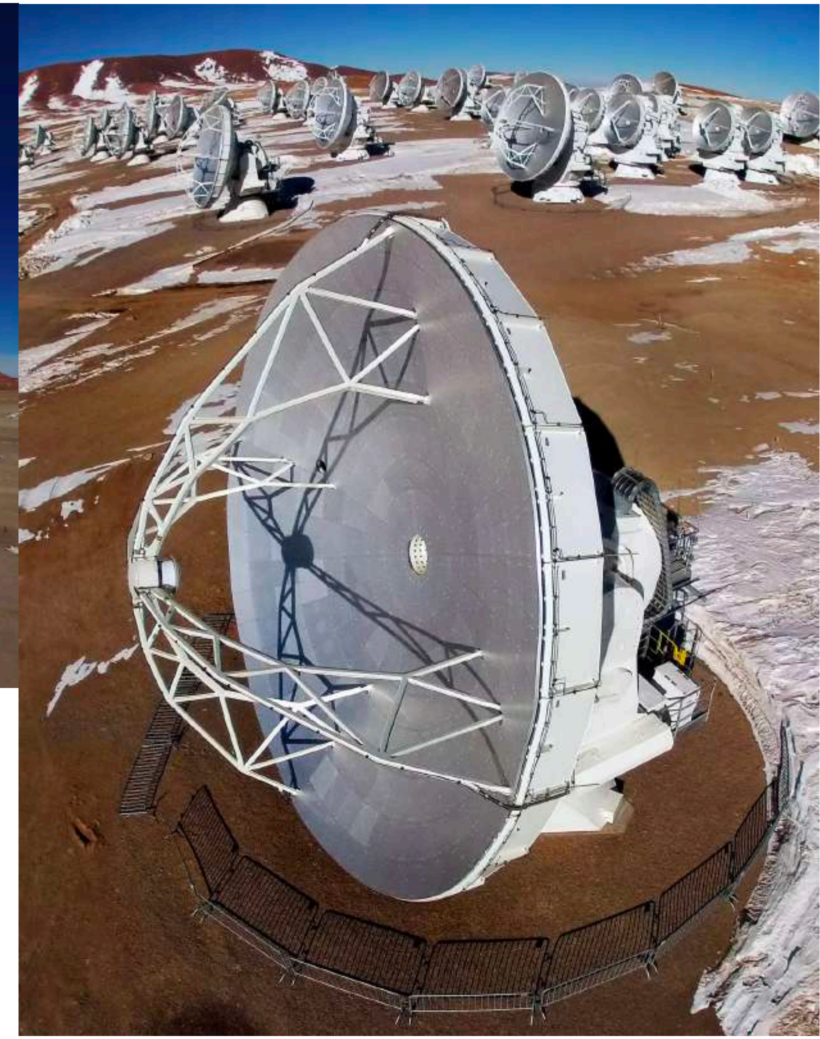
Lecture 8: Radio Astronomy - Deconvolution



Images of ALMA from
https://www.almaobservatory.org/en/images/

Rutgers Physics 346: Observational Astrophysics
March 23, 2021

Homework for Thursday, Mar. 25

Due: Quiz #8 will appear on Canvas Assignments at 4:40pm, due at noon Wednesday.

Due: Worksheet 2 will appear on Canvas tonight, due at 11:59pm Thursday.

Do: Look over options for Project 2 when they appear on Canvas (will broadcast an Announcement). We'll do signups for Project #2 options at the beginning of Thursday's Project Meeting and then start working. Reminder: if you worked on radio imaging (Projects 1-3) you'll now pursue optical spectroscopy (Projects 7-9). If you worked on optical imaging (Projects 4-6), you'll now pursue radio spectroscopy (Projects 10-12).

Quiz #7: Suppose that an array of telescopes has 8 antennas. What is the maximum number of distinct baselines a configuration of this array could deliver?

A: An antenna cannot create a baseline with itself, and each pair should count only once, so the answer is 8x7/2 = 28 baselines. This can also be calculated as combinations of 2 out of 8 = 8!/(6!)(2!) .

Aperture synthesis: summary

- any pair of antennas in an array corresponds to a **baseline** B, defined by a length and an orientation

- when viewed "from the point of the source" at a particular time, this becomes a **projected baseline** B_p

- the x and y components of a projected baseline and the observing wavelength λ determine the dimensionless coordinates $u = B_{p,x}/\lambda$ and $v = B_{p,y}/\lambda$

- a given (u,v) pair defines a two-dimensional **spatial frequency**; the power measured at that (u,v) corresponds to how "ripply" the sky brightness distribution is at a particular direction and periodicity

Aperture synthesis: summary

- **when we have many pairs of antennas (i.e., baselines) and make measurements at many different times, we can measure power at many spatial frequencies**

- **angular resolution ~ $\lambda/B_{p,max}$ – so longer baselines (and shorter wavelengths) give higher resolution!**

- **largest angular scale ~ $\lambda/B_{p,min} \geq \lambda/D$ (for D = diameter of an antenna) – so smaller antennas do a better job of recovering smoothly extended emission!**

- **if the sky brightness distribution includes emission that is larger than λ/D, the interferometer can't see it**

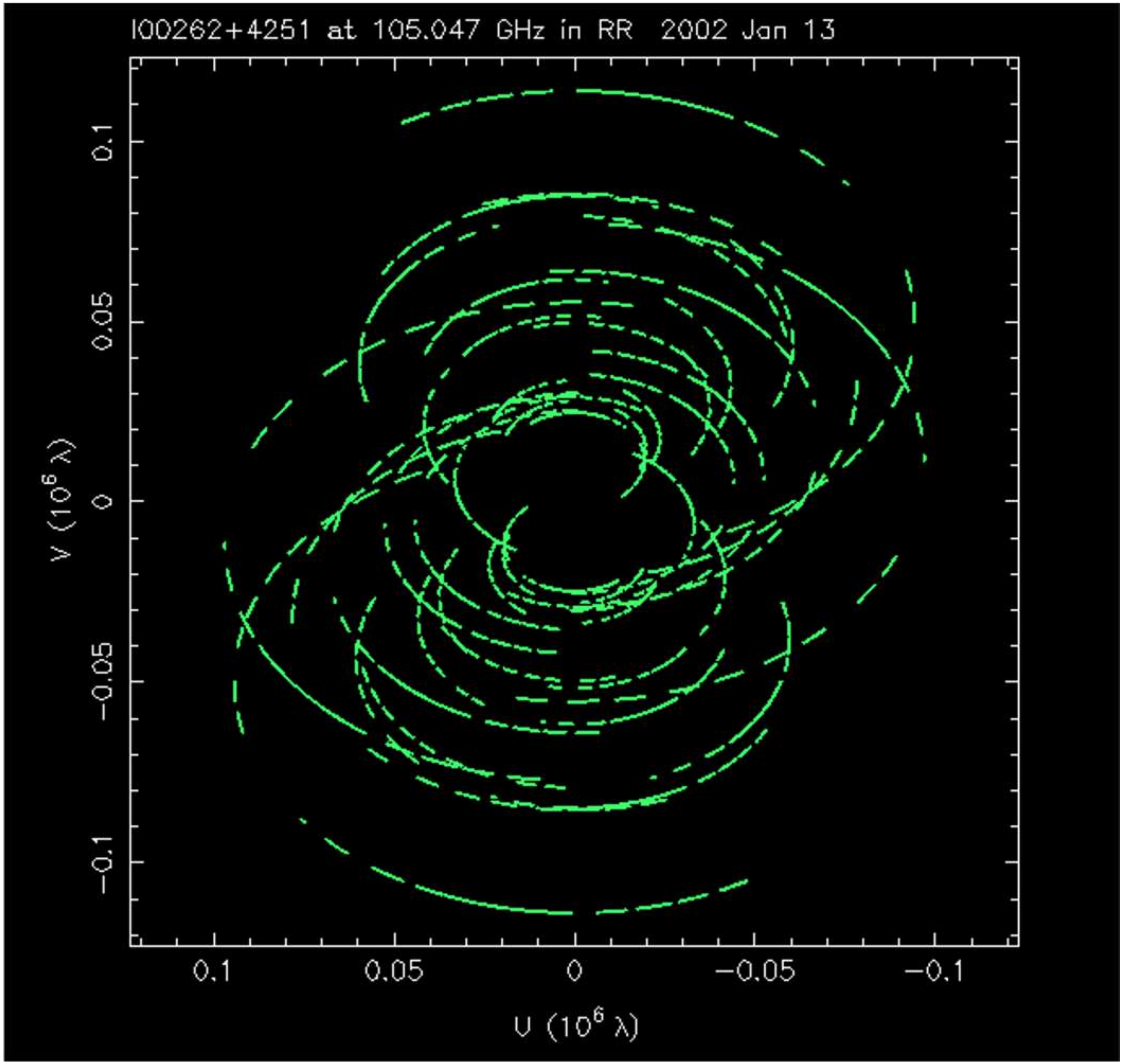
Example of *uv* sampling

Each arc is produced by a different baseline as the Earth rotates under the source.

Longer baselines reach larger (*u*,*v*).

Central gap due to nonzero antenna diameter.

More sampling is always better!



A crucially important Fourier pair

The complex number that is the response of a (projected) baseline with a given (u,v) to a given sky brightness distribution* $J(l,m)$ (where l and m are small angles on the sky in x and y directions) is the visibility $V(u,v)$.

In an ideal world, we'd know $V(u,v)$ at all possible $\{(u,v)\}$, and we could calculate $J(l,m)$ with a simple 2D Fourier transform: (remember $u = B_{p,x}/\lambda$ and $v = B_{p,y}/\lambda$)

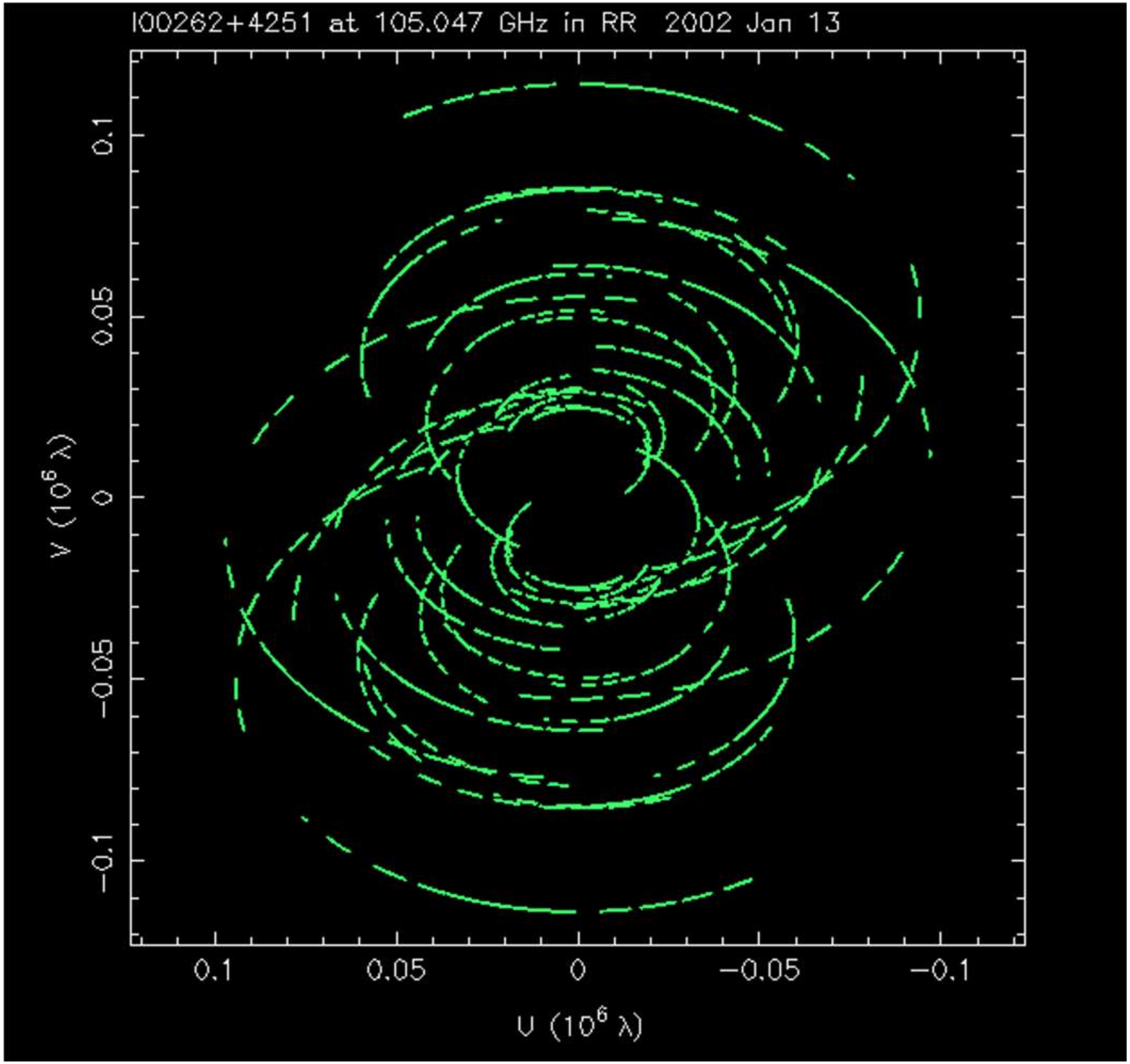
$$J(l,m) = \iint V(u,v)\, e^{ul+vm}\, du\, dv$$

In the real world, however, we don't have measurements of V at all possible $\{(u,v)\}$ – and this is a problem!

***Sky brightness distribution usually written as capital "*I*"...**

Incomplete *uv* sampling

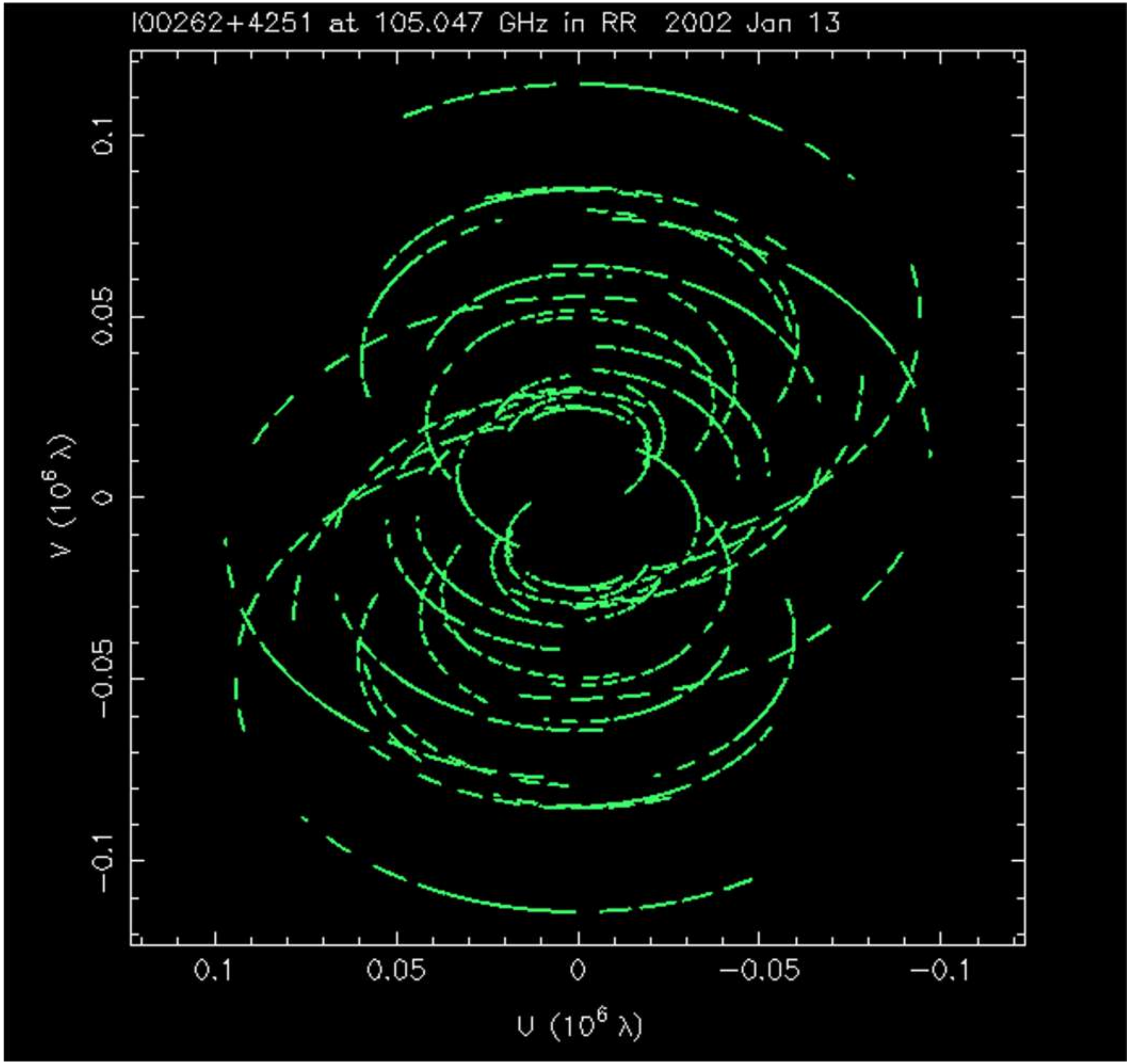
Question: how many possible sky brightness distributions are consistent with a given set of visibilities measured at the {(*u*,*v*)} represented by the green points?



Incomplete *uv* sampling

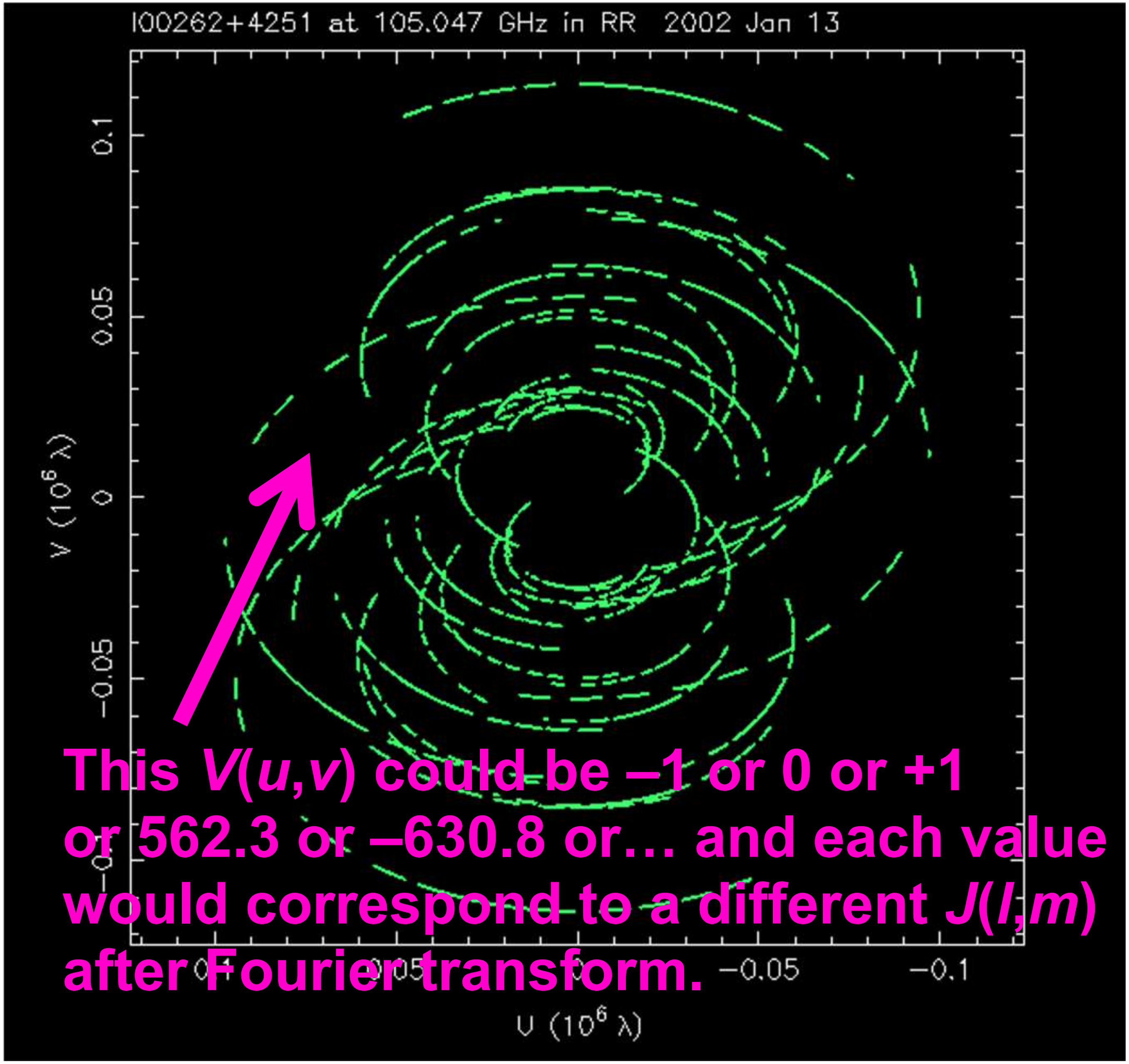
Question: how many possible sky brightness distributions are consistent with a given set of visibilities measured at the {(u,v)} represented by the green points?

Answer: infinitely many!



Incomplete *uv* sampling

Question: how many possible sky brightness distributions are consistent with a given set of visibilities measured at the {(*u*,*v*)} represented by the green points?



This *V*(*u*,*v*) could be –1 or 0 or +1 or 562.3 or –630.8 or… and each value would correspond to a different *J*(*l*,*m*) after Fourier transform.

Answer: infinitely many!

Incomplete *uv* sampling

Instead of being in the ideal world where we know $V(u,v)$ for all $\{(u,v)\}$ and obtain the sky brightness distribution as

$$J(l,m) = \iint V(u,v)\, e^{ul+vm}\, du\, dv$$

what we actually know in the real world is

$$J_D(l,m) = \iint V(u,v)\, S(u,v)\, e^{ul+vm}\, du\, dv$$

where $S(u,v)$ is the "sampling function" and has a value of

1 at all $\{(u,v)\}$ where we have data

0 at all $\{(u,v)\}$ where we do not have data

$J_D(l,m)$ is referred to as the "dirty image" or "dirty map."

Incomplete *uv* sampling

Let's look a bit more closely at this equation:

$$J_D(l,m) = \iint V(u,v)\, S(u,v)\, e^{ul+vm}\, du\, dv$$ **the "dirty map"**

This is effectively a convolution:

$$J_D(l,m) = J(l,m) * B(l,m)$$

in terms of

$$J(l,m) = \iint V(u,v)\, e^{ul+vm}\, du\, dv$$ **sky brightness distribution**

$$B(l,m) = \iint S(u,v)\, e^{ul+vm}\, du\, dv$$ **the "dirty beam"**

Incomplete *uv* sampling

Let's look a bit more closely at this equation:

$$J_D(l,m) = \iint V(u,v)\, S(u,v)\, e^{ul+vm}\, du\, dv$$ **the "dirty map"**

This is effectively a convolution:

$$J_D(l,m) = J(l,m) * B(l,m)$$

we can calculate this

we want this

in terms of

we have measured this

$$J(l,m) = \iint V(u,v)\, e^{ul+vm}\, du\, dv$$ **sky brightness distribution**

$$B(l,m) = \iint S(u,v)\, e^{ul+vm}\, du\, dv$$ **the "dirty beam"**

Convolution and deconvolution

Recall from before: convolution is "smearing" two functions together.

Deconvolution is the process of "unsmearing" the dirty beam from the dirty map in order to estimate the sky brightness distribution.

we can calculate this

$$J_D(l,m) = J(l,m) * B(l,m)$$

we want this

we have measured this

The challenge: since there are infinitely many possible sky brightness distributions consistent with measured visibility data and the dirty map $J_D(l,m)$, we have to assume a (Bayesian) prior to guide deconvolution.

image credit: D. Gary

Illustration of dirty map and beam

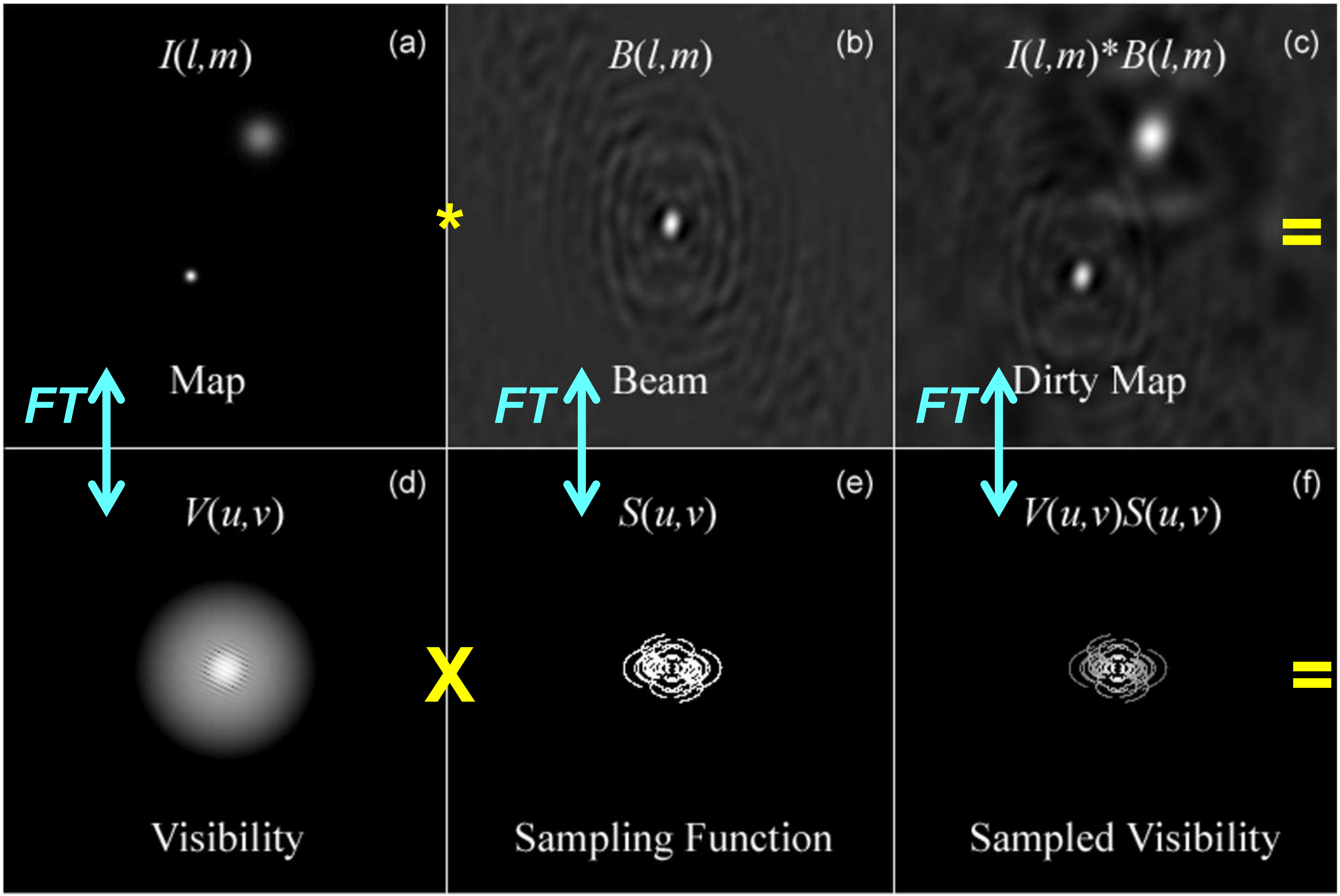


image credit: D. Gary

Illustration of dirty map and beam

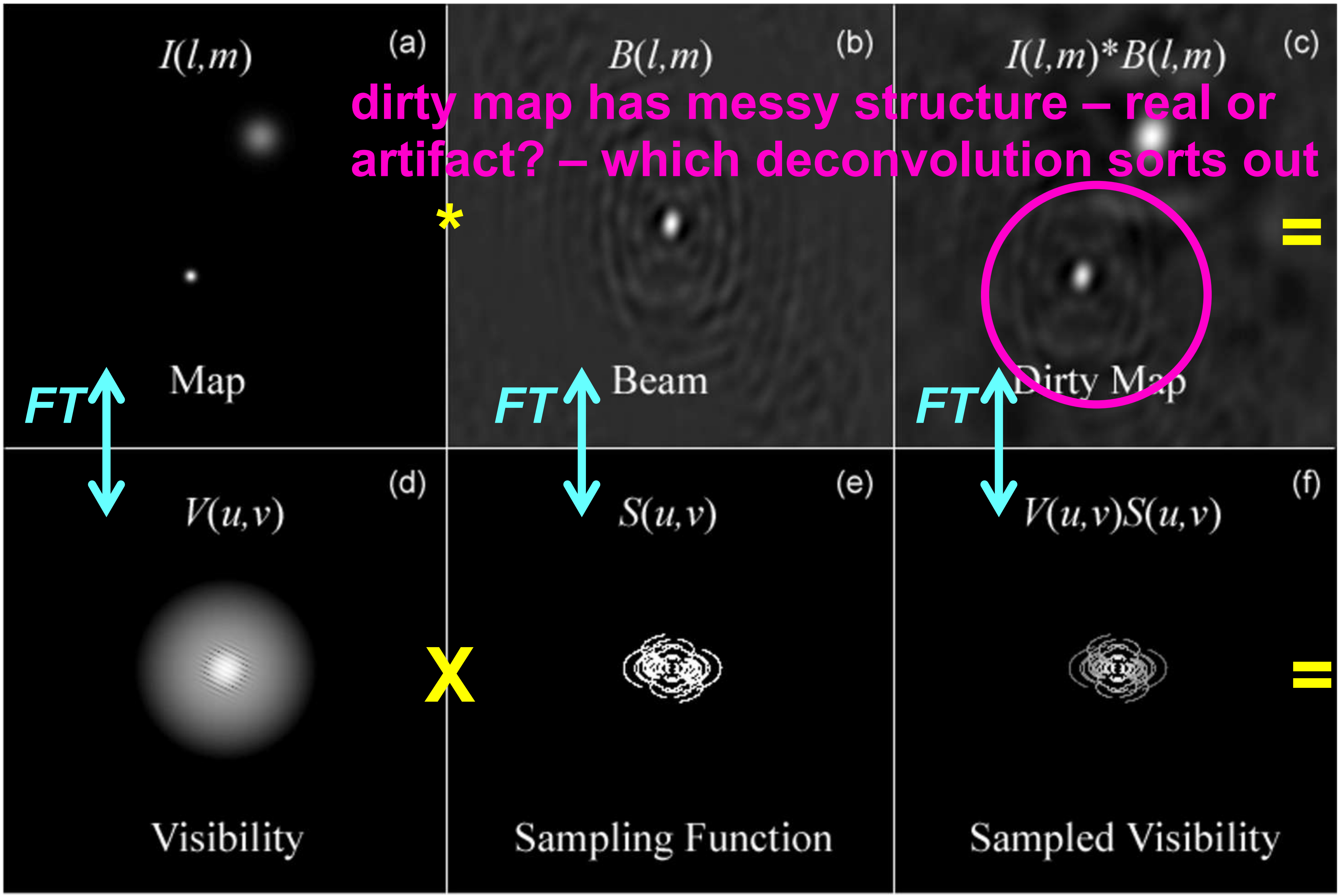
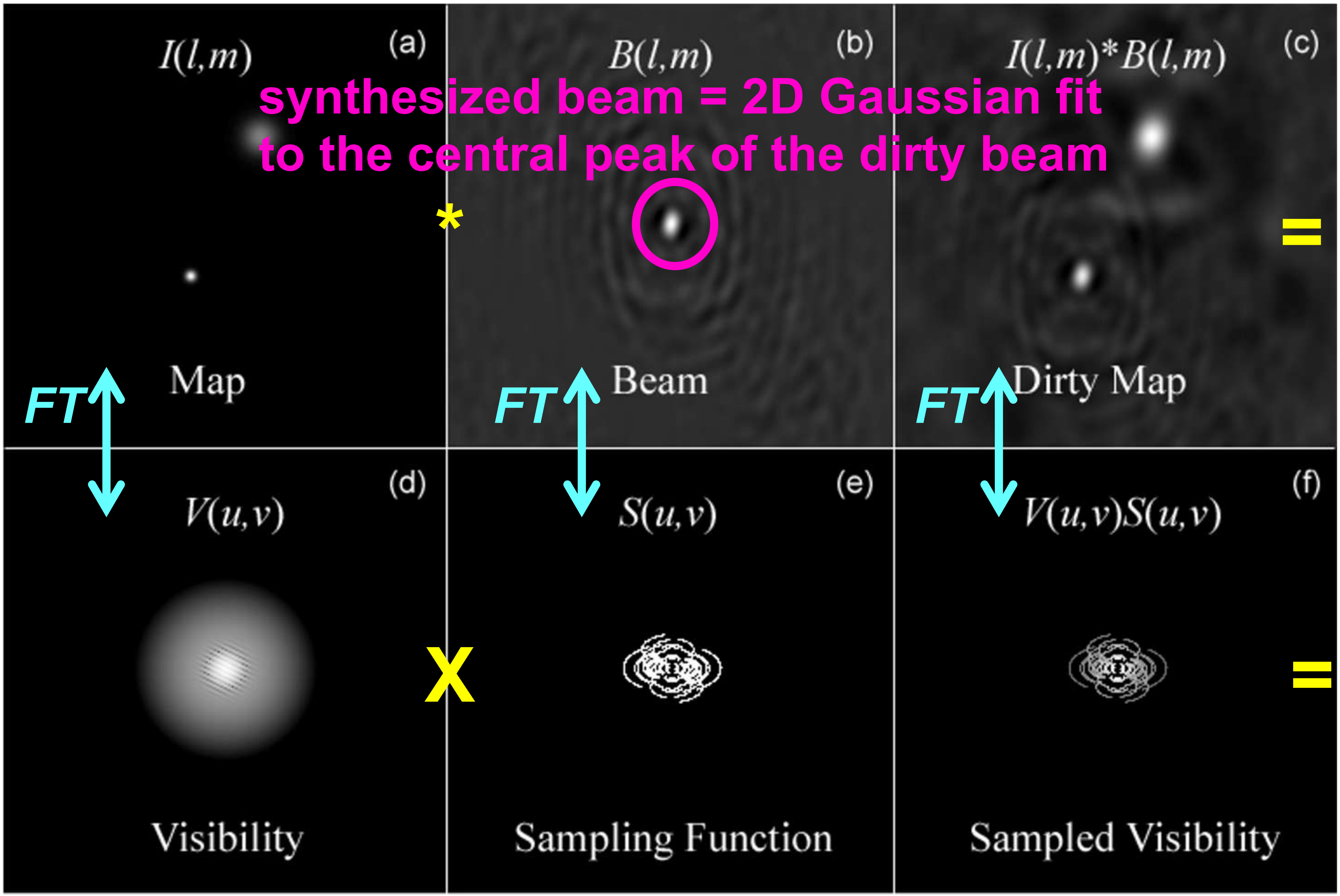


image credit: D. Gary

Illustration of dirty map and beam



Convolution and deconvolution

Recall from before: convolution is "smearing" two functions together.

Deconvolution is the process of "unsmearing" the dirty beam from the dirty map in order to estimate the sky brightness distribution.

we can calculate this

$$J_D(l,m) = J(l,m) * B(l,m)$$

we want this

we have measured this

The challenge: since there are infinitely many possible sky brightness distributions consistent with measured visibility data and the dirty map $J_D(l,m)$, we have to assume a (Bayesian) prior to guide deconvolution.

Possible prior # 1: smoothness

One possible prior is that the sky brightness distribution is maximally smooth (no sharp edges, etc.).

This prior guides the maximum entropy method (MEM) algorithm for deconvolution:

+ leads to reconstructed images that are smooth… good for imaging structures that extend over wide areas on the sky, but not so good for compact structures
+ algorithm also tends to run away and keep adding more and more emission to reconstructed image, unless a constraint on total flux is available

Possible prior # 2: point sources

Another possible prior is that the sky brightness distribution is well described as a sum of point sources.

This prior guides the CLEAN algorithm for deconvolution:

(1) identify the peak pixel remaining in the dirty map
(2) take the 2D dirty beam, shift it to be centered on the peak pixel, and rescale it to 5–10% of the peak pixel height
(3) subtract the shifted/scaled dirty beam from the dirty map, and record the corresponding delta function
(4) are the residuals only noise? if NO: go to step (1) and repeat; if YES: go to step (5)

(5) define "clean image" as sum of all delta functions (convolved with synthesized beam) and final residuals

More about CLEAN

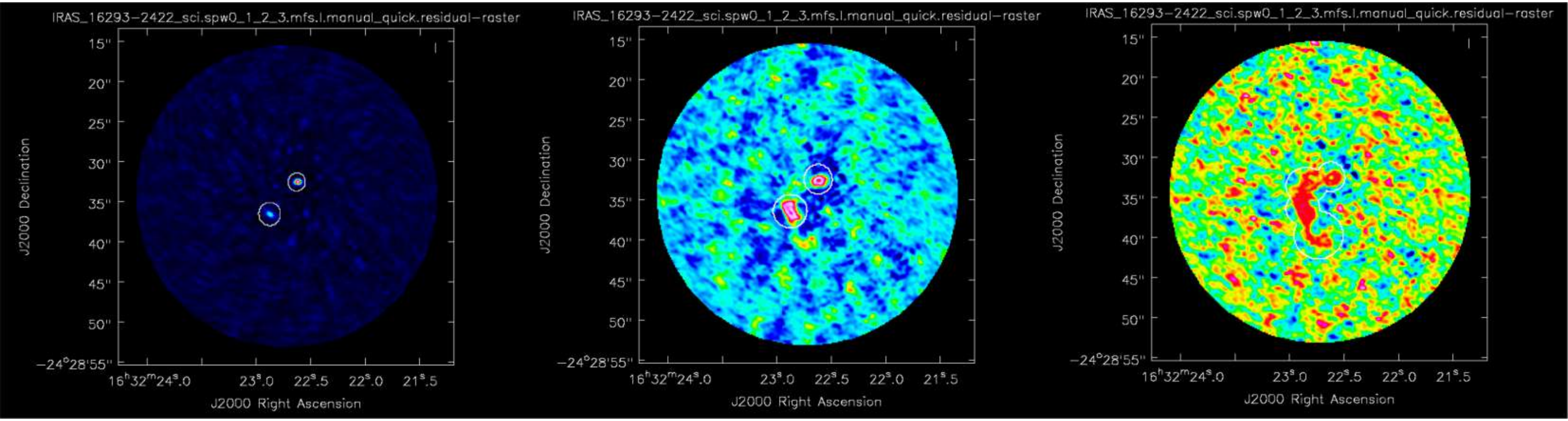
The CLEAN algorithm works quite well for sky brightness distributions that are well described in terms of point sources –but not so well for smooth, extended emission. (Fancier algorithms like "multi-scale CLEAN" can help if both point source and extended emission are present.)

One big question: when do we stop repeating steps (1) through (4)?

+ It is possible to underclean – then there will be artifacts of the dirty beam left in your "clean" map.
+ It is possible to overclean – then some of the "clean components" included in your clean map will actually be noise peaks.
+ This question and other details of CLEANing lead to experimentation and arguments among astronomers...

image credit: ALMA

CLEAN in progress



Left: initial dirty map
Middle: residuals after two clean cycles
+ emission peaks now lower relative to noise
Right: residuals after many clean cycles
+ low-level extended emission is now visible,
motivating use of larger “clean boxes” in which
to subtract clean components

Visibility weighting

The 2D Fourier transform of the visibility data is most easily calculated on a regular grid of points – so one of the key steps in processing aperture synthesis data is regridding.

A question now arises: how do we **weight** the individual visibility measurements when regridding the data?

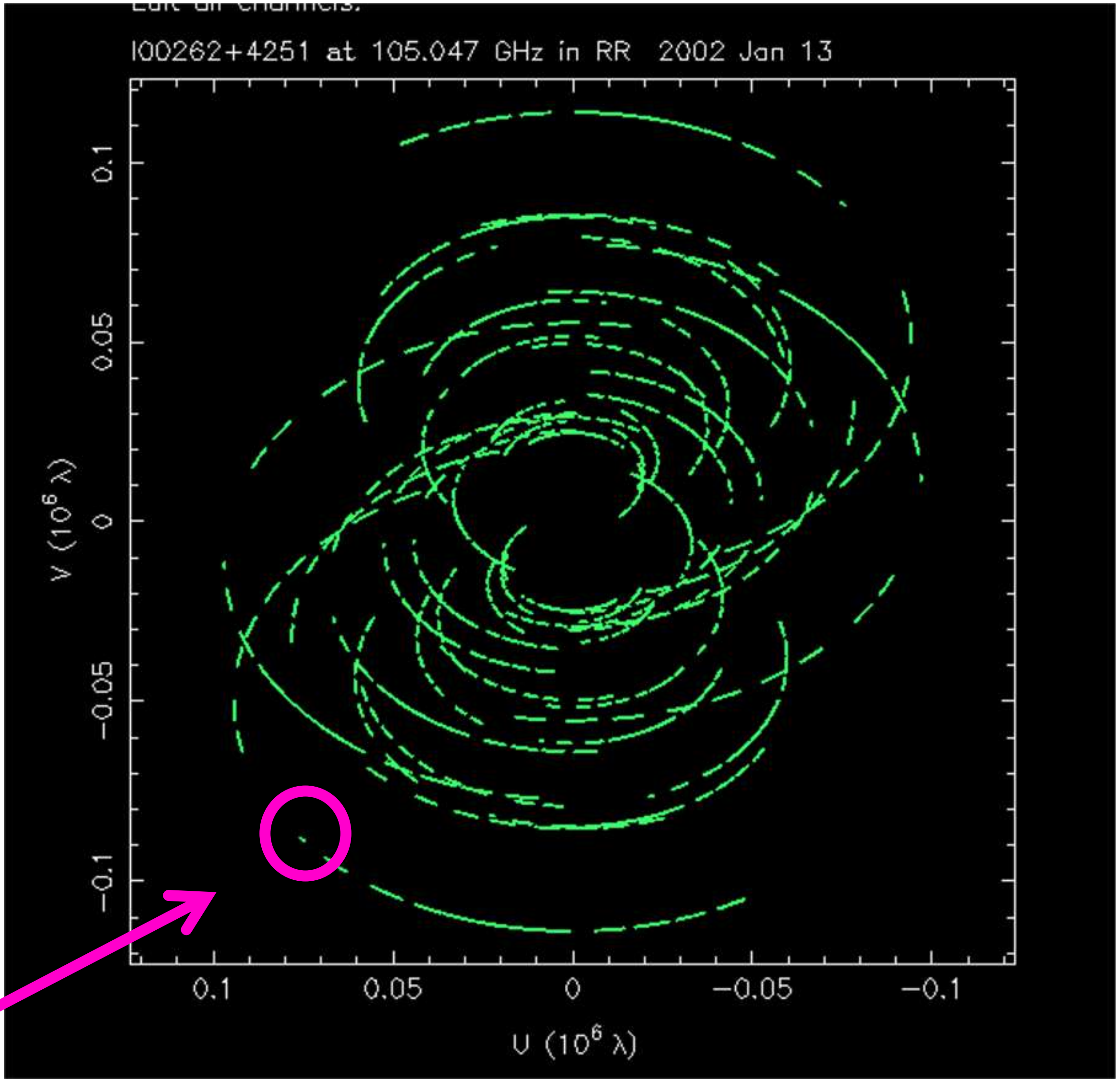
(1) better quality data should get more weight – so we almost always weight visibilities as $(T_{sys})^{-2}$ where system temperature is a measure of noise

(2) we can also think about whether we want to weight visibilities based on their **density**, i.e., how many other visibilities are measured in the same part of the *uv* plane

Visibility weighting

"Natural" weighting:
every visibility gets same density weight
+ optimizes S/N, limits resolution

"Uniform" weighting:
each visibility is weighted inversely proportional to local density
+ optimizes resolution, limits S/N

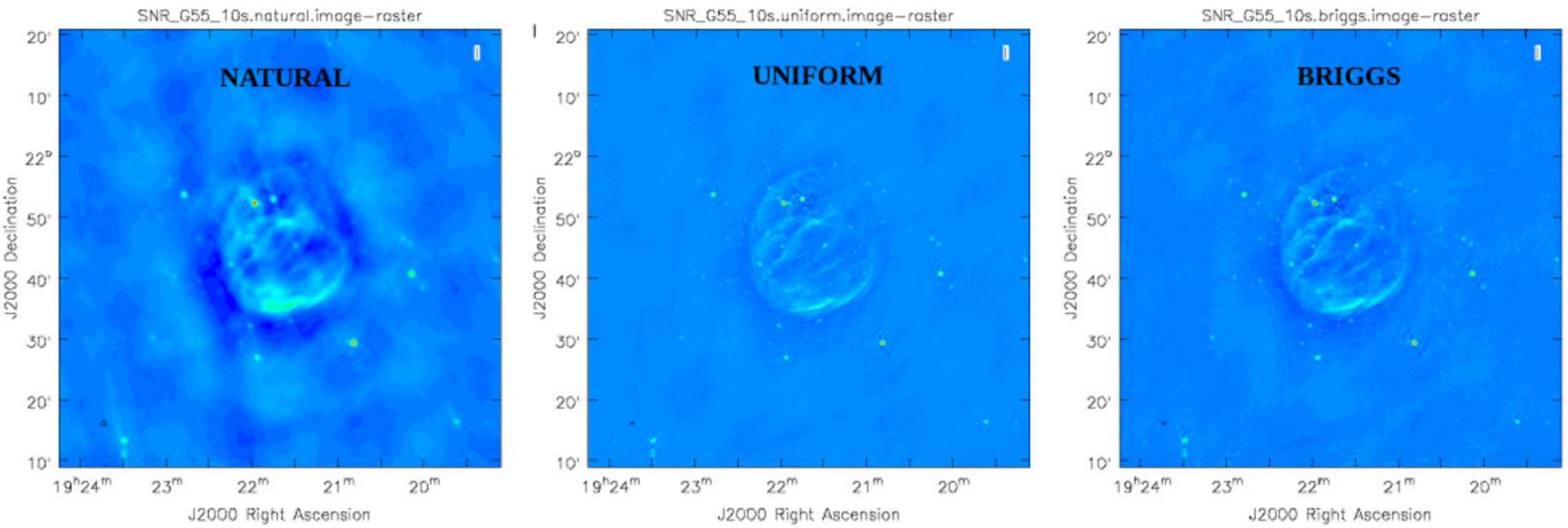


would get high weight for uniform weighting – and more emphasis on long baselines means higher resolution

image credit: NRAO

Visibility weighting

The possibility of different weighting schemes means that we can use the same visibility data to produce (and then deconvolve) multiple dirty maps!



"Briggs" weighting is intermediate between natural and uniform (named after the late Dan Briggs).