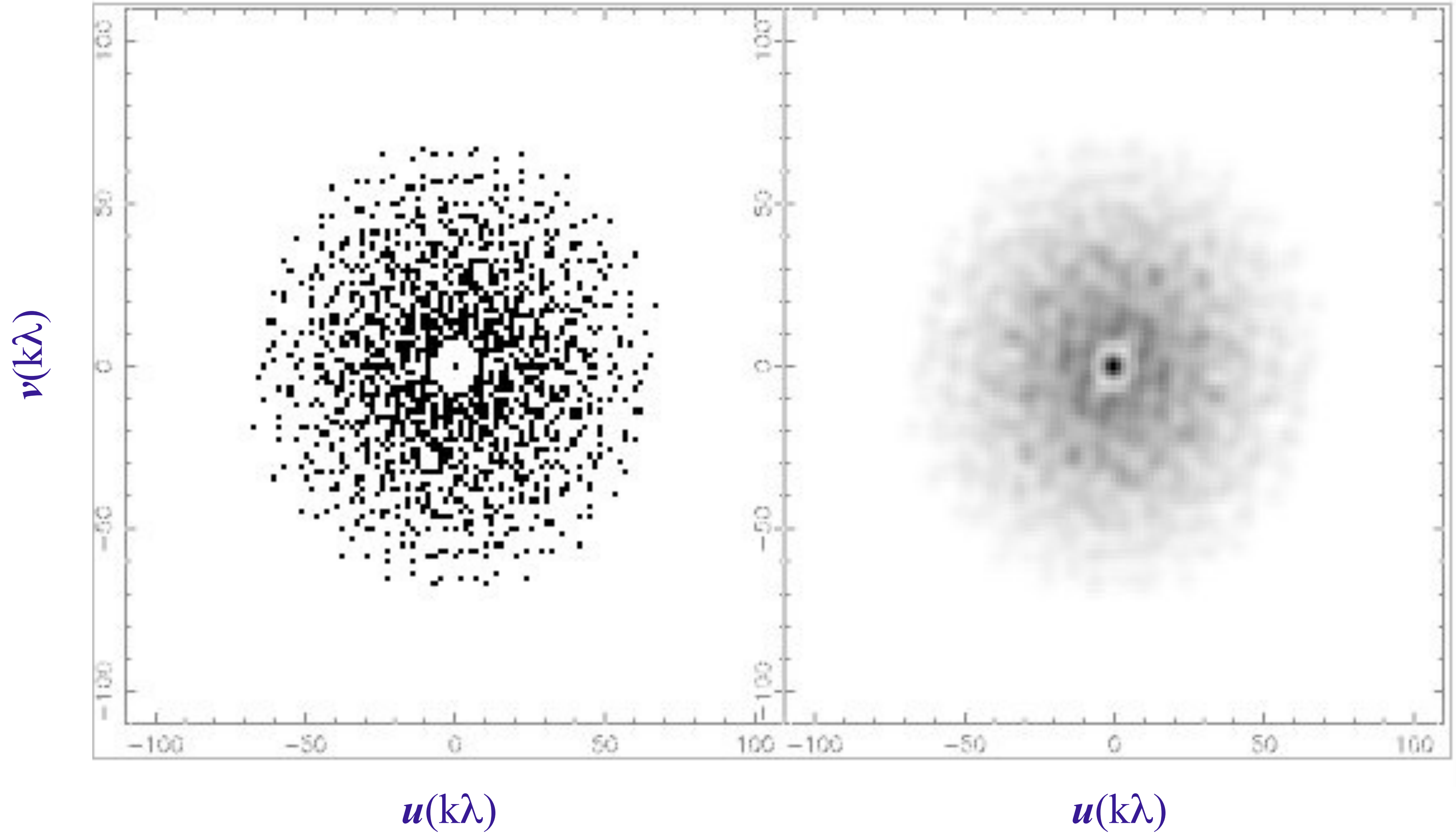
See talks by D.Wilner, U.Rao

Effects of Missing Short & Zero Spacings

Interferometer + Single Dish

nominal uv coverage: (baseline)/λ

What you are really measuring:

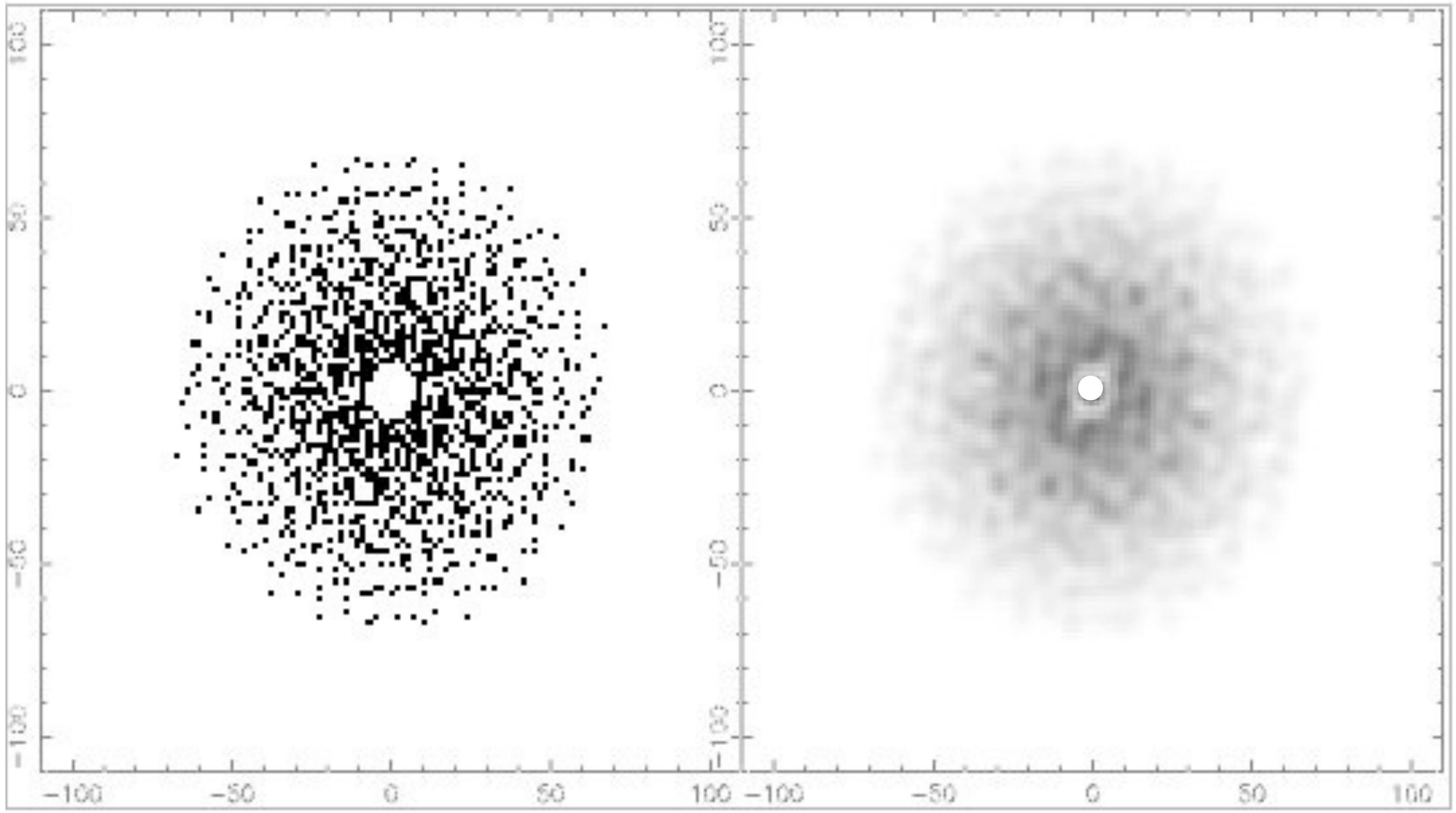


Effects of Missing Short & Zero Spacings

Interferometer ~~+ Single~~ Dish

nominal uv coverage: (baseline)/λ

What you are really measuring:



v(kλ)

u(kλ)

u(kλ)

NRAO

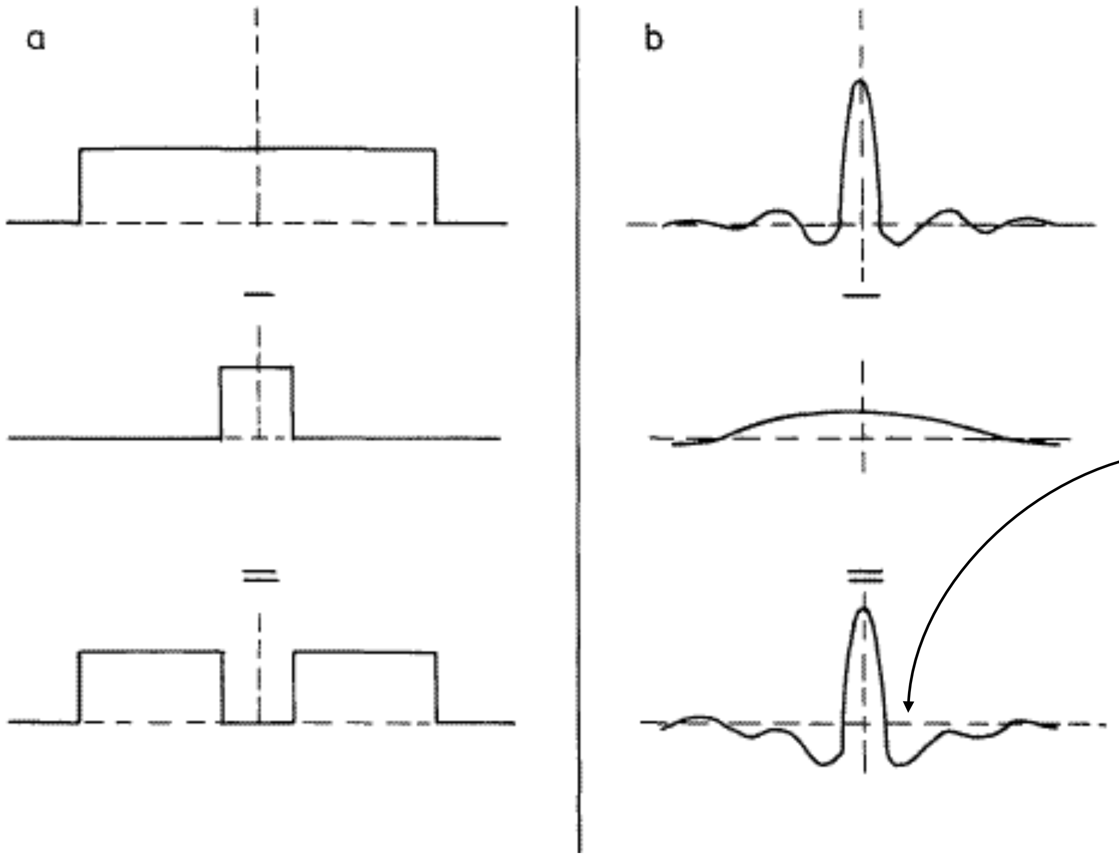
Effects of Missing Short & Zero Spacings

Interferometer ~~+ Single~~ Dish

- UV plane
- PSF

- Braun & Walterbos (1985)

- ideal
- Central hole
- Typical interfero-meter



- Negative 'bowl'

Effects of Missing Short & Zero Spacings

Interferometer ~~+ Single~~ Dish

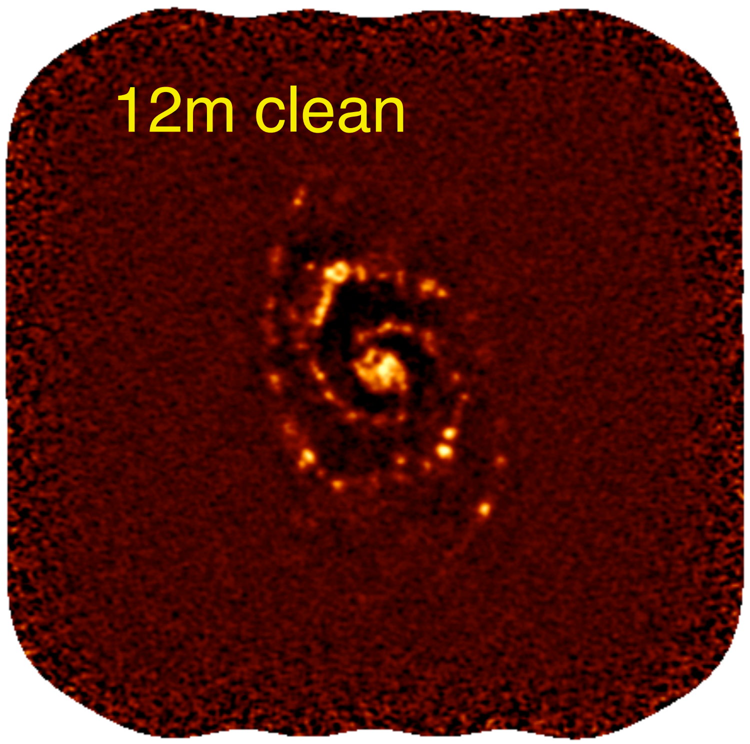
- UV plane
- PSF
- Braun & Walterbos (1985)

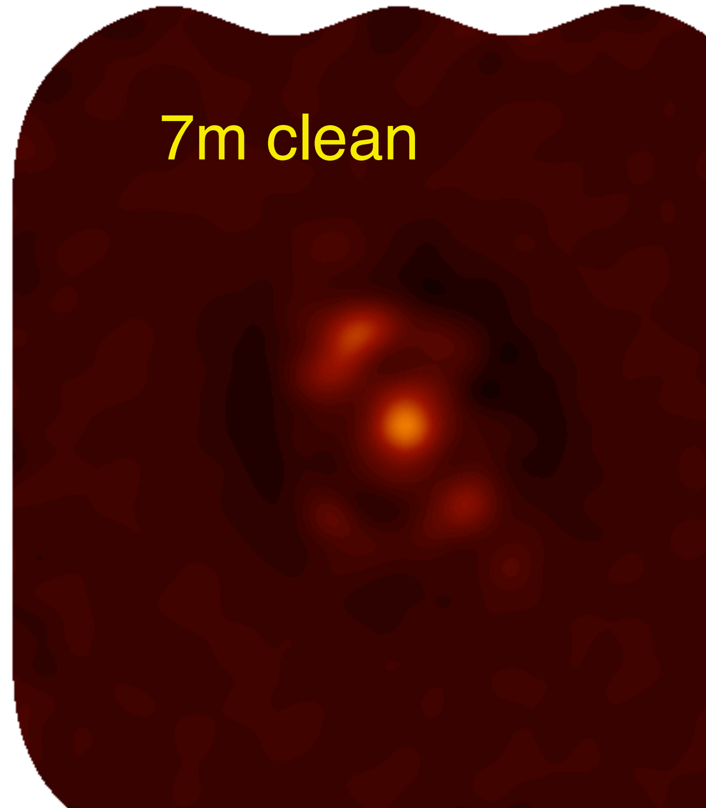
The "background level" in your map is unmeasured / variable: this is a big problem for measuring the fluxes of individual objects or regions.

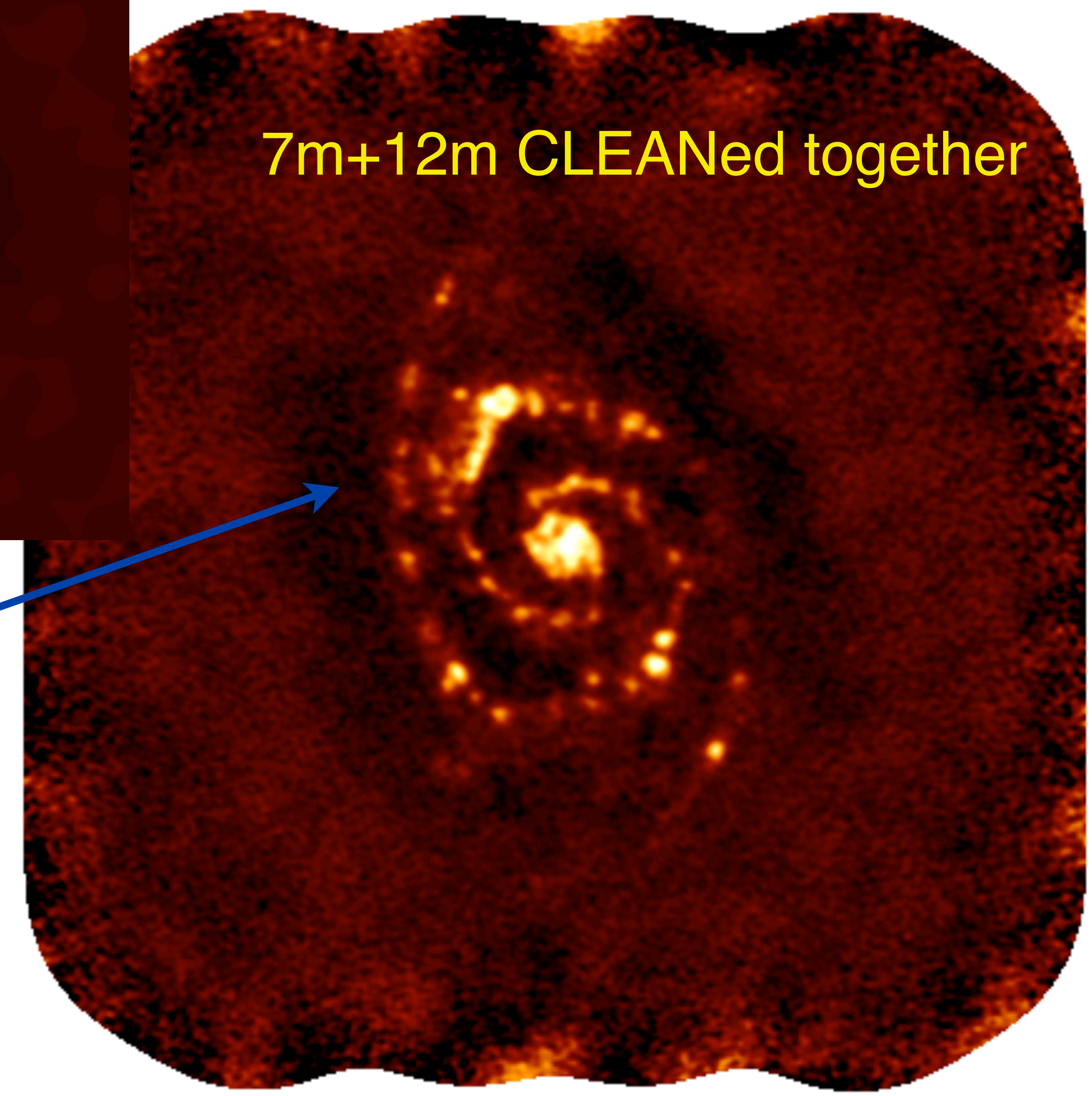
This matters because the science often comes from comparisons in different maps: the integrated line intensity in two transitions or lines; the continuum flux density at two widely separated frequencies.

(Often using data from completely different instruments...)

Effects of Missing Short & Zero Spacings



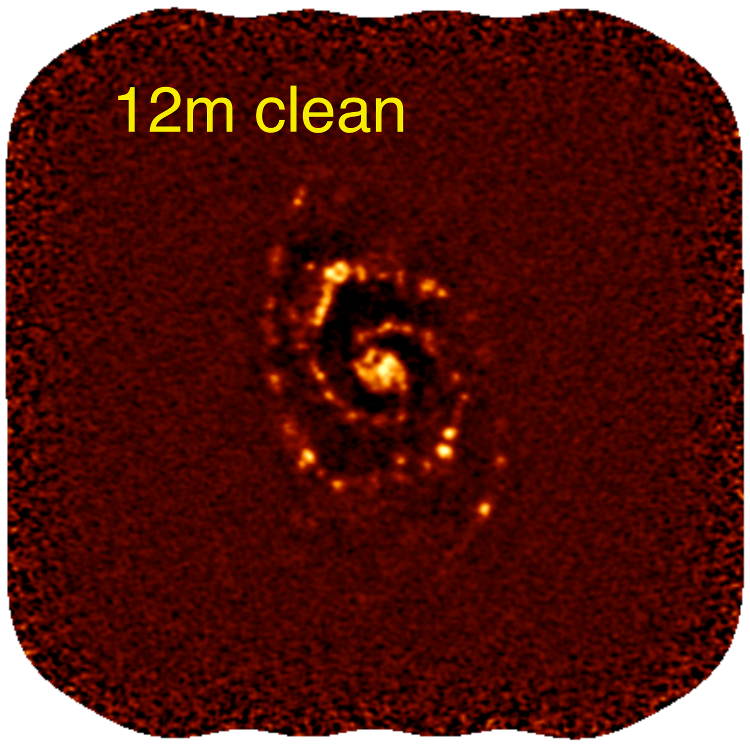


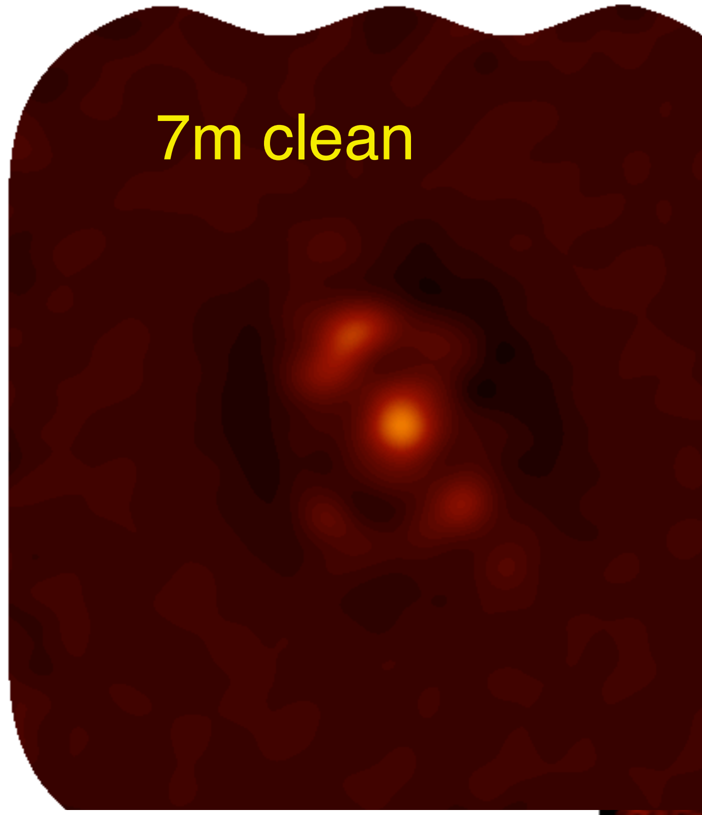


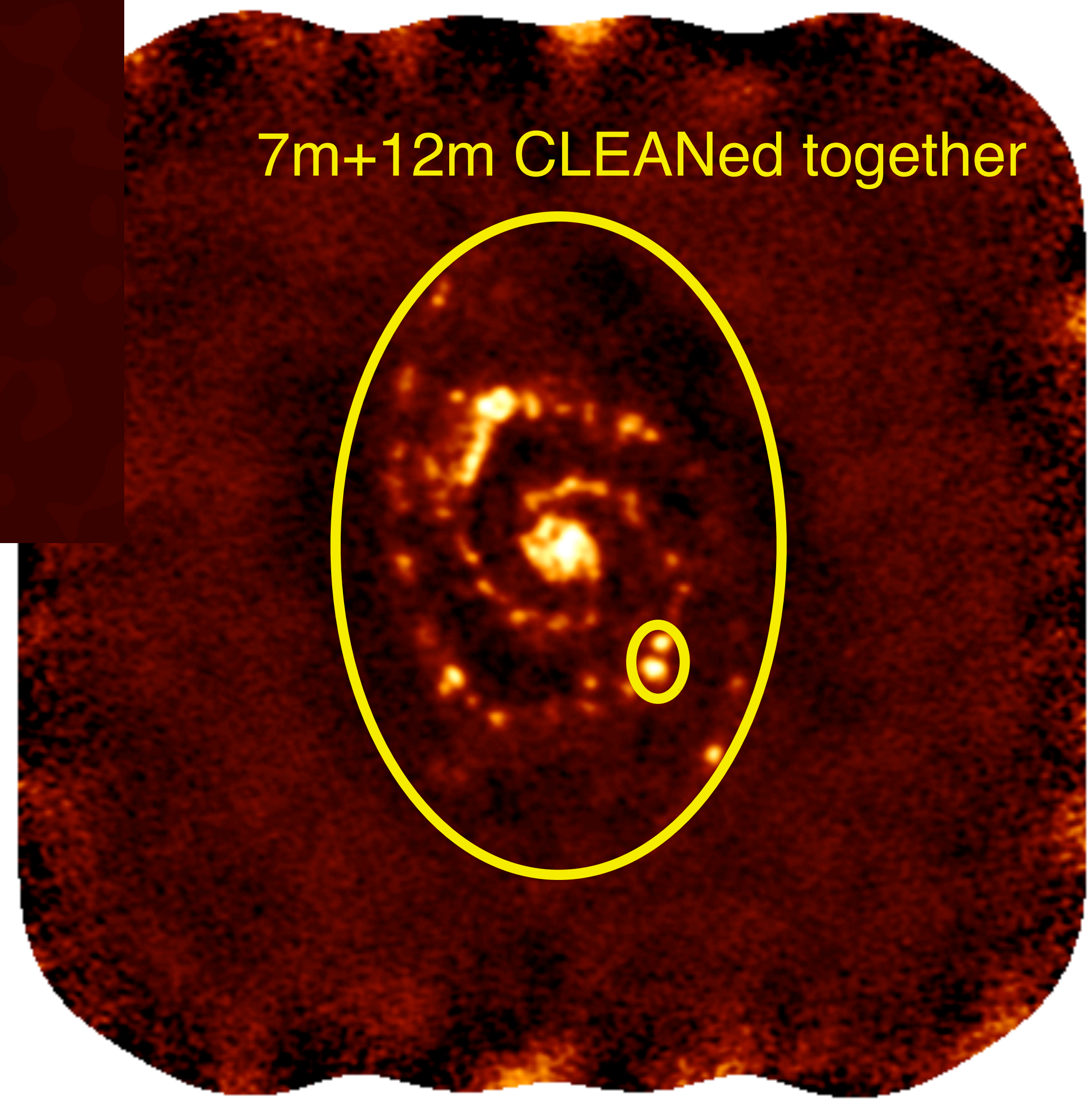
Combination of residual sidelobes (incomplete deconvolution) and poorly constrained short spacings.

https://casaguides.nrao.edu/index.php/Simalma_(CASA_5.1)

Effects of Missing Short & Zero Spacings







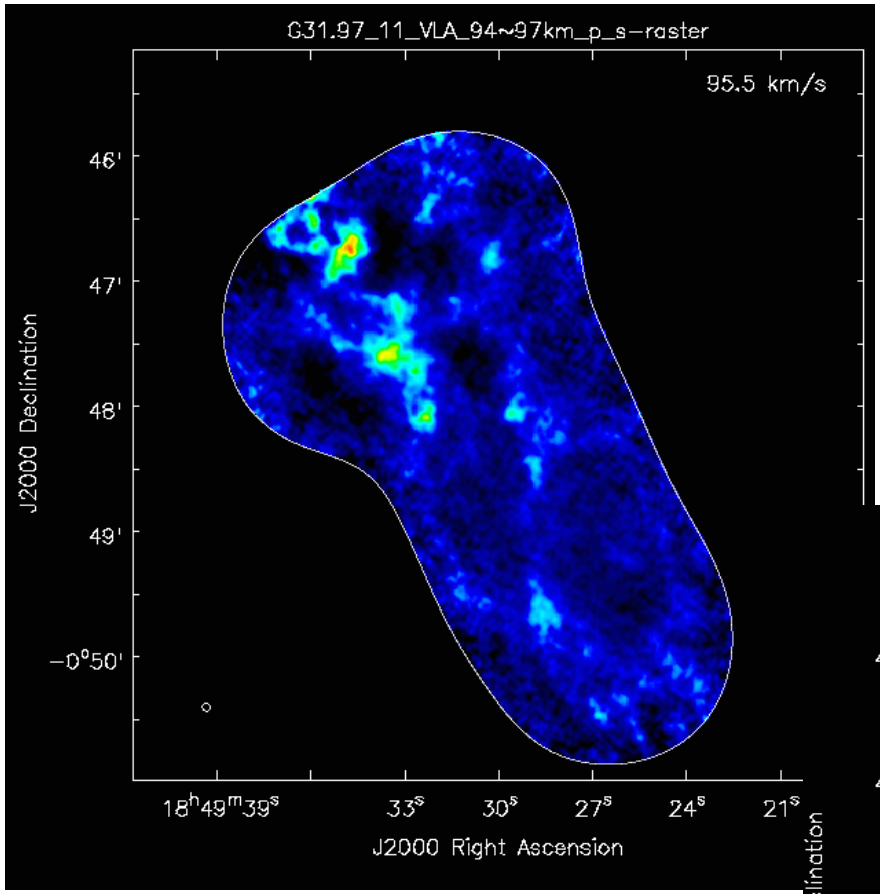
Measured total fluxes in any aperture will underestimate the true total fluxes.

maybe MS clean could do better but the ***real problem is that the short spacings are poorly constrained***.

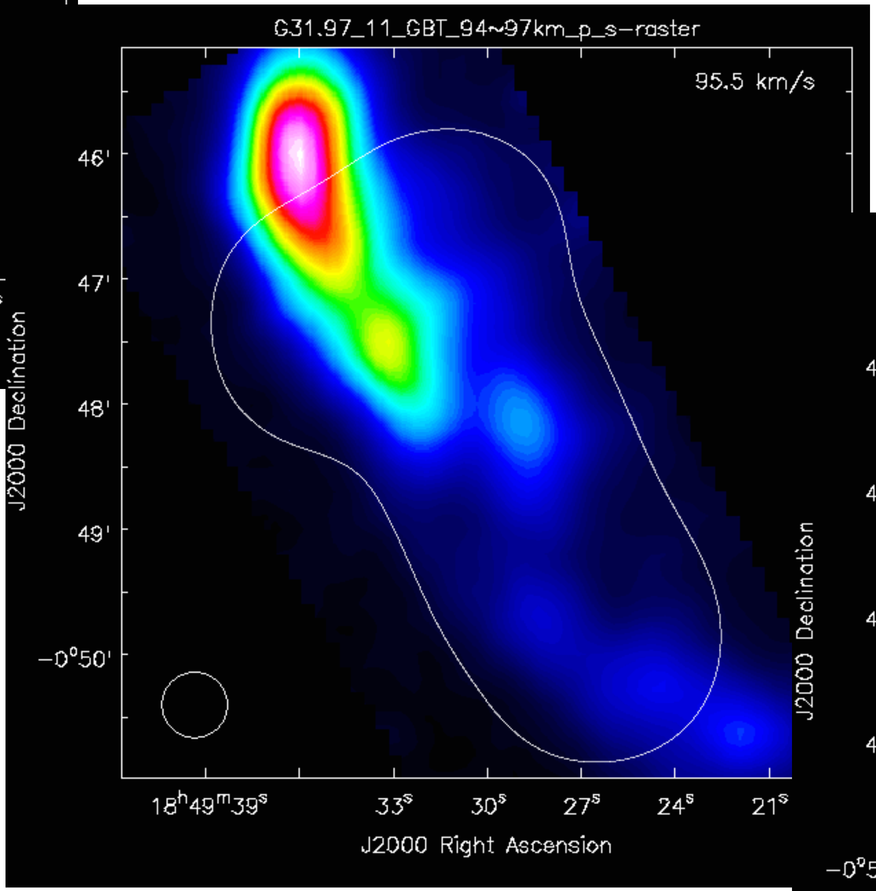
Add single dish data to the map!

https://casaguides.nrao.edu/index.php/Simalma_(CASA_5.1)

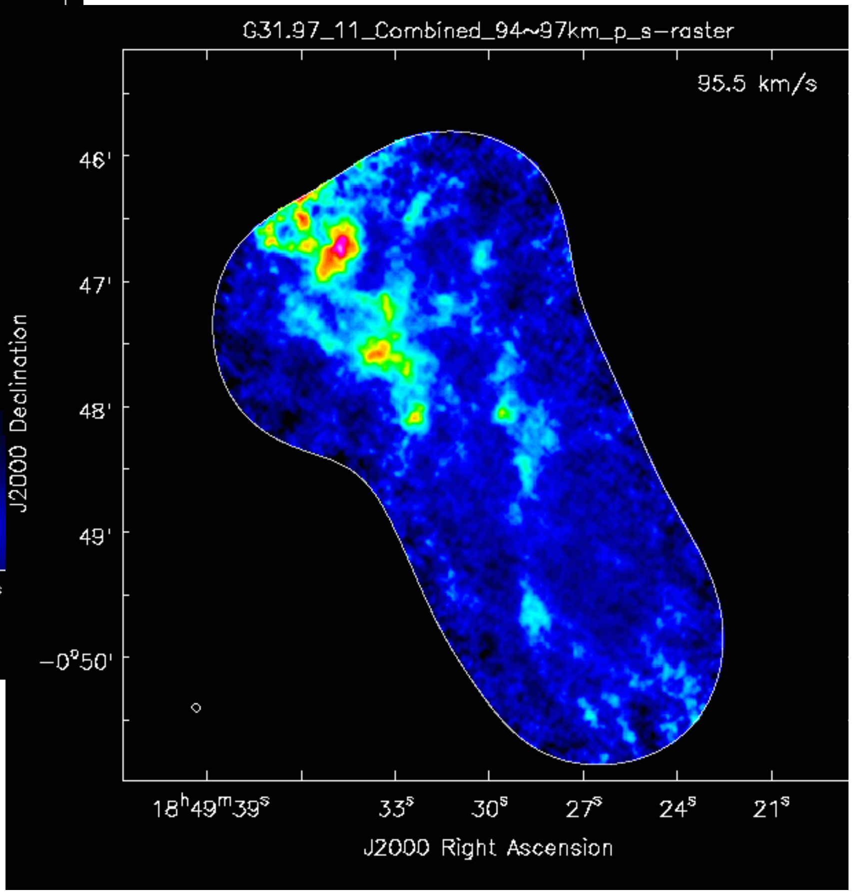
EVLA NH_3 (multi-scale CLEANed)



GBT NH_3

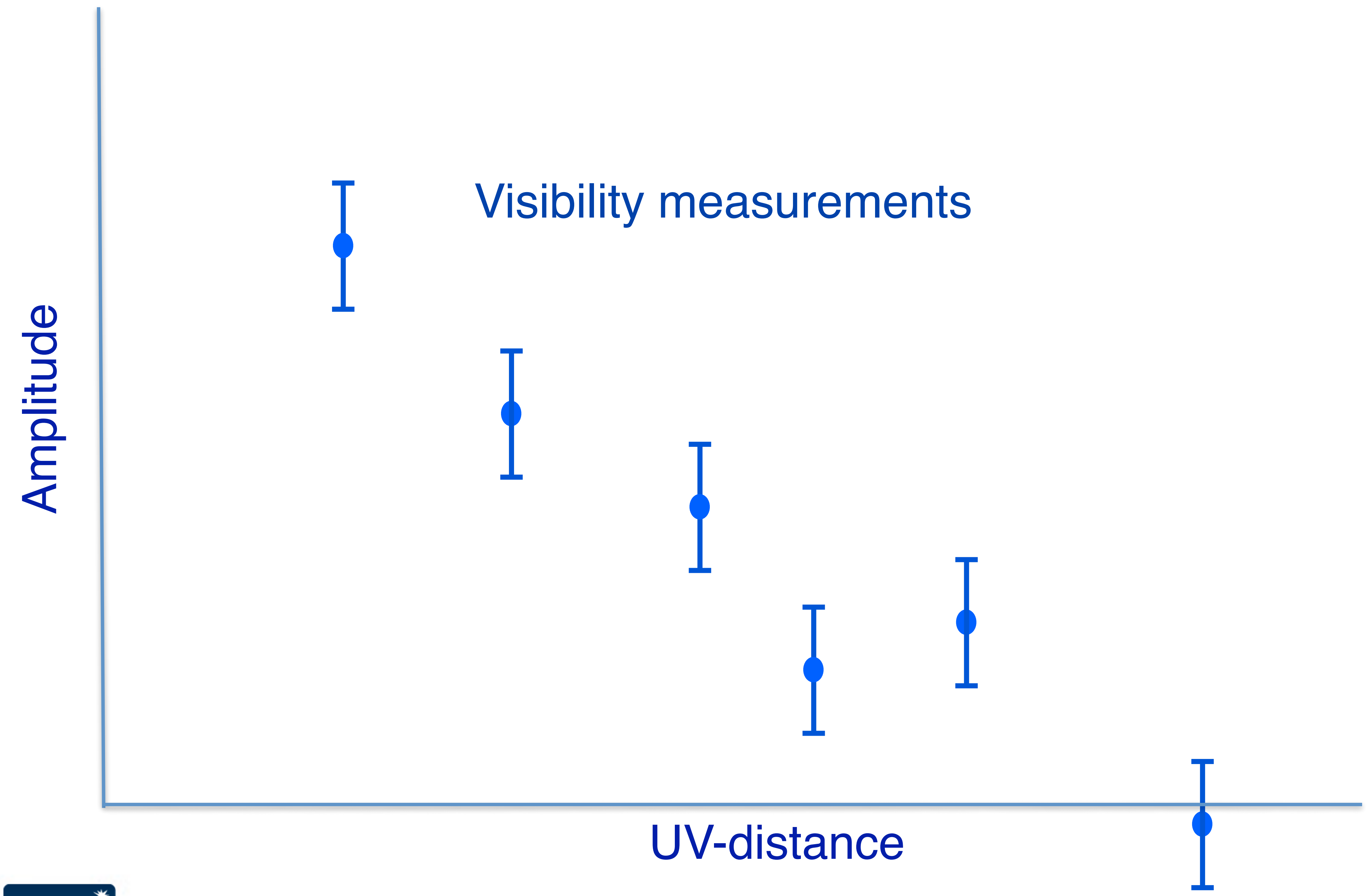


Feathered

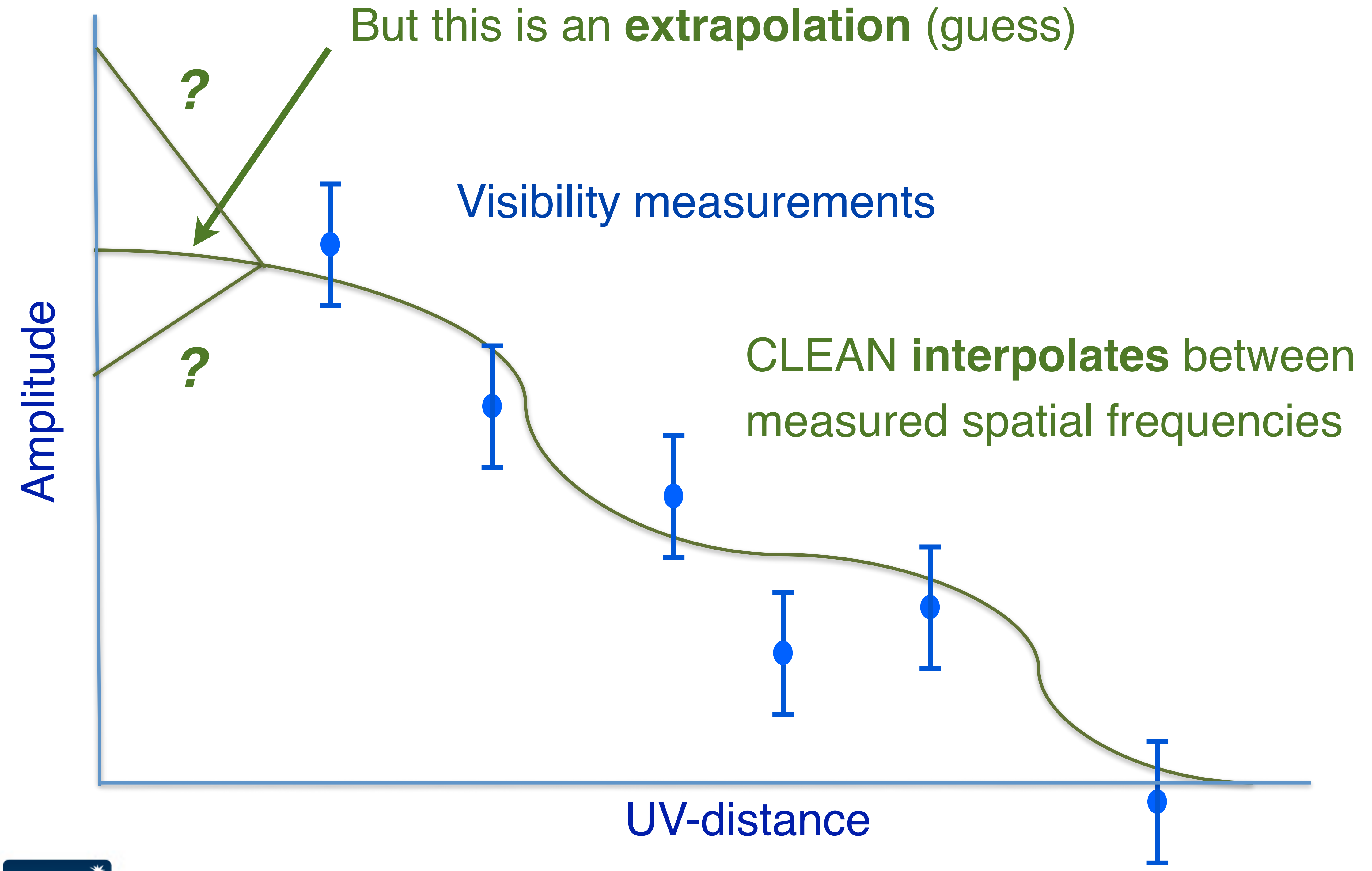


DiRienzo et al. (2015)

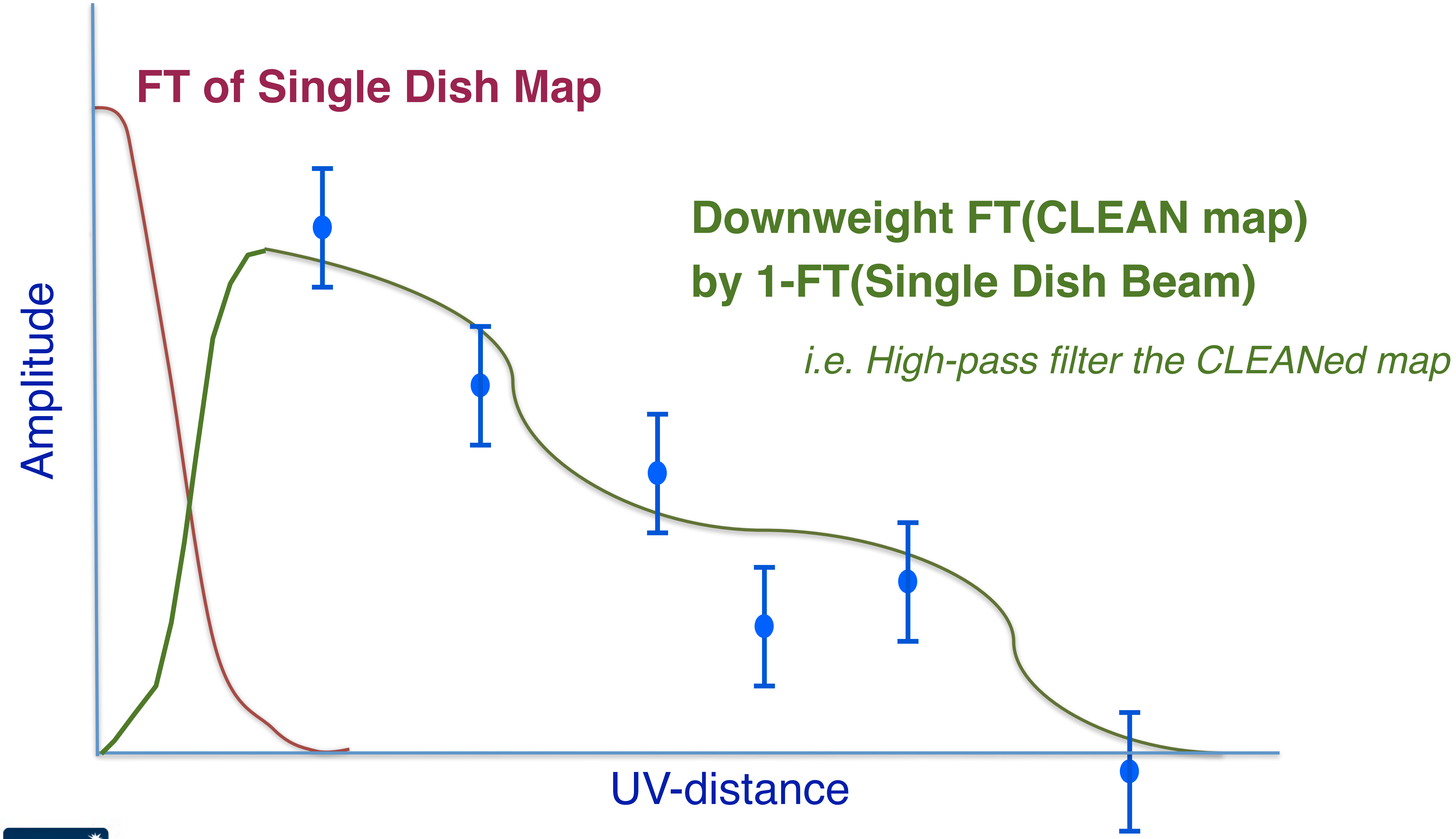
Feathering



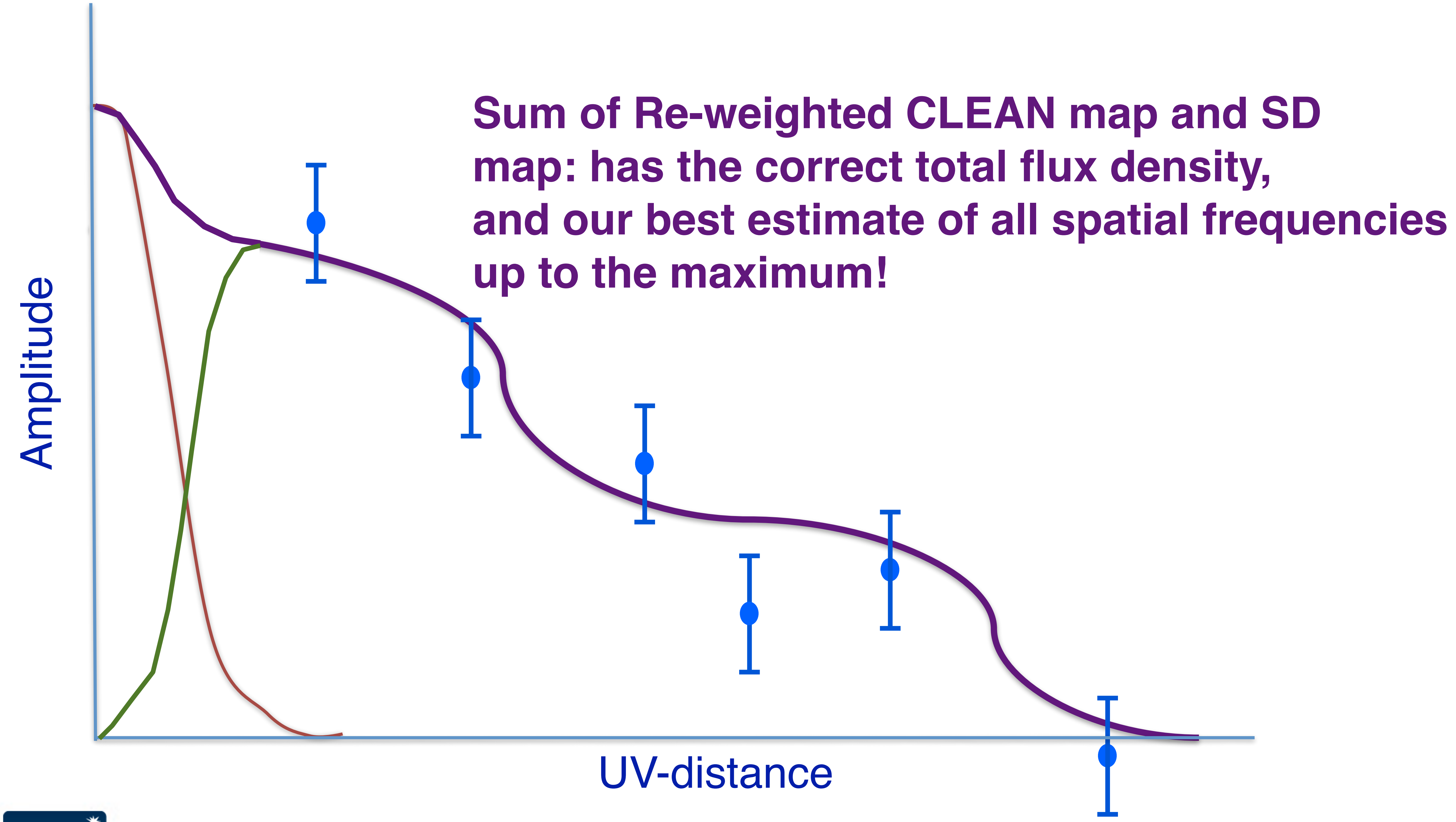
Feathering

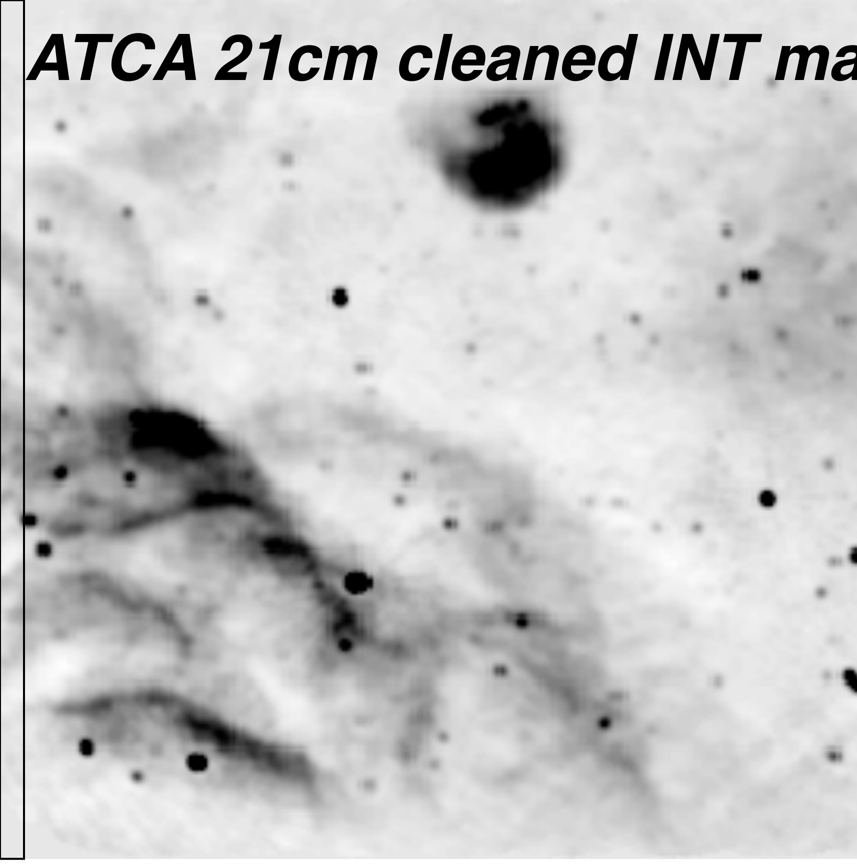

But this is an **extraplation** (guess)
?
?
Visibility measurements
CLEAN **interpolates** between measured spatial frequencies
Amplitude
UV-distance

Feathering

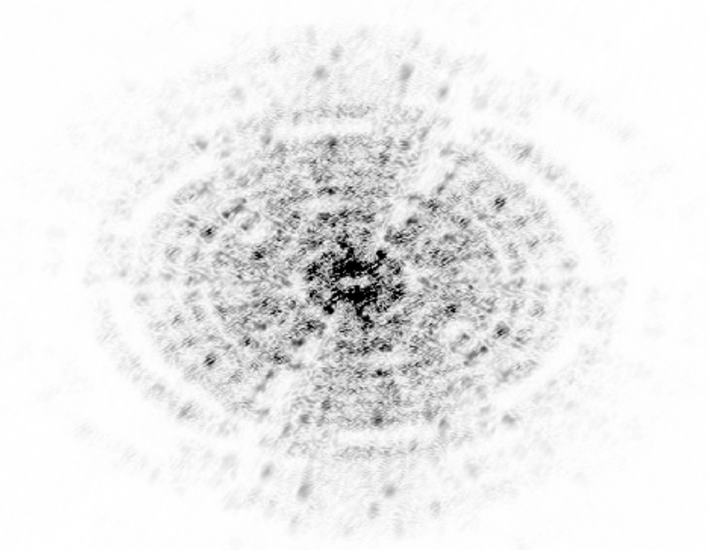

FT of Single Dish Map
Downweight FT(CLEAN map)
by 1-FT(Single Dish Beam)
i.e. High-pass filter the CLEANed map
Amplitude
UV-distance

Feathering

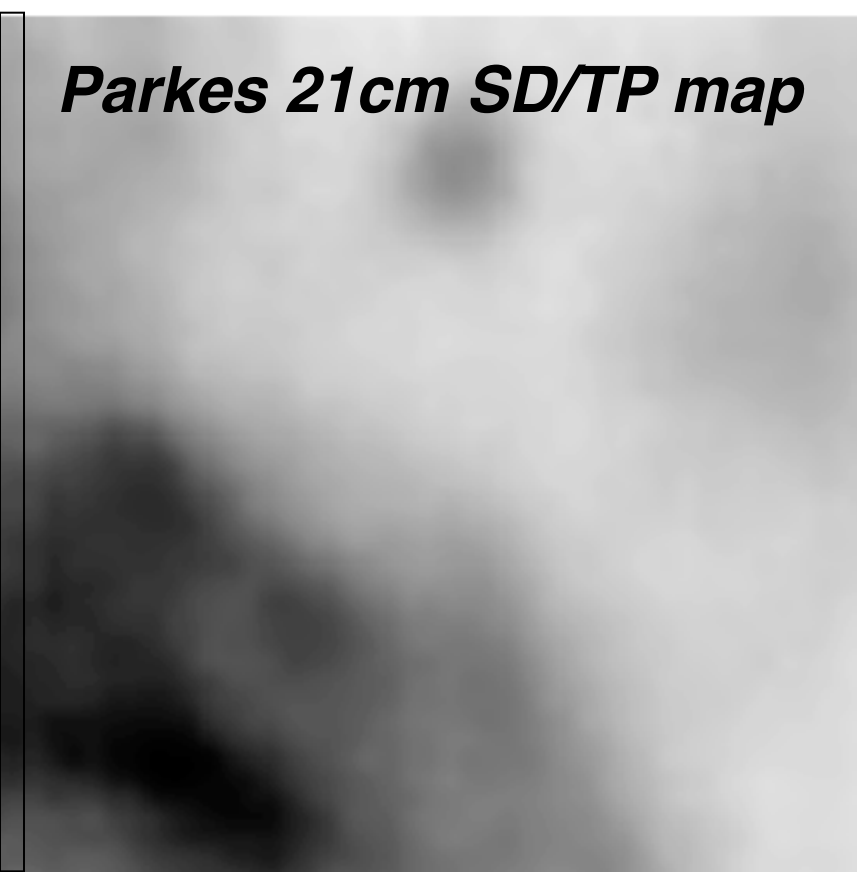

Sum of Re-weighted CLEAN map and SD map: has the correct total flux density, and our best estimate of all spatial frequencies up to the maximum!
Amplitude
UV-distance



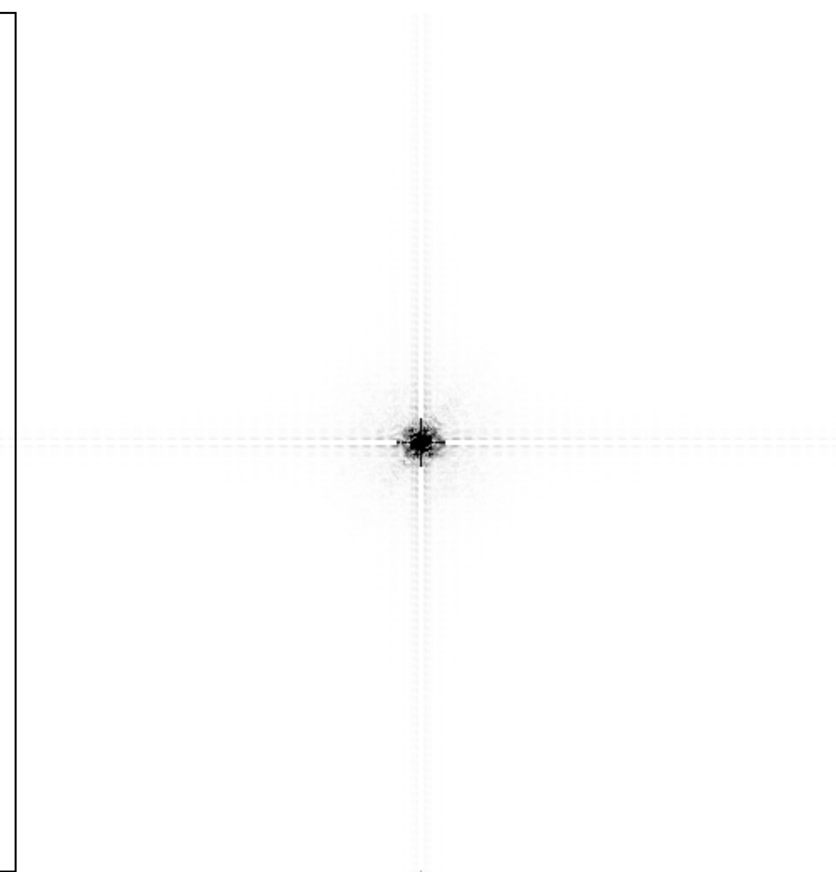
FT →



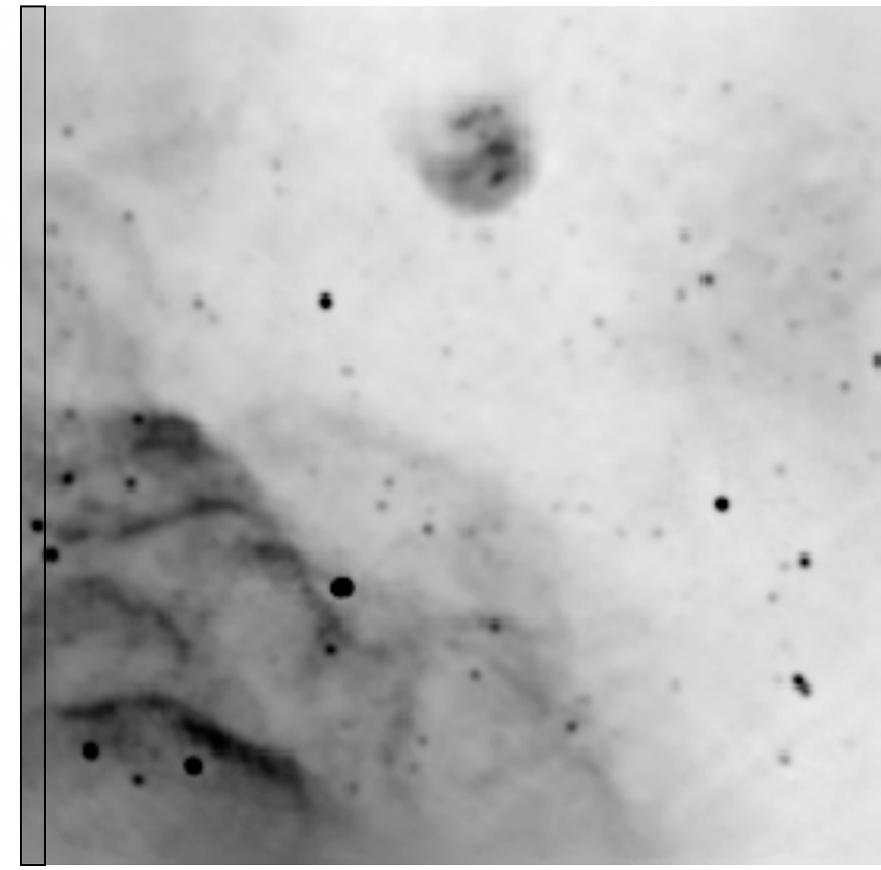
+



FT →

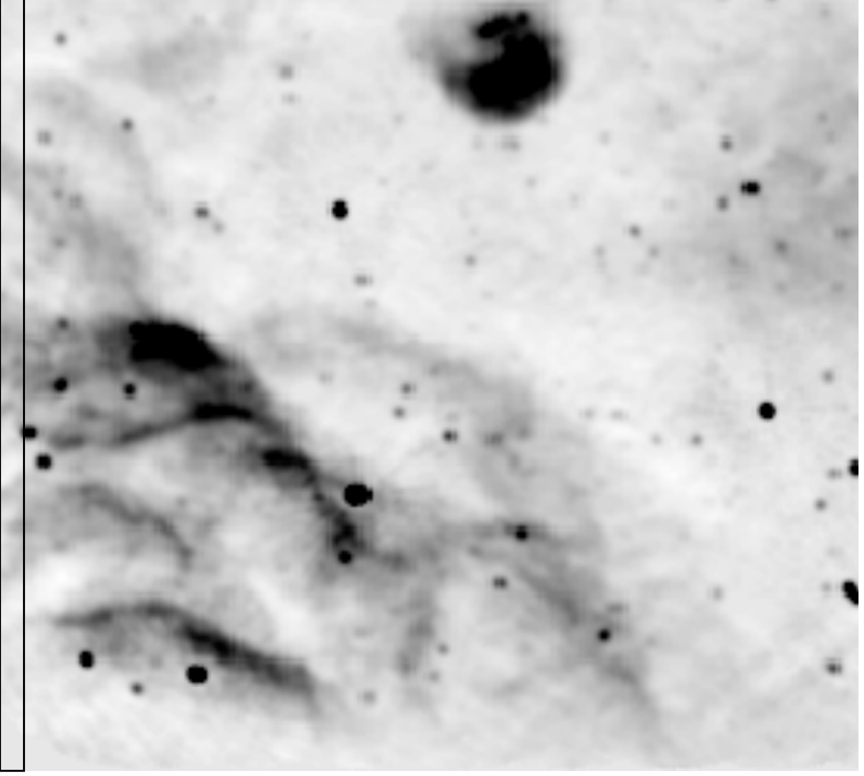


$FT^{-1}=$

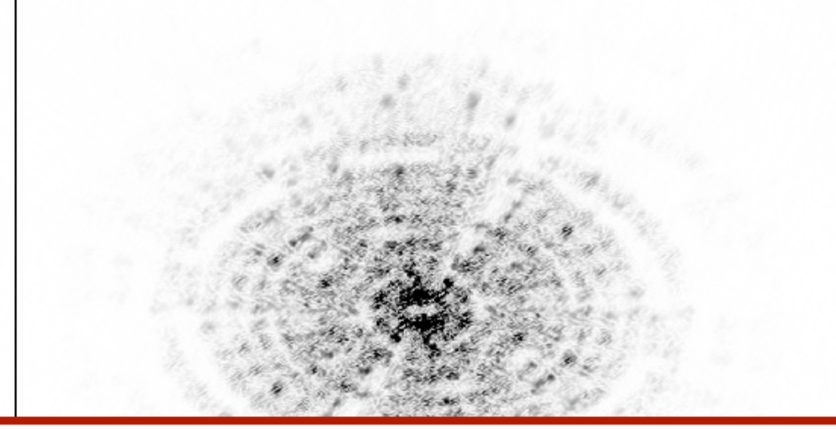


McClure-Griffiths et al.

ATCA 21cm cleaned INT map



FT →

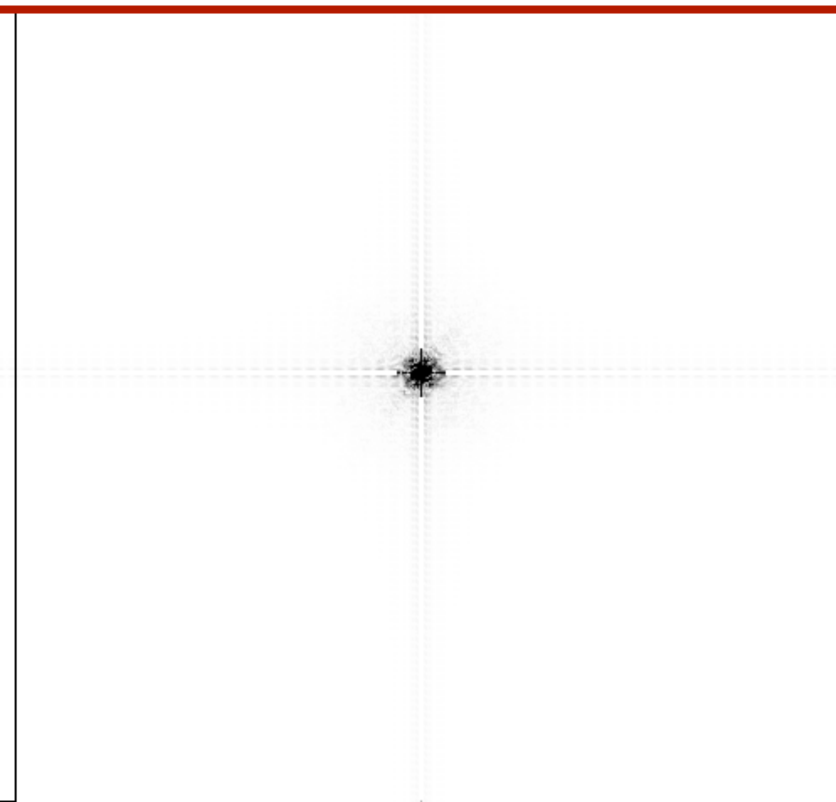


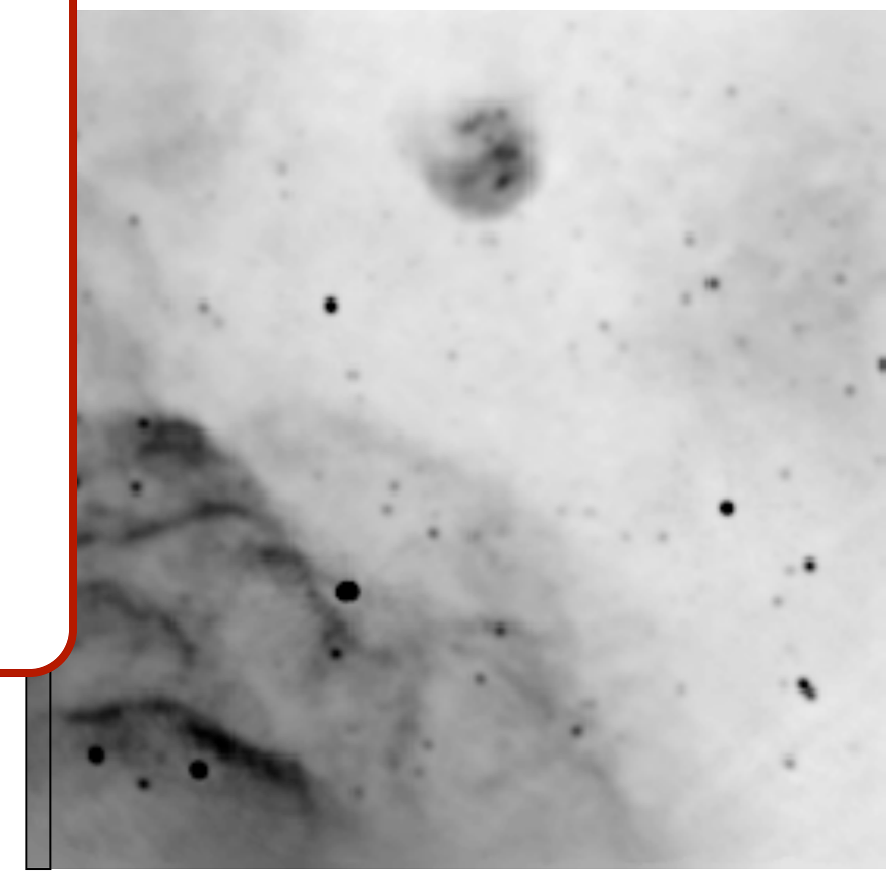
In CASA: Task feather()
*input low-res (SD) image
*high-res image
*SD calibration tweakable

Best to co-register pixels, velocity channels first.

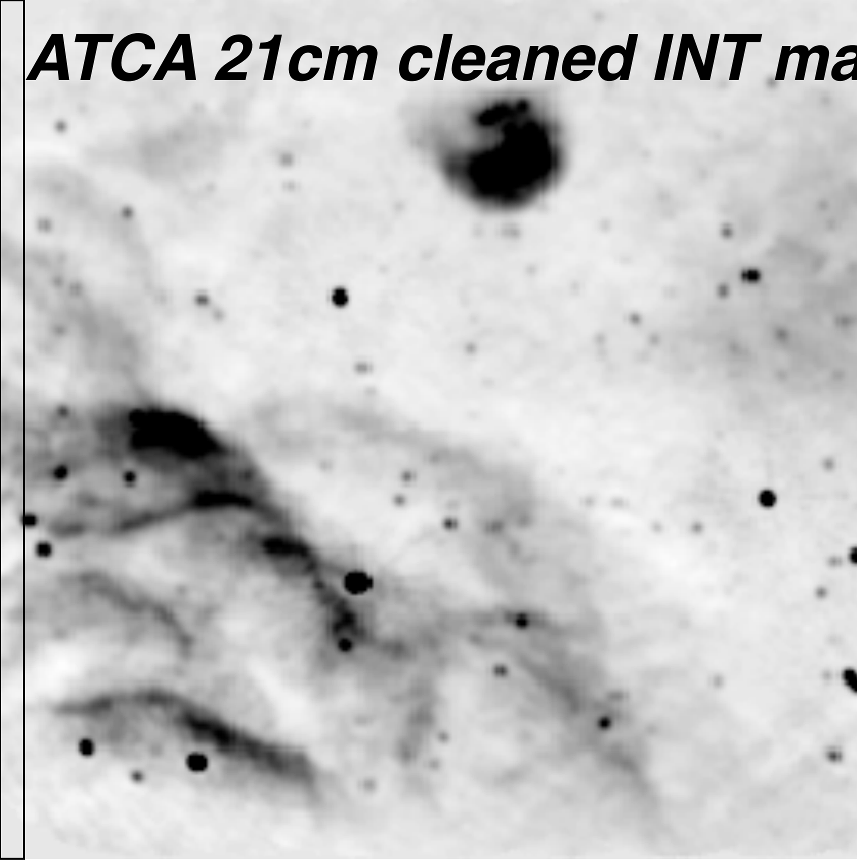
Parkes 21cm SD/TP map

FT →

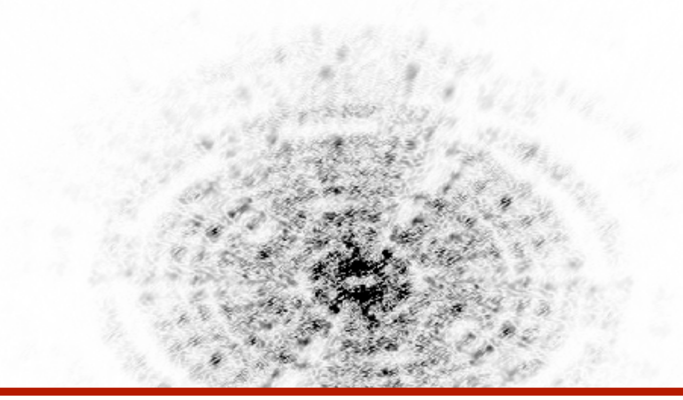




McClure-Griffiths et al.



FT →



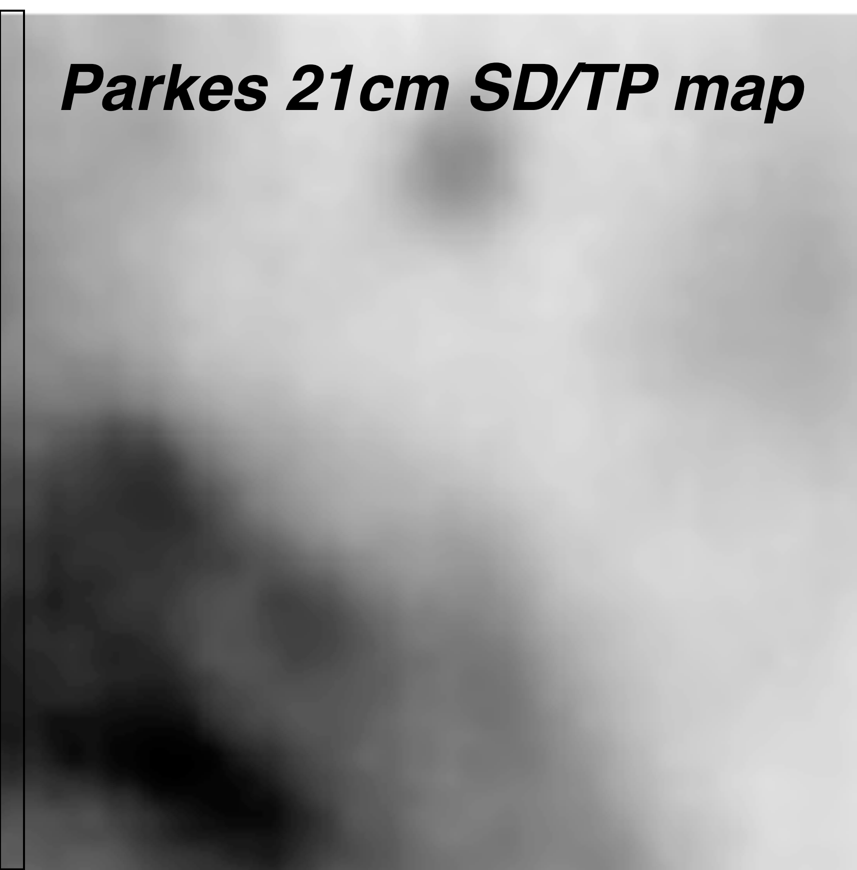
In CASA: Task feather()
*input low-res (SD) image
*high
*SD

Best to
velocity

Feathering is widely used and fairly robust but there are other approaches:

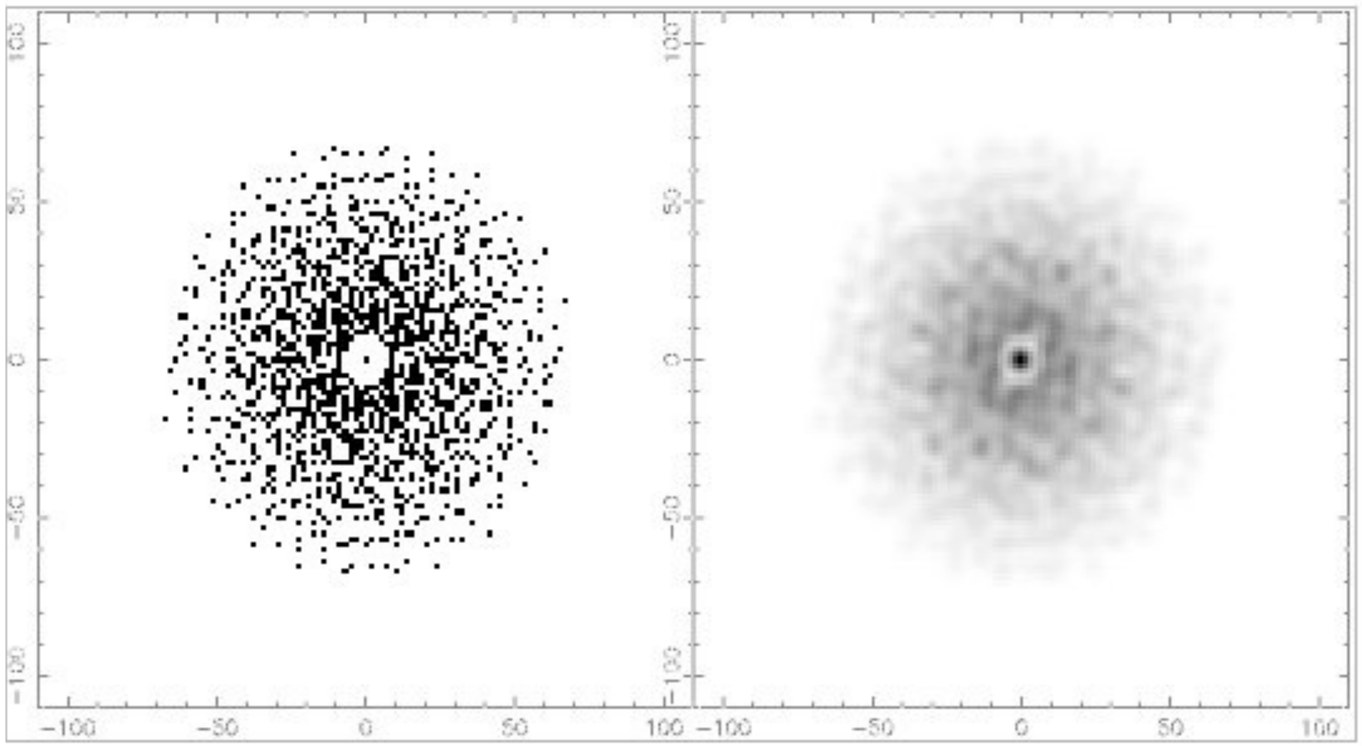
*MEM default image
*Turn SD into pseudo-visibilities, jointly deconvolve together (e.g., Koda et al.2011)

*See S. Stanimirovic article in *Single Dish Summer School* Proceedings



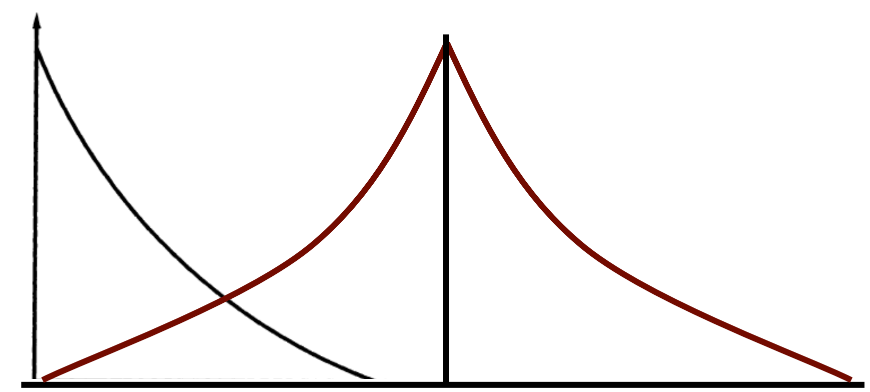
FT →

What Single Dish Data do I Need?



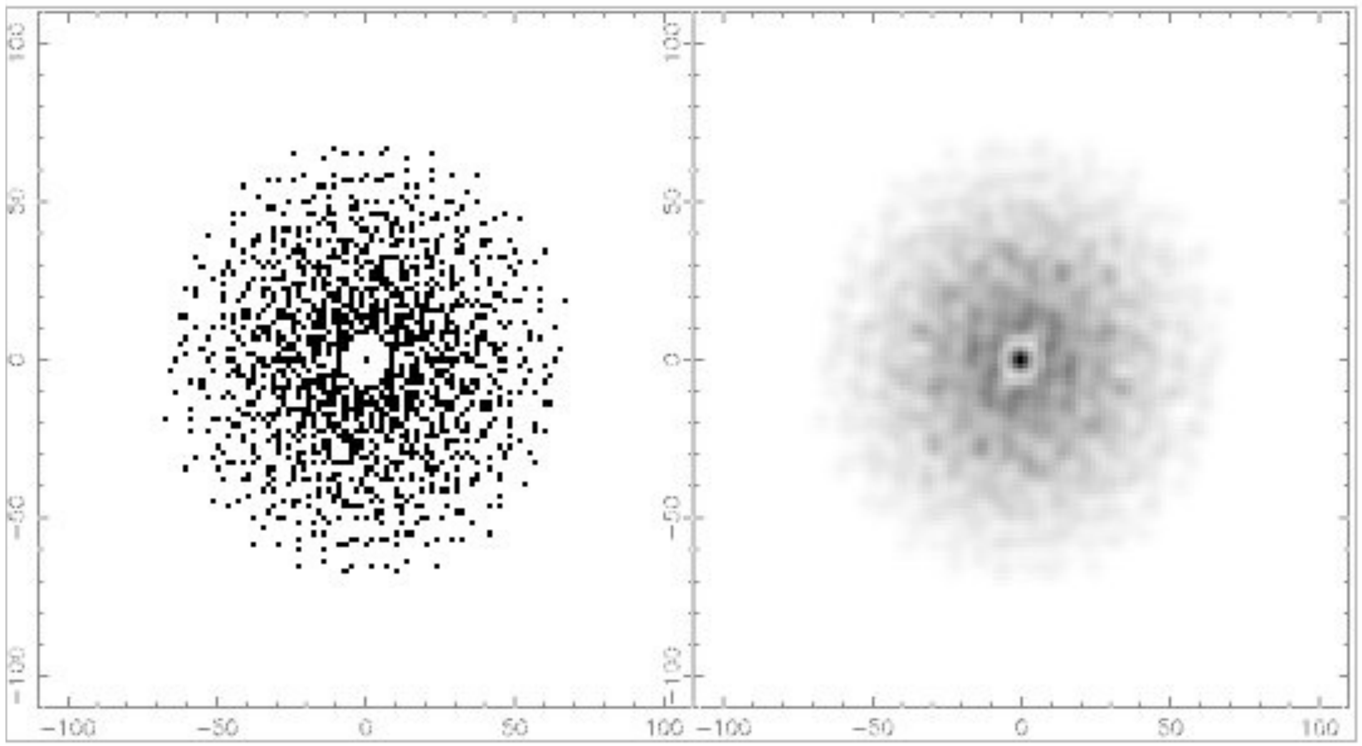
interferometer diameter D

single dish diameter D



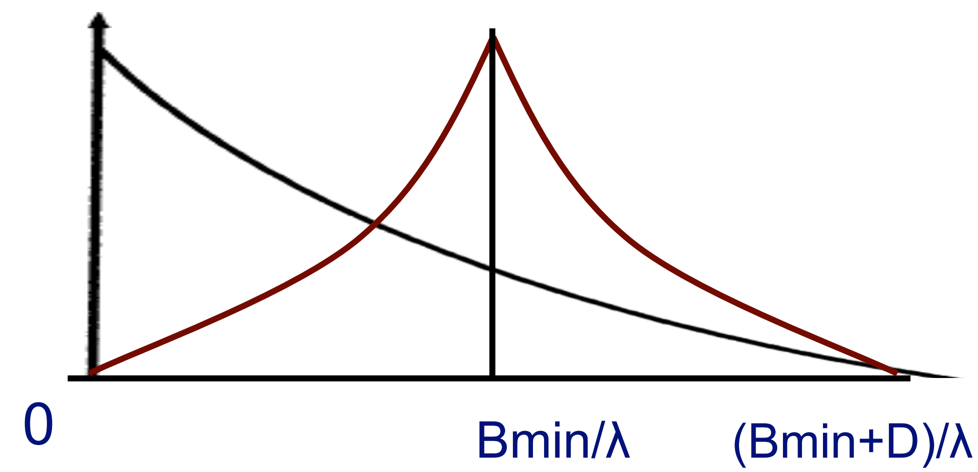
Problems:
****You still have a "hole" between (0,0) and Bmin***
****No common, well-measured spatial freq's***

What Single Dish Data do I Need?



interferometer diameter D

single dish diameter 2D



Problems:
****You still have a "hole" between (0,0) and Bmin***
****No common, well-measured spatial freq's***

To maximize flux recovery and image quality, you want a single dish of D > 1.5xB_{min}

NRAO

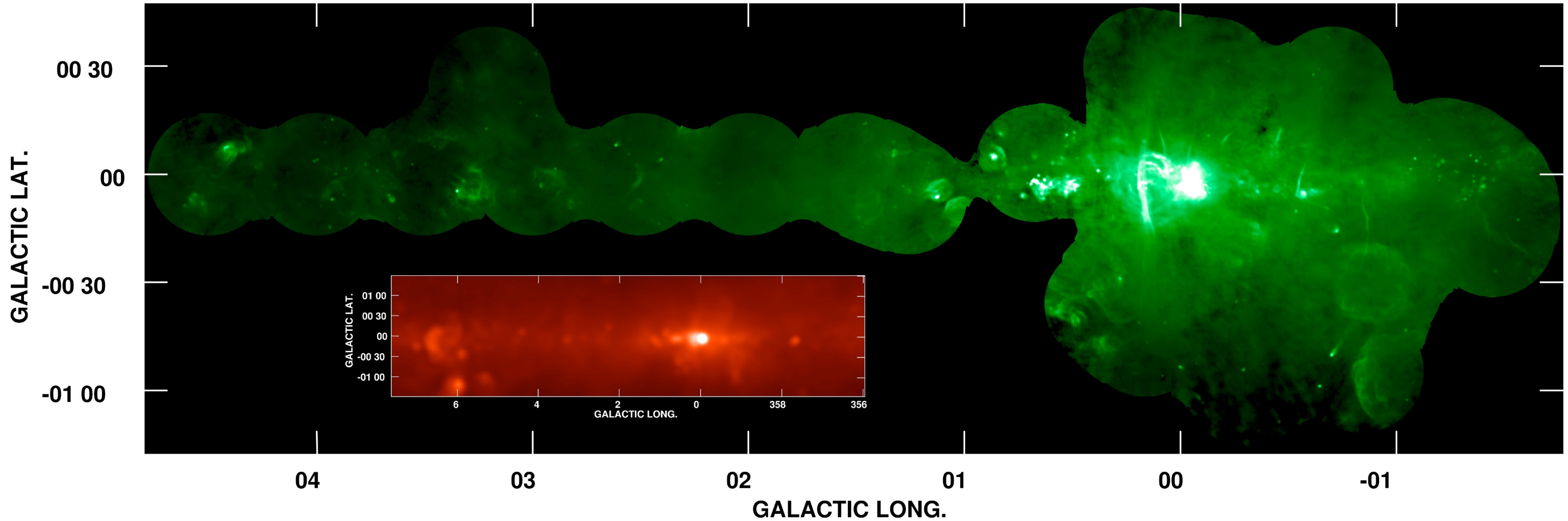
Single Dish Issues

- ***Striping***
  - – Scan rapidly and include signal-free “off” regions (spatial and/or spectral)
    - more of an issue for continuum than spectral line
  - – use appropriate calibration & imaging algorithms.
- ***Relative Calibration***
- ***Sidelobes***
  - – if significant, you may need to deconvolve the single-dish data before combination (e.g., single-dish clean)
  - – at short wavelengths, an “error beam” around the main beam is not uncommon
  - – at long wavelengths, aperture blockage can be an issue (clear aperture is better)
- ***SD Image may not have *all* spatial frequencies*** down to u=v=0 (e.g., millimeter-wavelength continuum)
- ***Pointing errors***
  - – minimize; smooth to mitigate somewhat

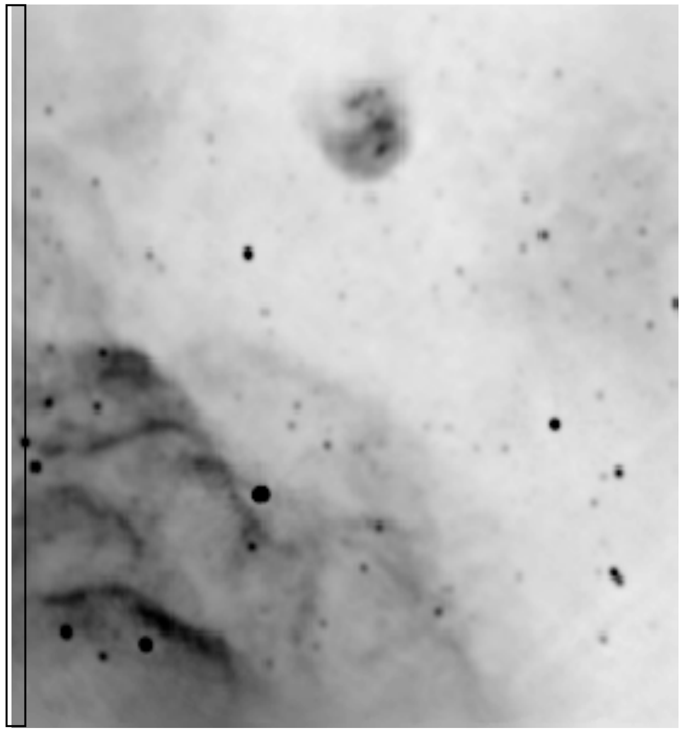
NRAO

Summary

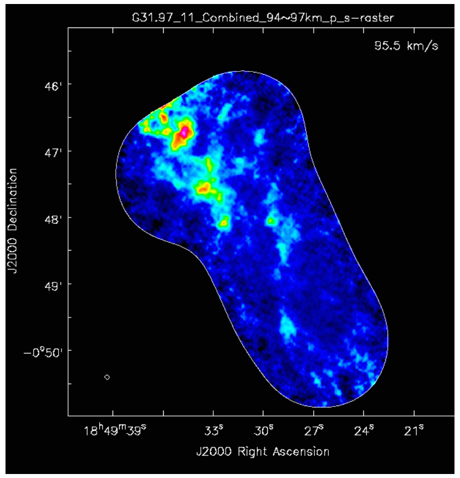
- Each visibility of an interferometer measures a range of spatial frequencies.
- By mosaicking, you can recover some of this information and make gorgeous, scientifically useful images.
  - Adding single dish data can make them even more useful.
- Imaging extended sources accurately can be tricky so get the best data you can, read the literature, experiment, and talk to some people who have done it before.



References & Acknowledgements



- Synthesis Imaging Summer School proceedings
  - mosaicking article by M. Holdaway
  - deconvolution article by T.Cornwell
  - previous lectures by J.Ott, D.Shepherd
- Single Dish Summer School
  - article by S.Stanimirovic
- Theory of Mosaicking: Ekers & Rots (1979)



- Joint Deconvolution: Saul, Stavely-Smith, & Brouw (1996)
- CLEANing: Jorsater & VanMoorsel (1995); Walter et al. (2008); Condon et al. (1998); MS Clean: Cornwell (2008)
- Joint Mosaic UV Gridding: Myers et al. (2003)
- Example of Pseudo-Visibility Joint Deconvolution approach to SD+INT combo: Koda et al. (2011)
- Heterogeneous array / SD relative integration times:
  - Pety-Guth et al. (2008); Kurono et al. (2009); Mason & Brogan (2013)
- Useful discussions with C.Brogan, U.Rao, J.Ott, & others

NRAO

Joint Deconvolution

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

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

- σ_p is the noise variance of an individual pointing; A(**x**) is the primary response function of an antenna (primary beam)
- W(**x**) is an apodization function to suppresses noise amplification at the edge

Sault, Staveley-Smith, Brouw (1996)

Joint Deconvolution

- 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
  – Provides “Ekers & Rots” information: more structure recovered.
  – Overlapping pointings require good knowledge of PB shape further out than the half power point