References:
1. Synthesis imaging in radio astronomy II (Chp. 2, 7)
2. Interferometry and synthesis in radio astronomy (Chp. 10)
3. Low Frequency Radio Astronomy and the LOFAR observatory (Chp. 8)
4. My lecture 10 and 11 in Astro. Tech II course:
http://www.ncra.tifr.res.in/~ruta/files/2022/2022-lecture-11-ruta.pdf
http://www.ncra.tifr.res.in/~ruta/files/2022/2022-lecture-10-ruta.pdf

Visibility gridding and imaging

Ruta Kale

Radio Astronomy School 2024, 19 November 2024

Outline

What is “imaging” ?

Why do we need to grid visibilities ?

How do we grid the visibilities ?

What are the implications of these steps on the image ?

Visibilities

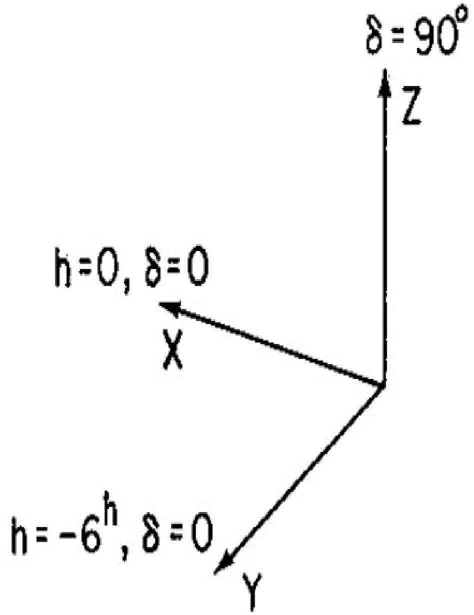
Output of the correlator !

A complex number : amplitude and phase

Recorded for *every* baseline, polarization, frequency channel and sampling time

Each such entry, *a visibility*, makes up *a* point in the 'uv'-plane (sampling plane)

Visibility is a complex quantity, the magnitude of which has the dimensions of spectral power flux density ($Wm^{-2}Hz^{-1}$).



$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \frac{1}{\lambda} \begin{pmatrix} \sin H_0 & \cos H_0 & 0 \\ -\sin\delta_0 \cos H_0 & \sin\delta_0 \sin H_0 & \cos\delta_0 \\ \cos\delta_0 \cos H_0 & -\cos\delta_0 \sin H_0 & \sin\delta_0 \end{pmatrix} \begin{pmatrix} L_X \\ L_Y \\ L_Z \end{pmatrix}$$

The “uv-plane”

Figure 2–11. Coordinate system for specification of baseline parameters. X is the direction of the meridian at the celestial equator, Y is toward the East, and Z toward the North celestial pole.

Projected baseline vectors in the uv-plane trace ellipses

$$u^2 + \left(\frac{v - (L_Z/\lambda)\cos\delta_0}{\sin\delta_0} \right)^2 = \frac{L_X^2 + L_Y^2}{\lambda^2}$$

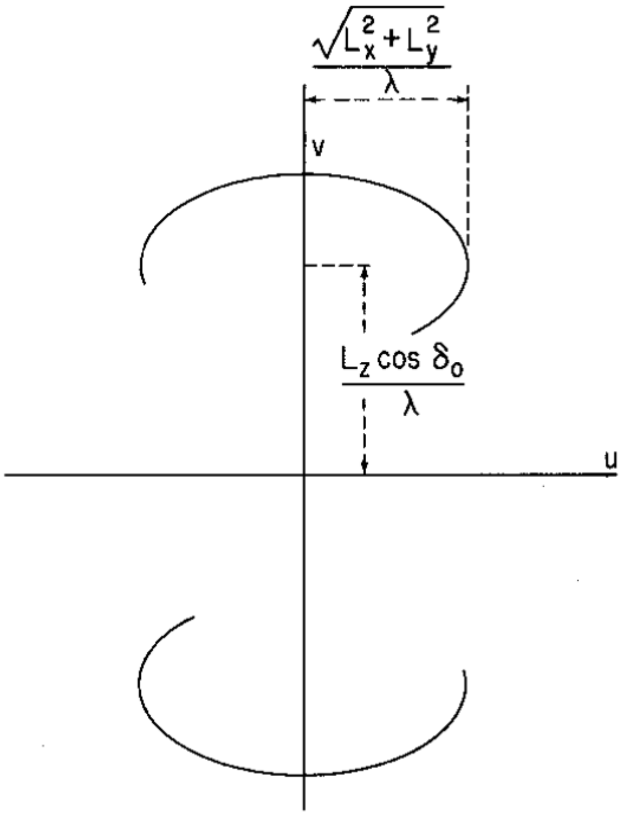
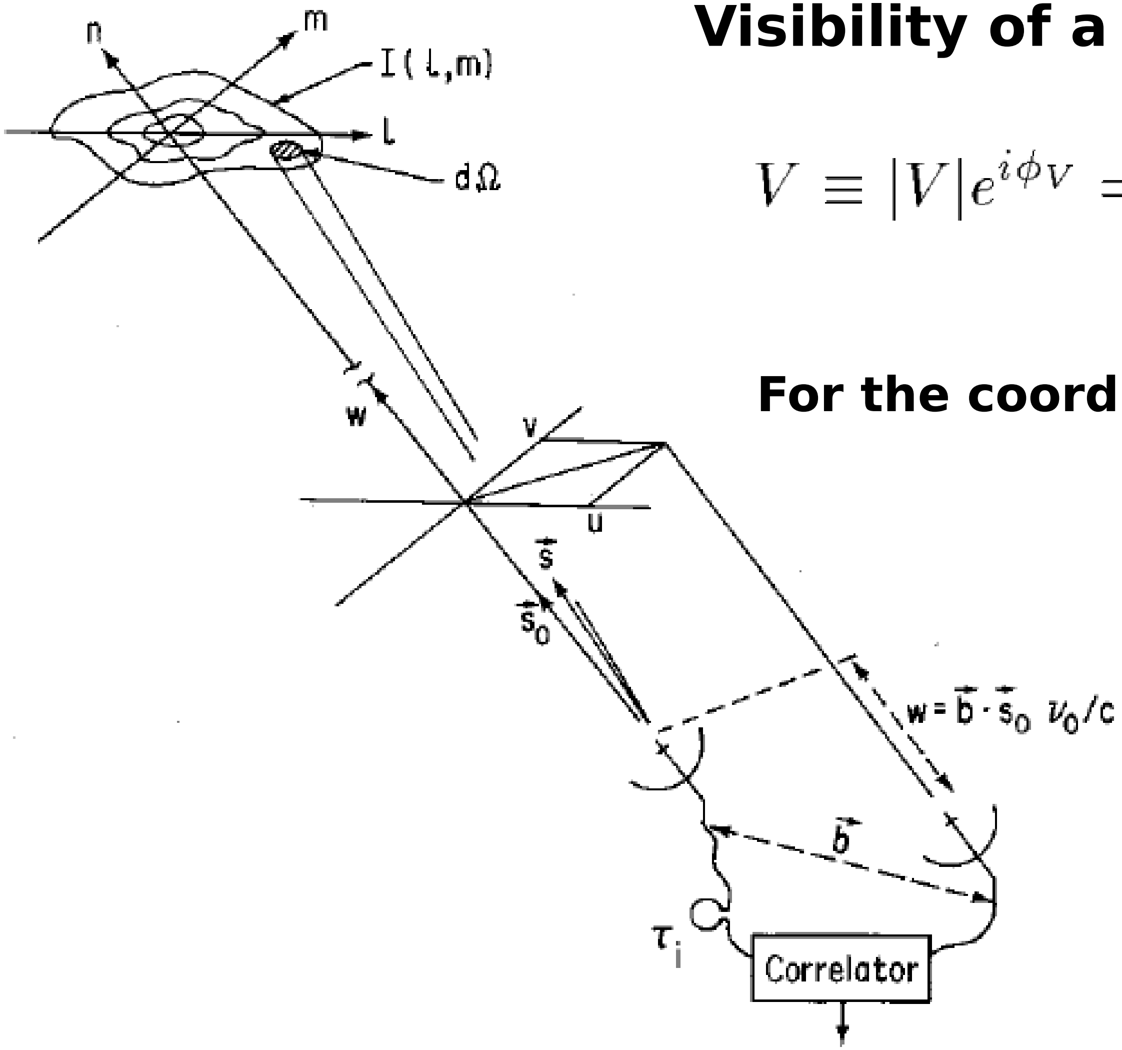


Figure 2–12. Elliptical loci representing the projection of the baseline vector onto the (u, v) plane as a source is tracked across the sky. The lower curve corresponds to a reversal of the direction of the baseline vector, and represents the points for which the visibility is the complex conjugate of that measured on the upper curve. The axial ratio of the ellipses is equal to $\sin\delta_0$. For an East–West baseline $L_Z = 0$, and a single ellipse is centered on the (u, v) origin.

Visibility of a source:



$$V \equiv |V| e^{i\phi_V} = \int_S \mathcal{A}(\boldsymbol{\sigma}) I(\boldsymbol{\sigma}) e^{-2\pi i \nu \mathbf{b}\cdot\boldsymbol{\sigma}/c} \, d\Omega$$

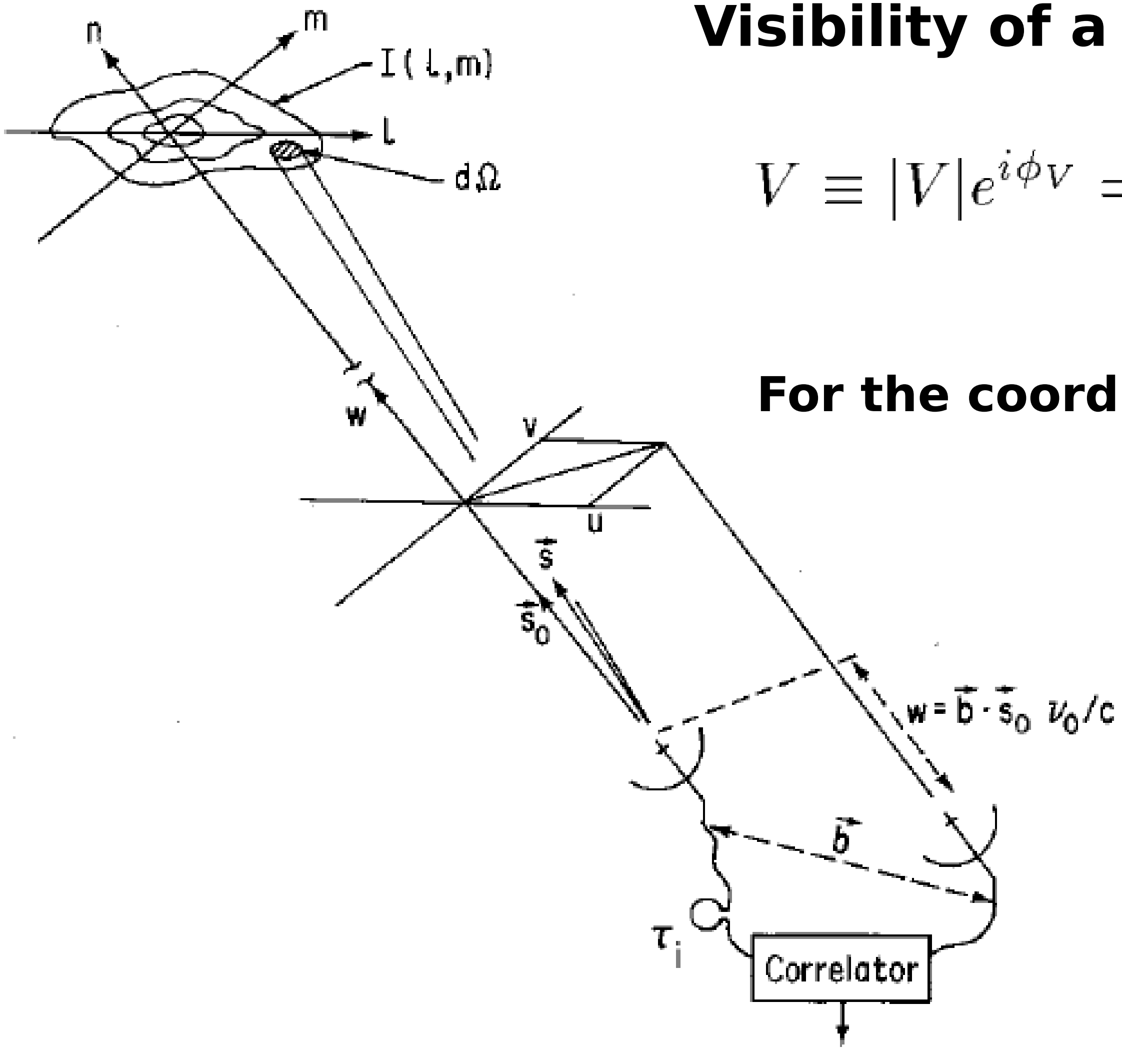
For the coordinate system in the figure:

$$\frac{\nu\, \mathbf{b}\cdot\mathbf{s}}{c} = ul + vm + wn$$

$$\frac{\nu\, \mathbf{b}\cdot\mathbf{s}_0}{c} = w,$$

$$d\Omega = \frac{dl\, dm}{n} = \frac{dl\, dm}{\sqrt{1-l^2-m^2}}$$

$$V(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{A}(l,m) I(l,m) e^{-2\pi i\left[ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right]} \frac{dl\, dm}{\sqrt{1-l^2-m^2}}$$



Visibility of a source:

$$V \equiv |V| e^{i\phi_V} = \int_S \mathcal{A}(\boldsymbol{\sigma}) I(\boldsymbol{\sigma}) e^{-2\pi i \nu \mathbf{b}\cdot\boldsymbol{\sigma}/c} \, d\Omega$$

For the coordinate system in the figure:

$$\frac{\nu\,\mathbf{b}\cdot\mathbf{s}}{c} = ul + vm + wn$$

$$\frac{\nu\,\mathbf{b}\cdot\mathbf{s}_0}{c} = w,$$

$$d\Omega = \frac{dl\,dm}{n} = \frac{dl\,dm}{\sqrt{1-l^2-m^2}}$$

$$V(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{A}(l,m) I(l,m) e^{-2\pi i\left[ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right]} \frac{dl\,dm}{\sqrt{1-l^2-m^2}}$$

Inverse transform will give us the sky brightness I

Imaging: FT relationship

FT relationship exists between the sky brightness I, the primary beam pattern A and the visibility V observed with an interferometer:

$$\mathcal{A}(l,m)I(l,m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

2-D relationship holds while:

$$\left|\frac{\Delta\nu}{c}\,\mathbf{b}\cdot(\mathbf{s}-\mathbf{s_0})\right| \ll 1$$

$$\left|w(l^2+m^2)\right| \ll 1$$

Observations are confined to a small region of the sky.

Imaging

$$\mathcal{A}(l,m)I(l,m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

Primary beam

The primary beam correction can be made at the final stage after obtaining the image – thus we will drop this in the equations henceforth. Let I itself denote the sky brightness modified by the primary beam.

Imaging

$$I(l,m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

Discrete measurements:

$$(u_k, v_k),\ k = 1, \ldots, M$$

M depends on the number of antennas in an array
For an array of 30 antennas like the GMRT, M ?

Imaging

$$I(l,m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

Image

Visibilities

In practice you have the following:

$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

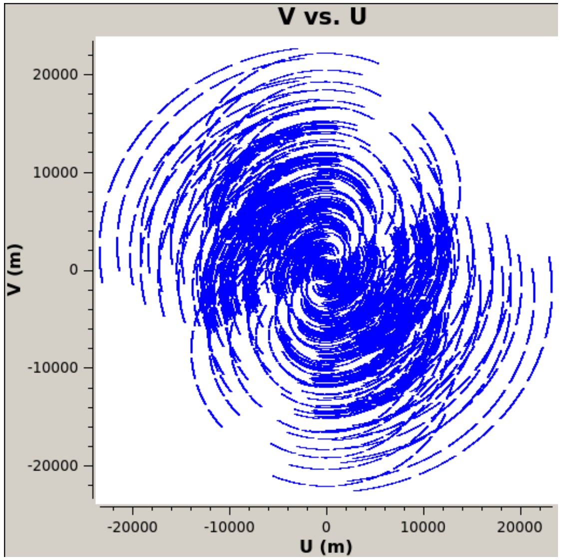
“Dirty” image

Sampling function

Observed visibilities

Imaging

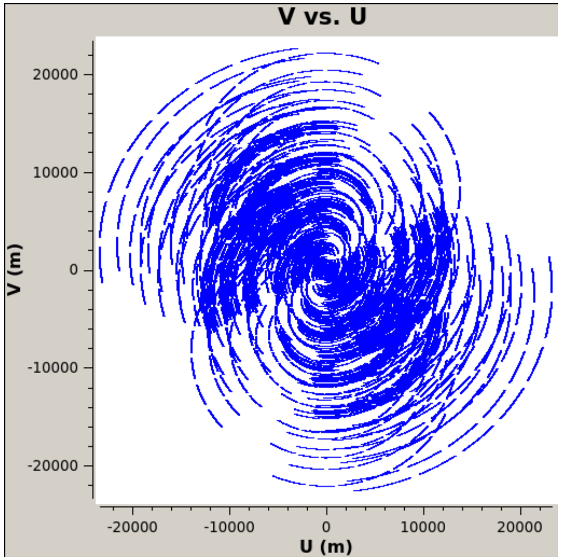
$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$



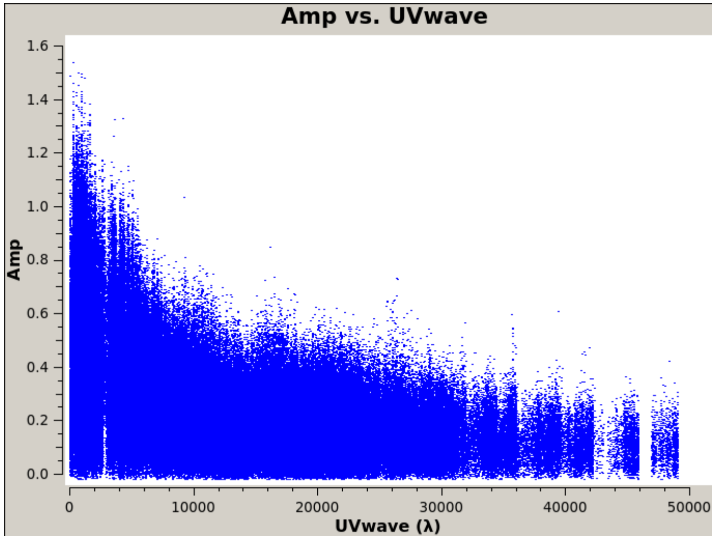
Sampling

Imaging

$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$



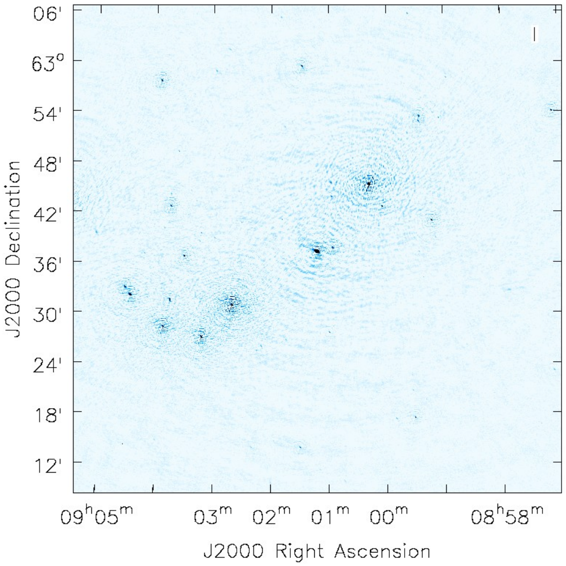
Sampling



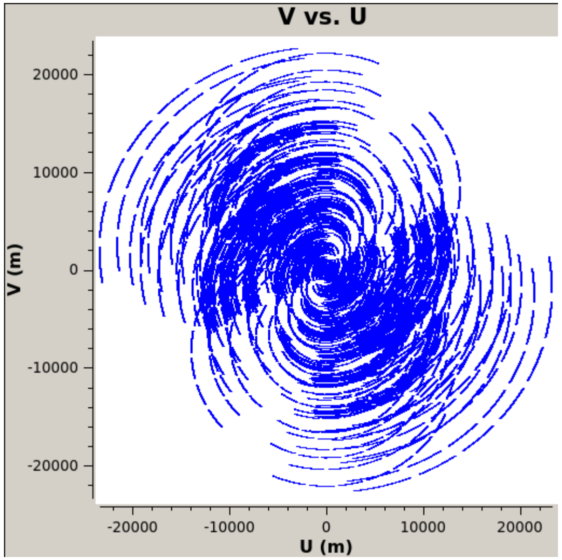
Observed visibilities (complex numbers) Only amp. Shown.

Imaging

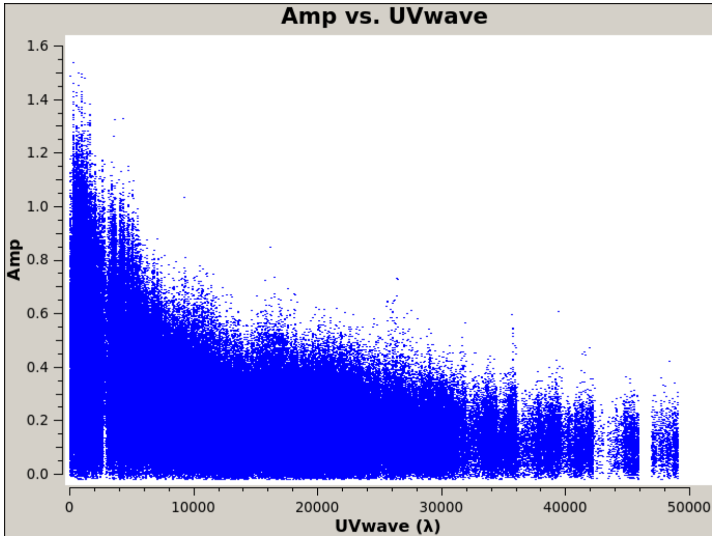
$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$



“Dirty” image



Sampling



Observed visibilities (complex numbers) Only amp. Shown.

Imaging: two methods

$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

Direct Fourier Transform

Discrete Fourier Transform (using Fast Fourier Transform)

Imaging

$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$

Direct Fourier Transform:

$$\frac{1}{M}\sum_{k=1}^{M} V'(u_k,v_k)e^{2\pi i(u_k l+v_k m)}$$

To be evaluated at every point of a NxN grid.

Number of multiplications needed to evaluate are $\sim 2MN^2$

M and N are of the same order and thus the number of multiplications needed are $\sim N^4$

Direct Vs Discrete Fourier Transform

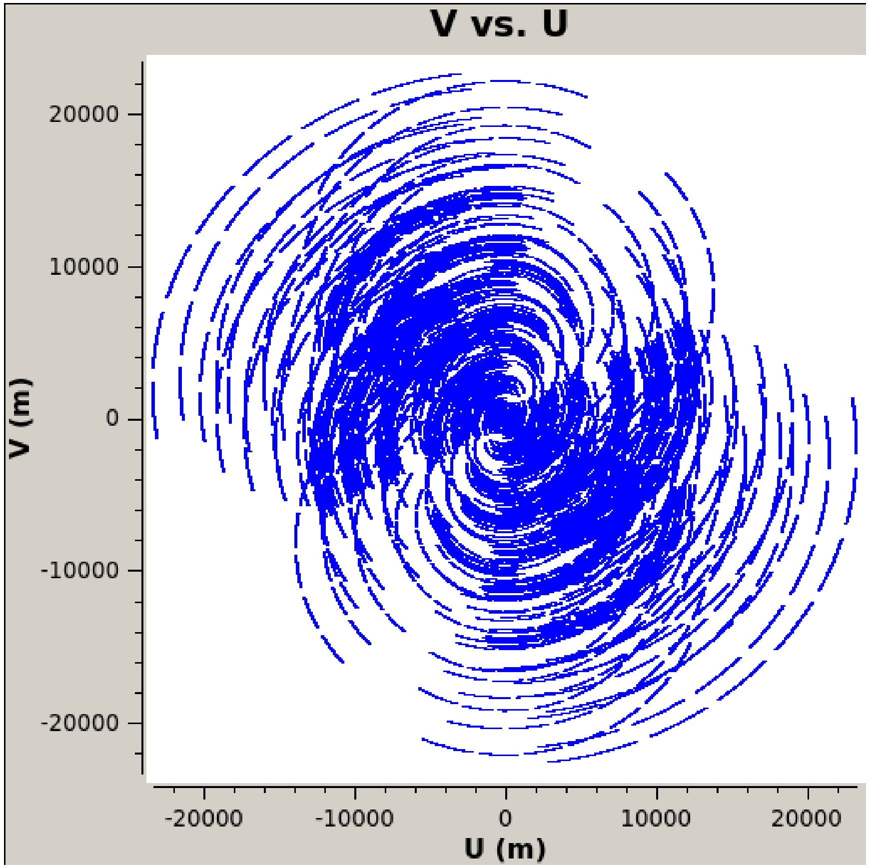
Due to computational advantages, fast algorithms to find the Discrete Fourier Transform (DFT) are most commonly used in radio astronomy (algorithm for DFT: Fast Fourier Transform - FFT).

Application of FFTs requires *bringing data to a regular grid* and then perform the transform.

Only in special cases where number of antenna elements are few, the “direct Fourier Transform” is used.

Fast Fourier Transform

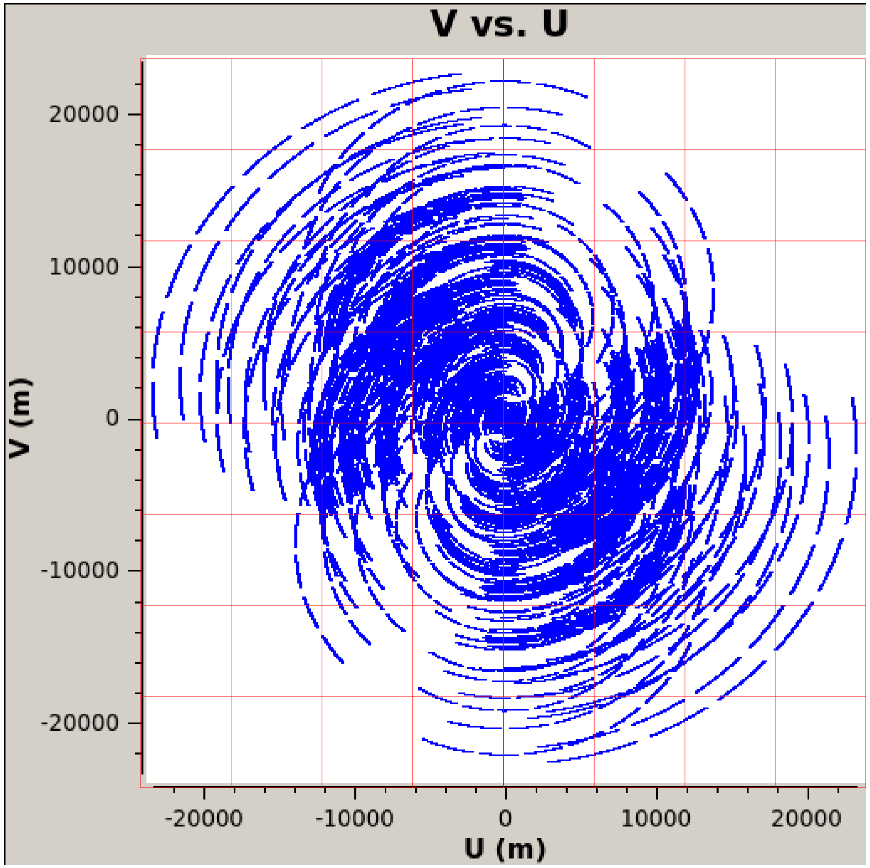
Sampling



Fast Fourier Transform

Requires the data to be on a regular grid.

Gridding



Grid shown is only illustrative.

Fast Fourier Transform

Requires the data to be on a regular grid.

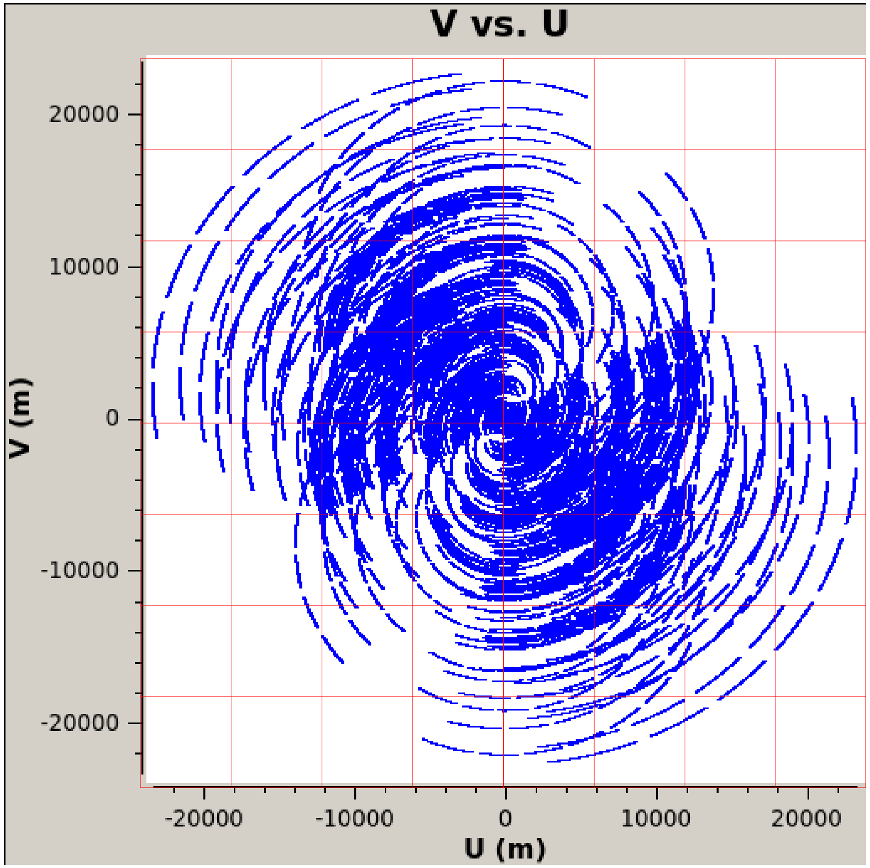
Gridding

To bring the data to a regular grid requires ~ N operations.

Further the FFT algorithms only require ~ $N^2 \log_2 N$ operations. (E. g. Cooley-Tukey algorithm)

Compare this with N^4 for the DFT case

In most common situations, FFTs are used.

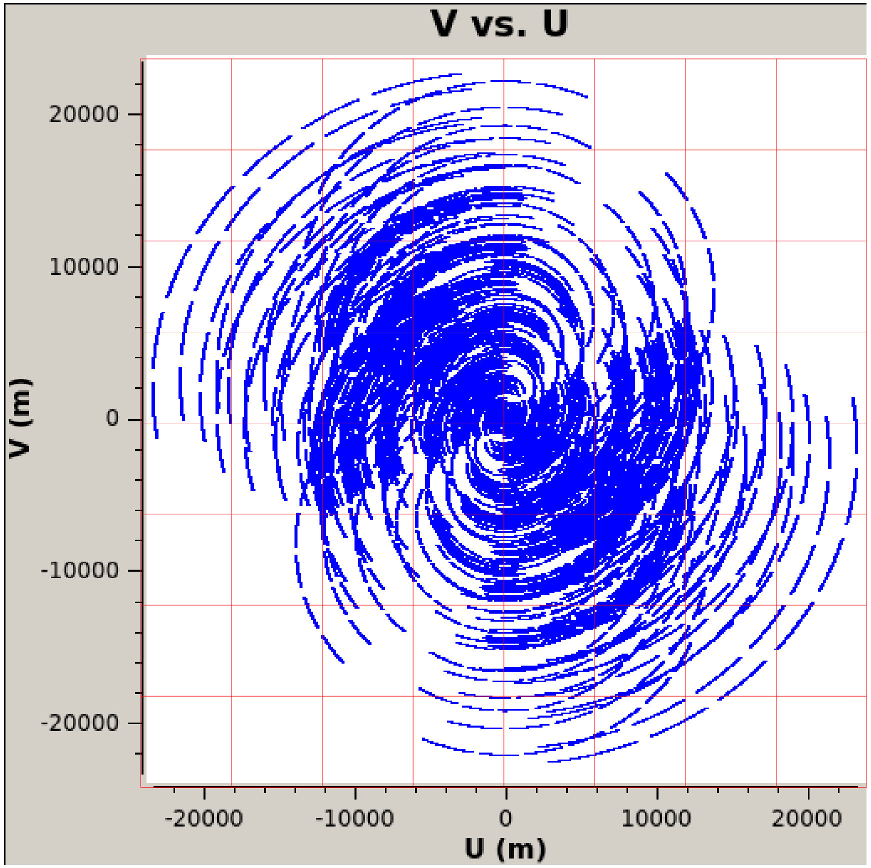


Grid shown is only illustrative.

Gridding the visibilities

Motivated by the fact that we want to take full advantage of the FFT algorithms.

We want the data on a “grid” that is uniformly spaced with a power of two points on each side.



Grid shown is only illustrative – grid size needs to be chosen with due consideration of error tolerence and computation

Zoom onto a single point in the uv-plane

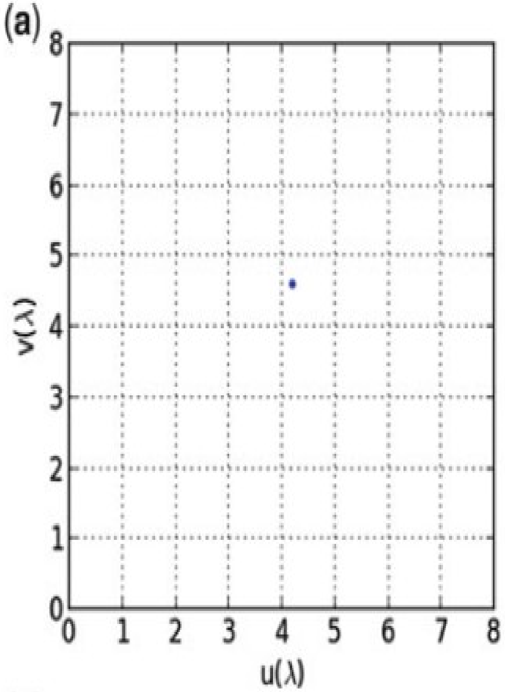
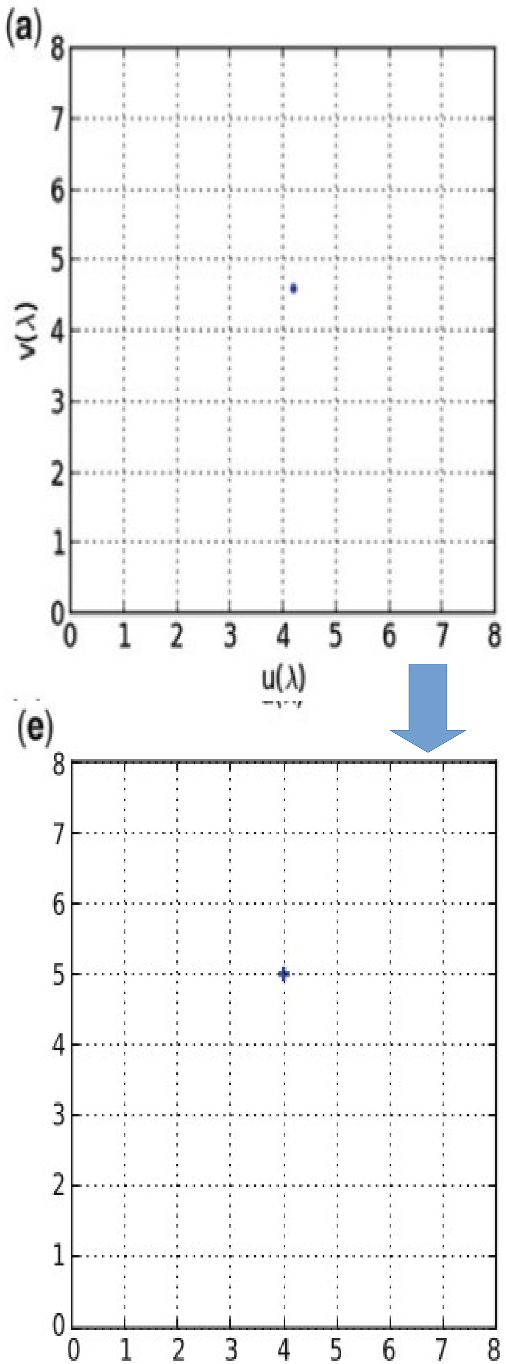


Illustration of gridding

Zoom onto a single point in the uv-plane



How to achieve this ?

Illustration of gridding

Zoom onto a single point in the uv-plane

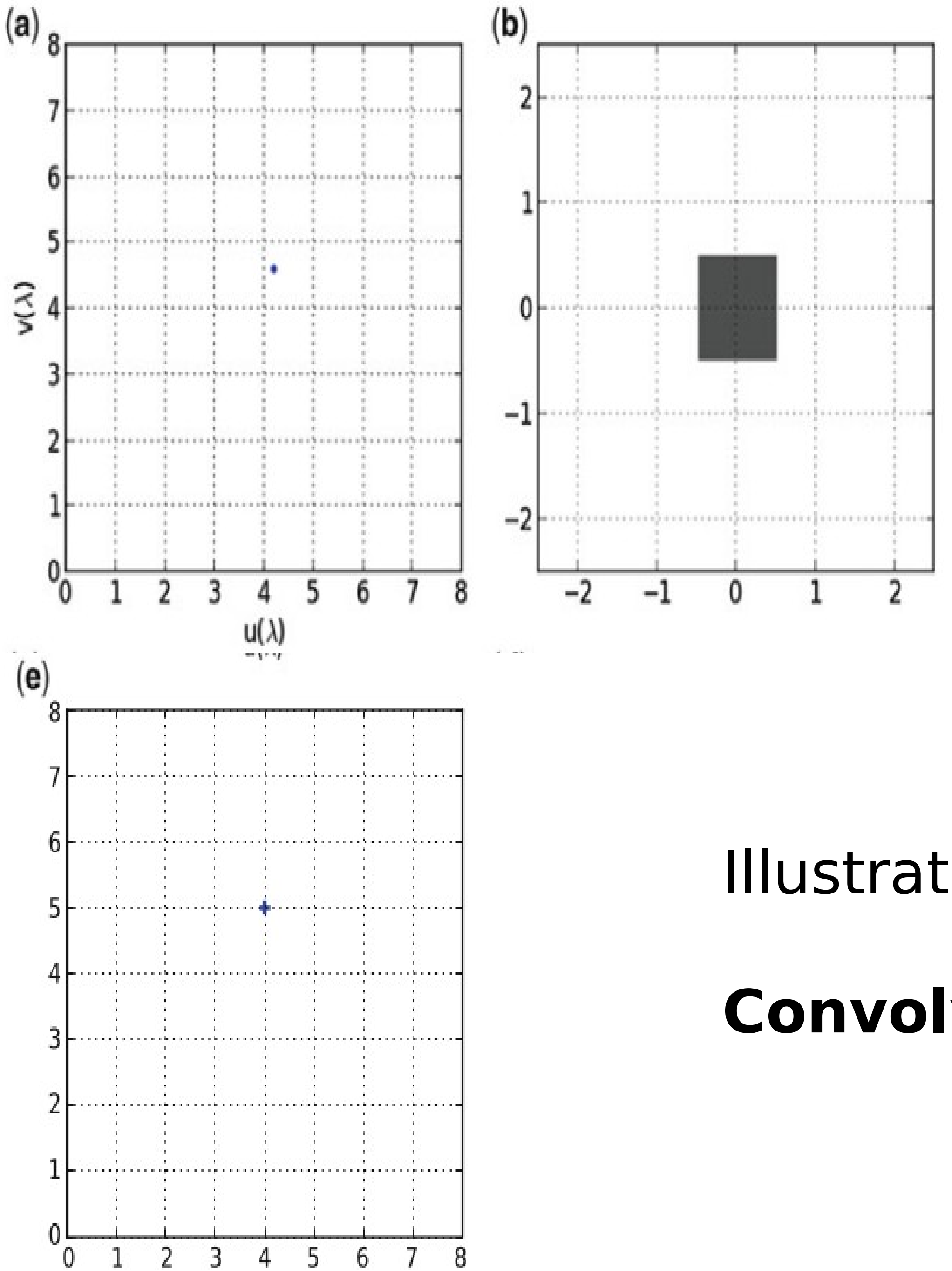
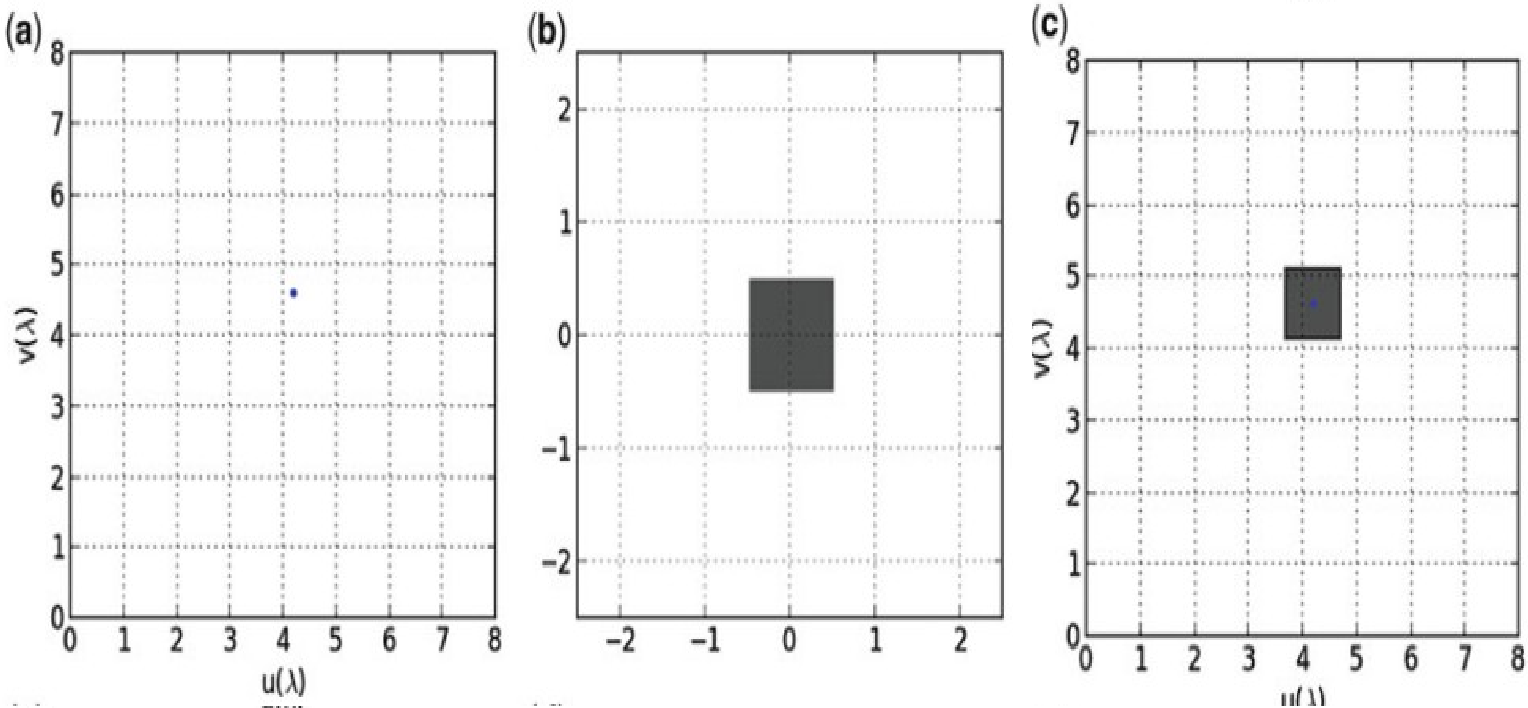


Illustration of gridding:

Convolve (a) by (b)

Zoom onto a single point in the uv-plane



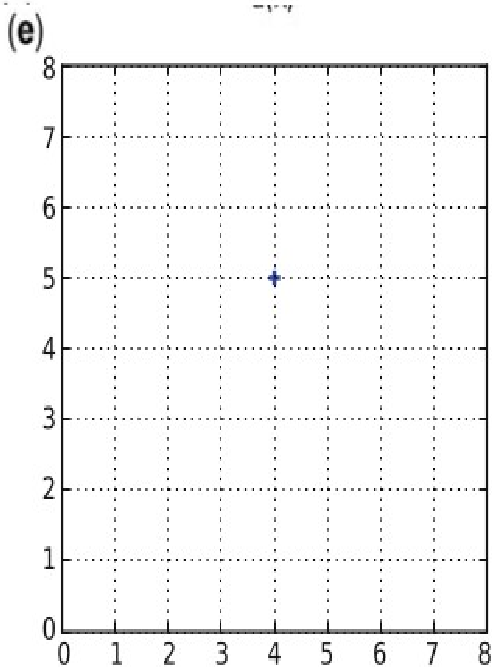
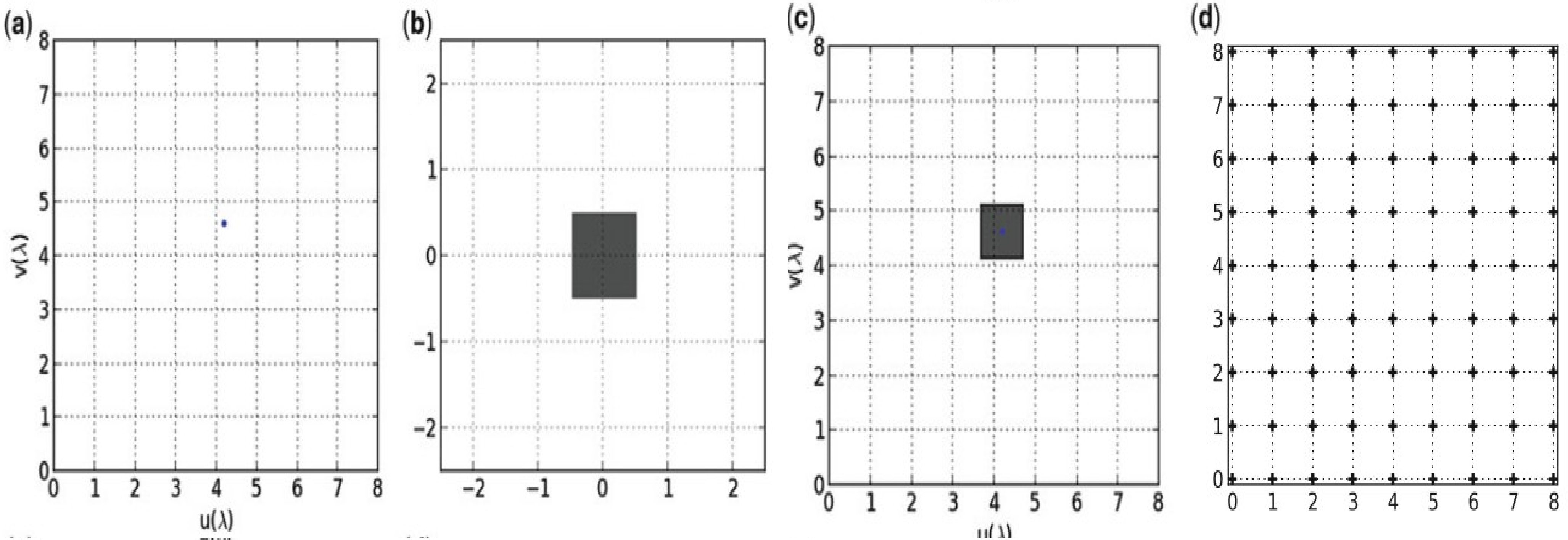


Illustration of gridding:

Convolve (a) by (b)

Zoom onto a single point in the uv-plane



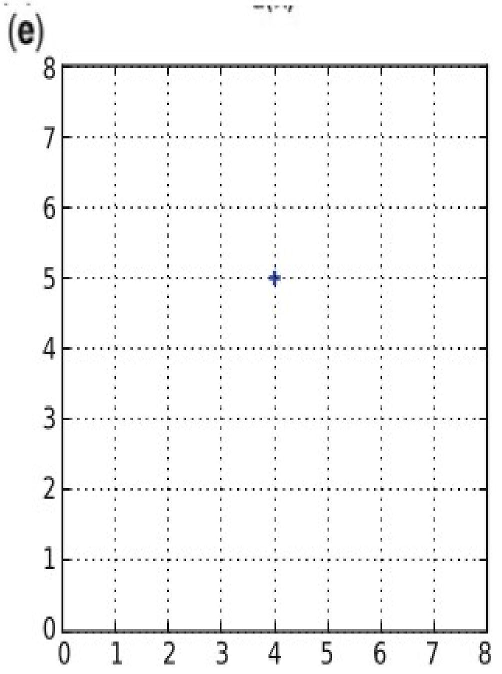
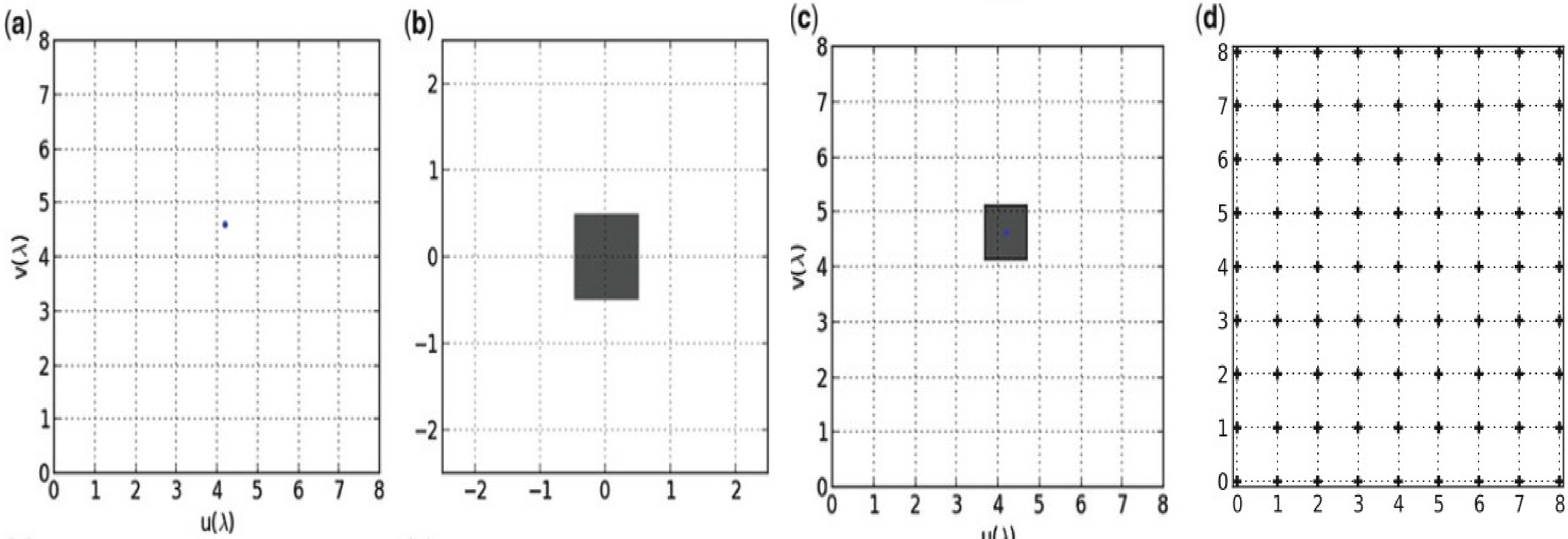
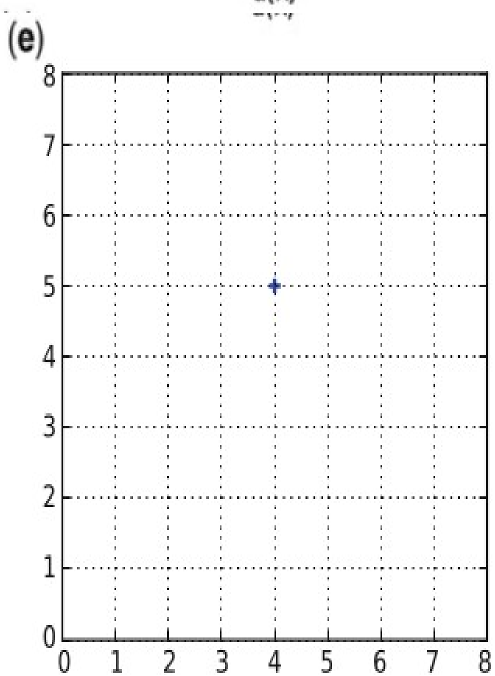


Illustration of gridding:

Convolve (a) by (b)
Multiply (c) and (d)
Output = Gridded visibility





Effect of this on the image?

Convolve (a) by (b) - multiplication with the FT of box

Multiply (c) and (d) – convolution with FT of sampling function

Output = Gridded visibility

Spacing of uv grid points is inversely proportional to the field of view of the image, the comb function that is applied in the image plane has a spacing proportional to the field of view.

This can fold sources outside the field of view into the image.

Replace the box function with a function that has a FT pair with more desirable properties !

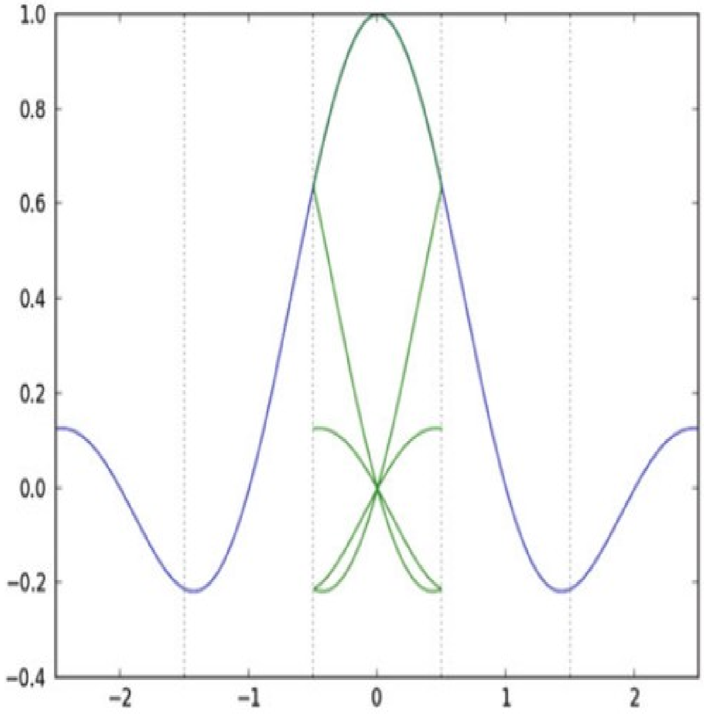


Fig. 8.4 Box function and aliasing. The Fourier transform of a box function is the sinc function (blue curve). Convolution in the *uv* domain by a box function is equivalent to multiplication in the image domain by a taper with considerable sidelobes. Subsequent sampling in the *uv* domain causes the sidelobes to show up in the main image (green curves)

What is the optimal function ?

The optimal function is the one that minimises the energy in the sidelobes in the image domain, under the constraint that its “support“ in the uv domain is within a certain limit.

Energy in desired region/Total energy

$$\frac{\int_{-A}^{A} |f(t)|^2 dt}{\int_{-\infty}^{\infty} |f(t)|^2 dt}$$

Slepian and Pollak 1961

Prolate Spheroidal Wave Function

The optimal function is the one that minimises the energy in the sidelobes in the iamge domain, under the constraint that its „support“ in the uv domain is within a certain limit.

Energy in desired region/Total energy

$$\frac{\int_{-A}^{A} |f(t)|^2 dt}{\int_{-\infty}^{\infty} |f(t)|^2 dt}$$

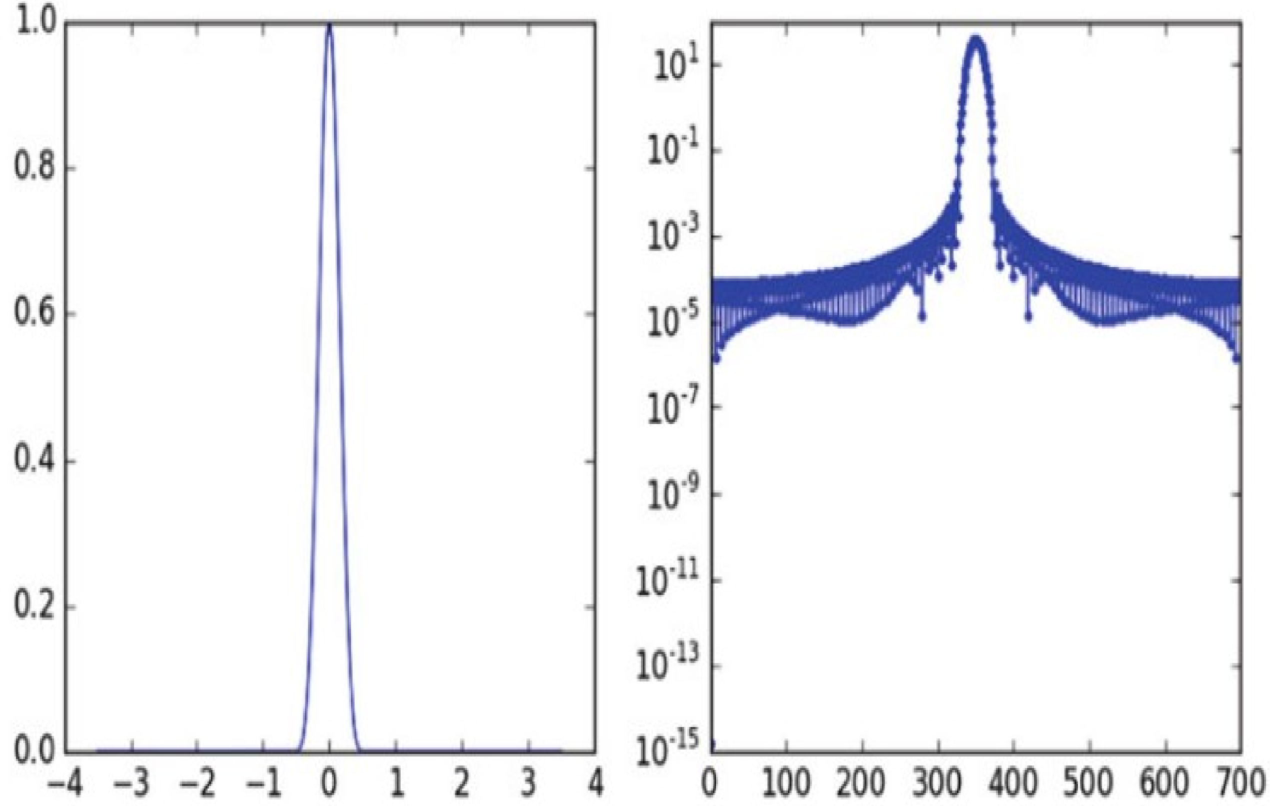


Fig. 8.5 The prolate spheroidal wave function. Left: the spheroidal function. Right: the Fourier transform of the spheroidal function

Slepian and Pollak 1961

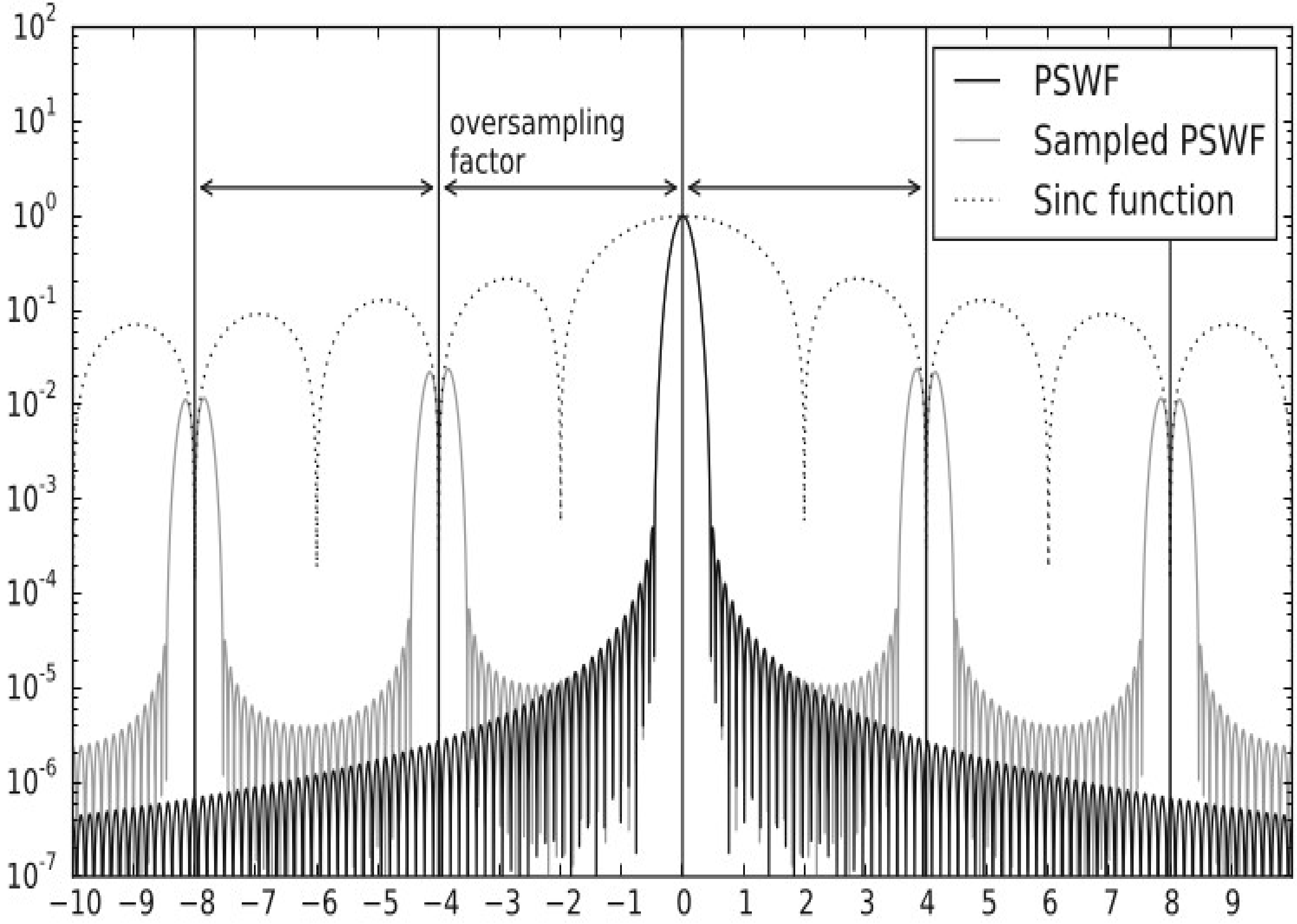
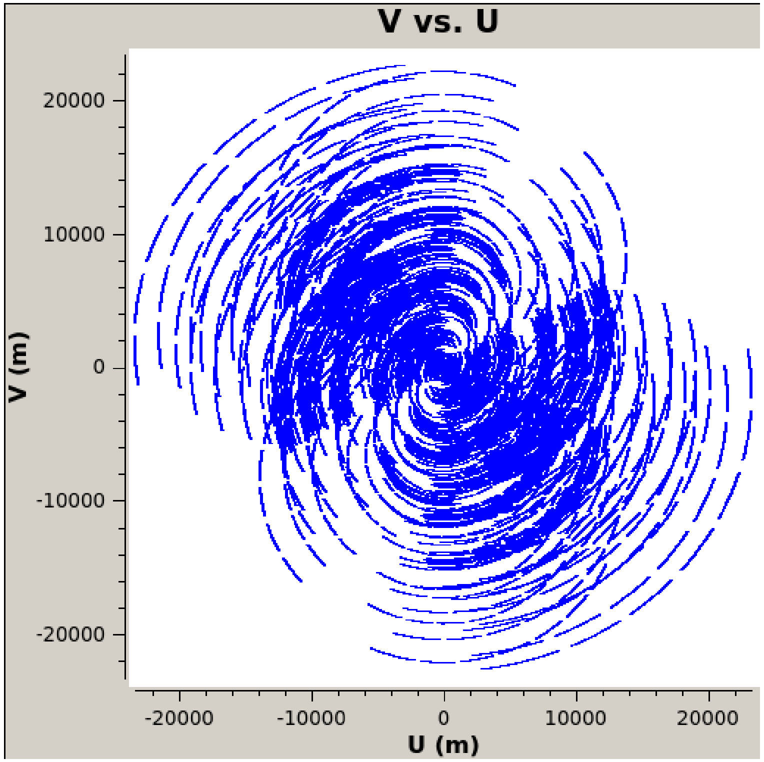


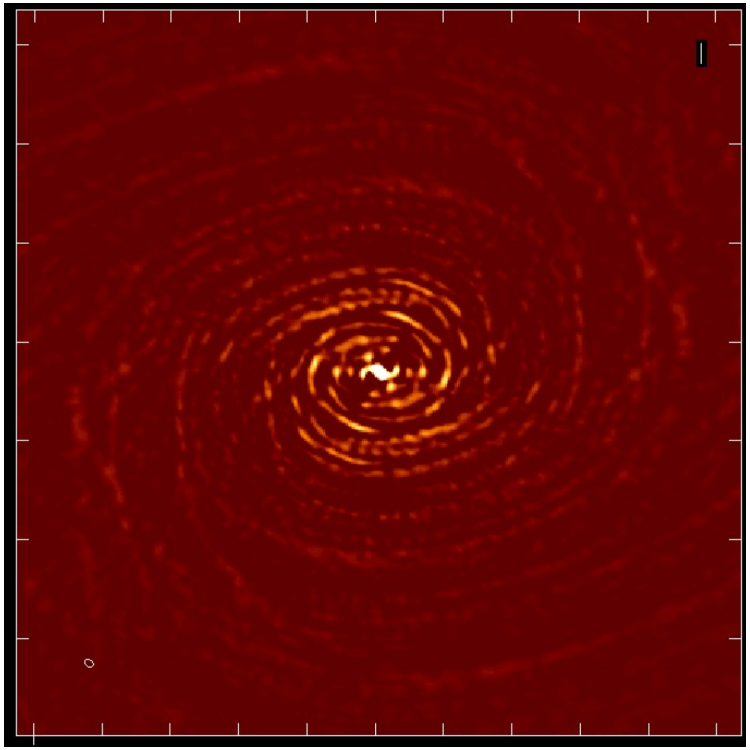
Fig. 8.7 Illustration of the aliasing induced by the gridding procedure, and its suppression by oversampling the convolution function

Increased oversampling – sidelobes will be pushed farther outwards and decreased in amplitude.

Synthesized beam



FT
→



Desirable characteristics: Low and uniform sidelobes; high resolution

No unique approach to get all of this. Choice according to the science requirement.

Weighting: control the shape of the beam

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

$$B = \mathfrak{F} S$$

Introduce a weighted sampling distribution:

Weighting: control the shape of the beam

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

$$B = \mathfrak{F} S$$

Introduce a weighted sampling distribution:

$$W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k)$$

T_k = tapering function
D_k = density weighting
R_k = reliability weight

Weighting: control the shape of the beam

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

$$B = \mathfrak{F} S$$

Introduce a weighted sampling distribution:

T_k = tapering function
D_k = density weighting
R_k = reliability weight

$$W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k)$$

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

Weighted visibilities

$$V^W = W V'$$

$$V^W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k) V'(u_k, v_k)$$

Weighting: control the shape of the beam

$$W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k)$$

T_k = tapering function
D_k = density weighting
R_k = reliability weight

$$V^W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k) V'(u_k, v_k)$$

If the sampling were a smooth function like a Gaussian we would have no sidelobes.
However it is like a bunch of delta functions – often with large gaps in between.

In an array: typically data points are dense in the inner region of the uv-plane and are sparse outside – gives rise to more weight to shorter spacings.

Briggs 1995 (PhD thesis: detailed treatment of weighting of visibilities)

Weighting: control the shape of the beam

$$W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u-u_k, v-v_k)$$

T_k = tapering function
D_k = density weighting
R_k = reliability weight

$$V^W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u-u_k, v-v_k) V'(u_k, v_k)$$

Tapering weights are used to downweight the data at the outer edge.
Density weights are used to lessen the effect of non-uniform density of sampling in the uv-plane.

The weights are factored into components arbitrarily - only for convenience.

Briggs 1995 (PhD thesis: detailed treatment of weighting of visibilities)

Weighting: control the shape of the beam

$$W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k)$$

T_k = tapering function
D_k = density weighting
R_k = reliability weight

$$V^W(u,v) = \sum_{k=1}^{M} R_k T_k D_k \delta(u - u_k, v - v_k) V'(u_k, v_k)$$

T_k = tapering function, separable into u and v dependent parts.

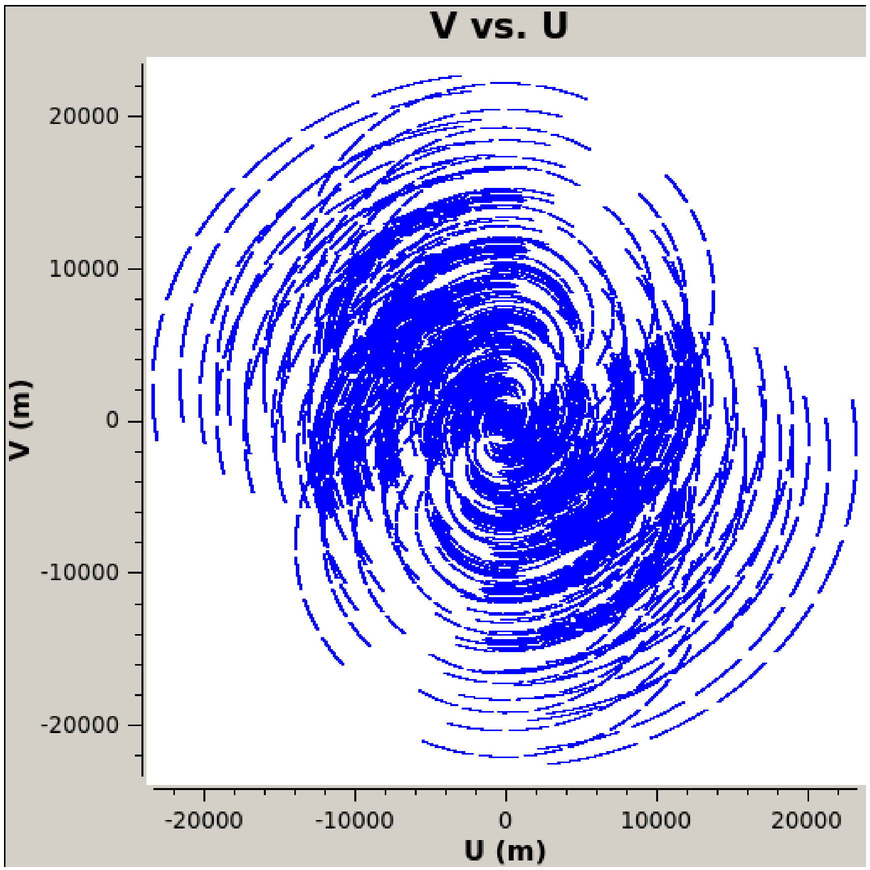
$$T(u,v) = T_1(u) T_2(v)$$

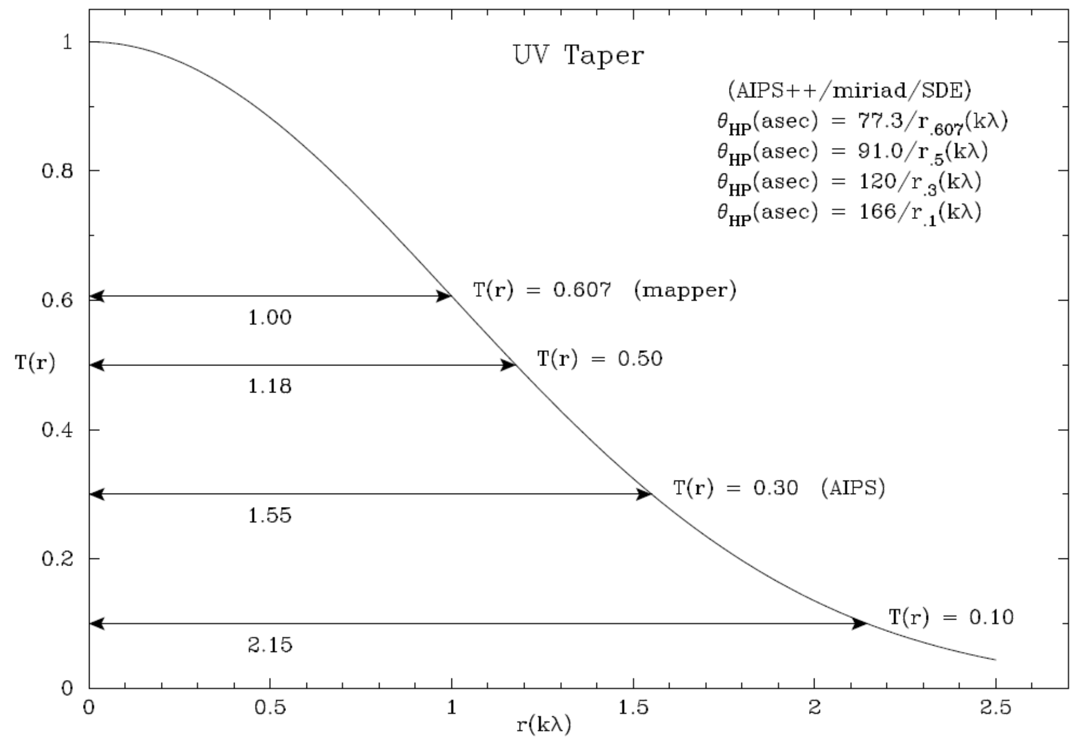
$$T_k = T(r_k) \qquad r_k \equiv \sqrt{u_k^2 + v_k^2}.$$

A Gaussian taper, for example:

$$T(r) = \exp(-r^2/2\sigma^2)$$

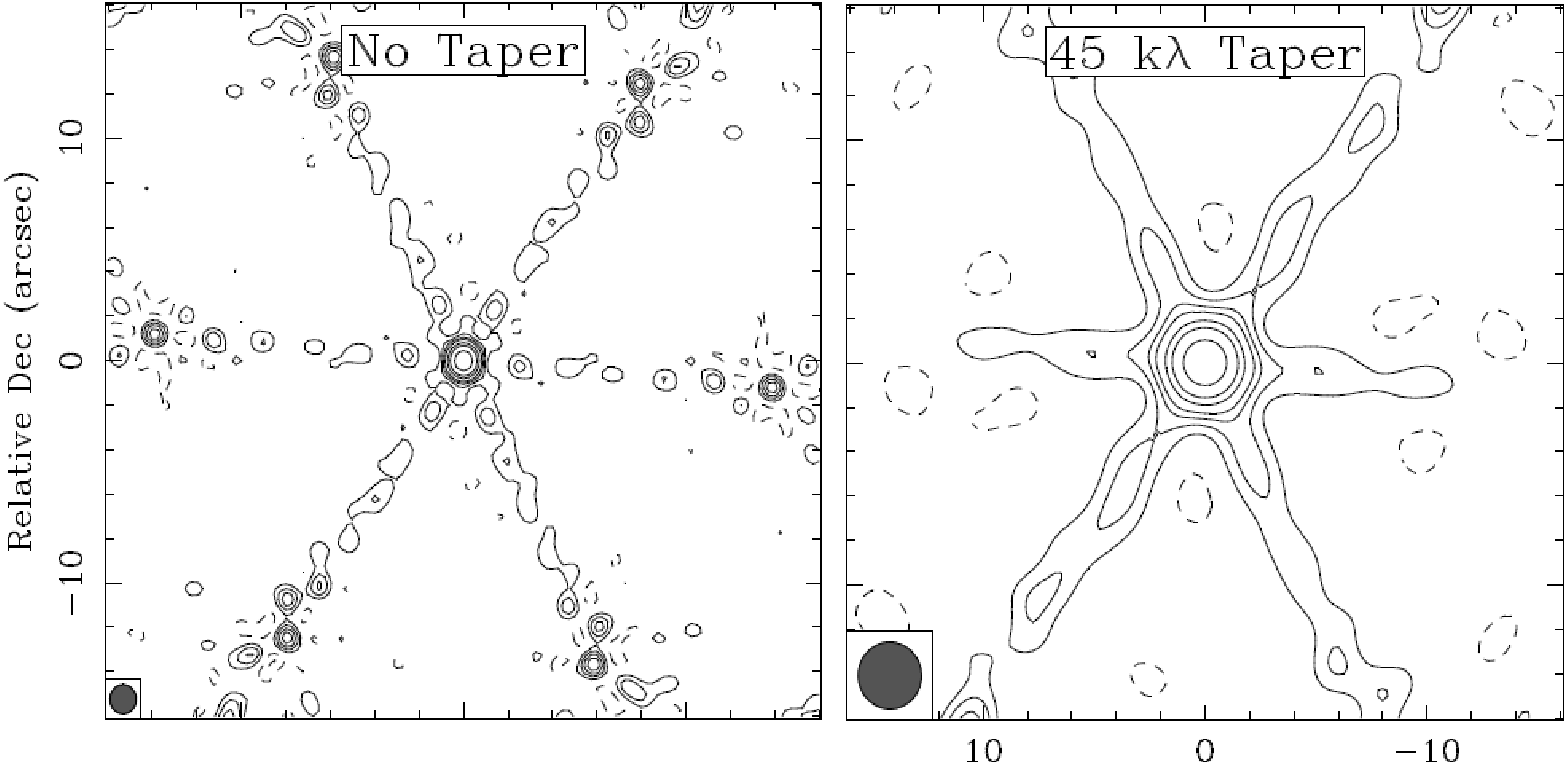
Tapering





The synthesized beam width will change depending on the choice of the taper.

Tapering example



Density weighting

Natural weights

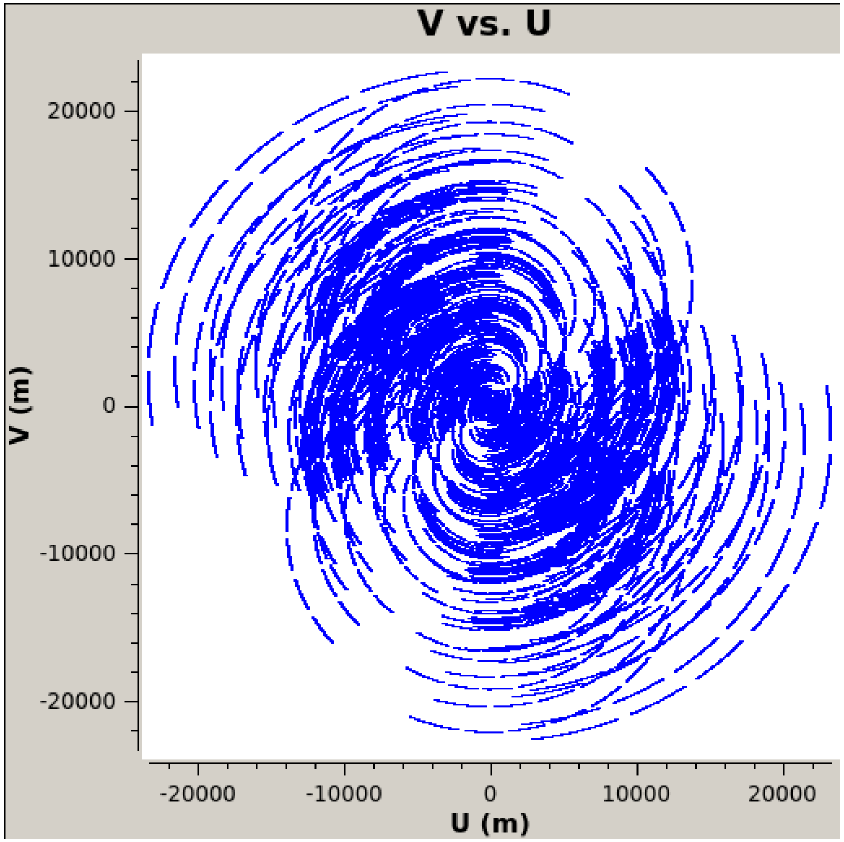
$$D_k = 1$$

Uniform weights

$$D_k = \frac{1}{N_s(k)}$$

$N_s(k)$ is the number of points within a symmetric region in (u,v) of width s centered on k^{th} point.

N_s is the number of points within a grid cell.



Density weighting

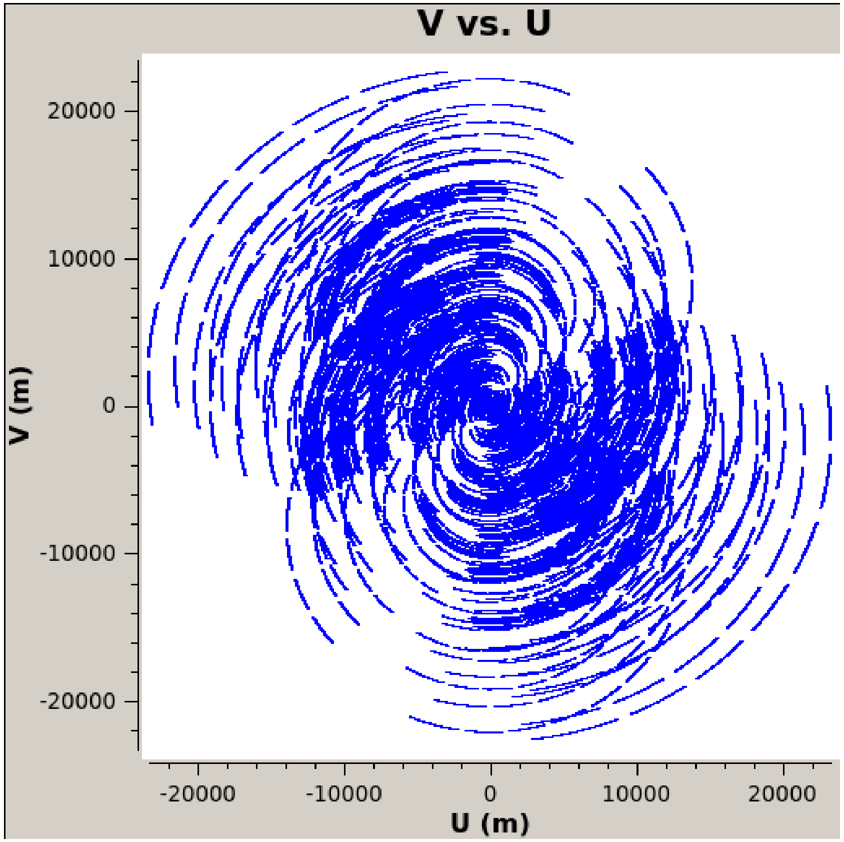
Natural weights

$$D_k = 1$$

Uniform weights

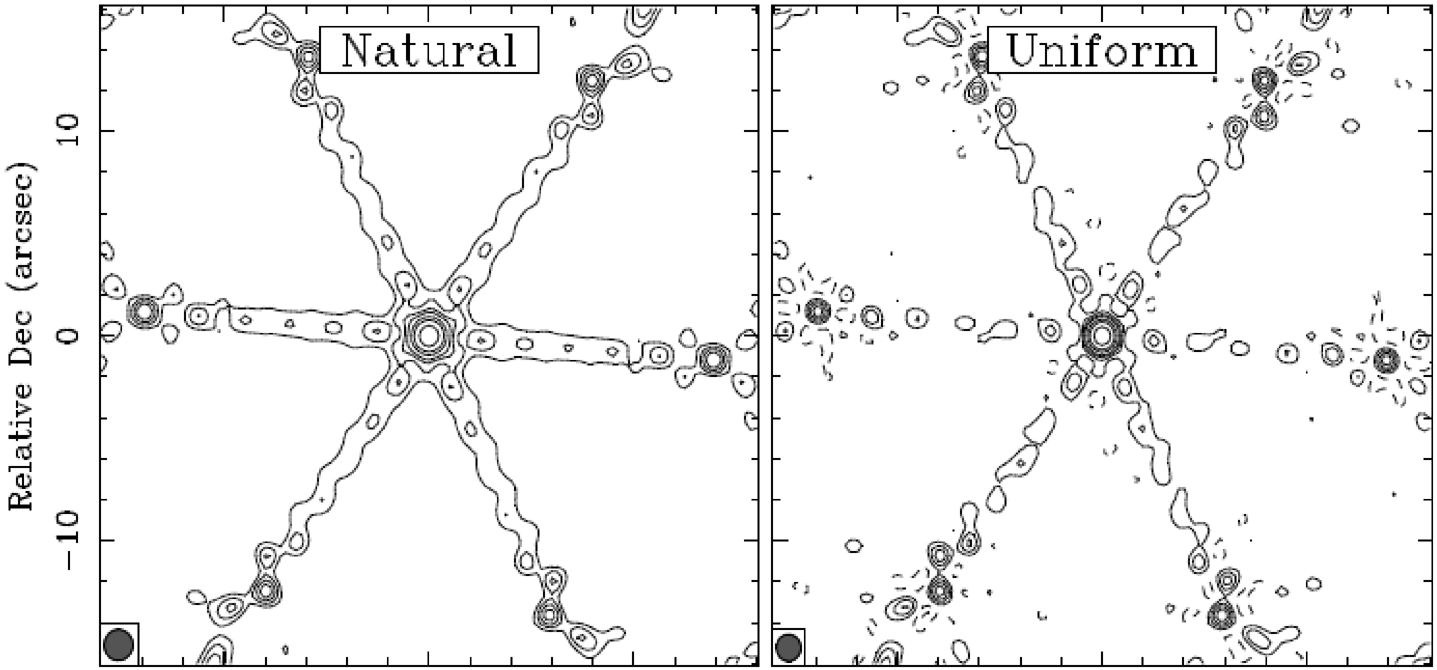
$$D_k = \frac{1}{N_s(k)}$$

$N_s(k)$ is the number of points within a symmetric region in (u,v) of width s centered on k^{th} point.



Robust weighting: hybrid form of weighting: uses minimisation of summed sidelobe power and thermal noise.

Density weights example

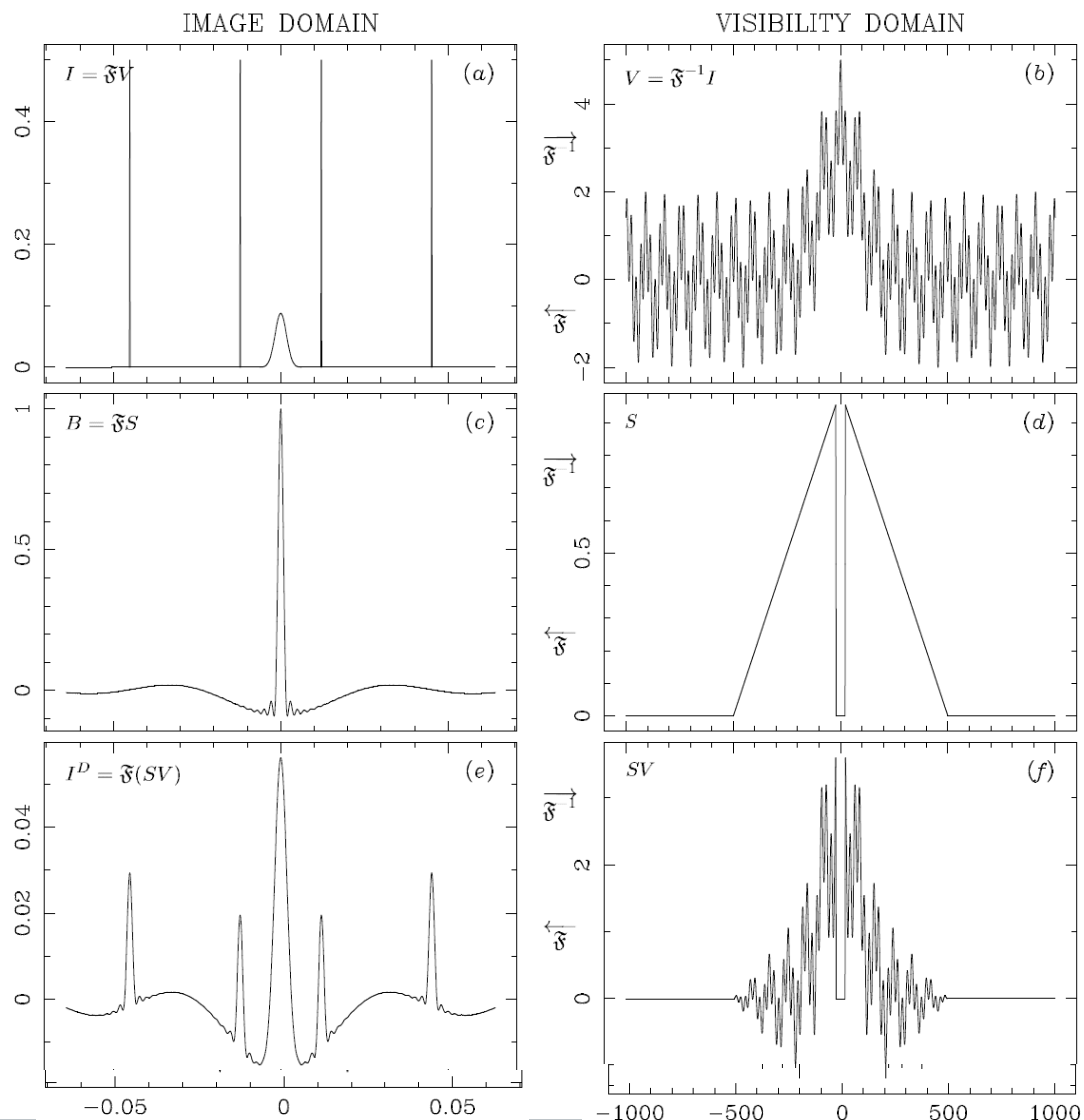


Graphical representation

Model source: symmetric

Synthesized beam

Dirty image if a direct FT is computed



Model Visibilities: Real and even due to symmetry

Sampling: central hole, falling density towards the outskirts

Sampled visibilities (these are the observed ones)

Radians

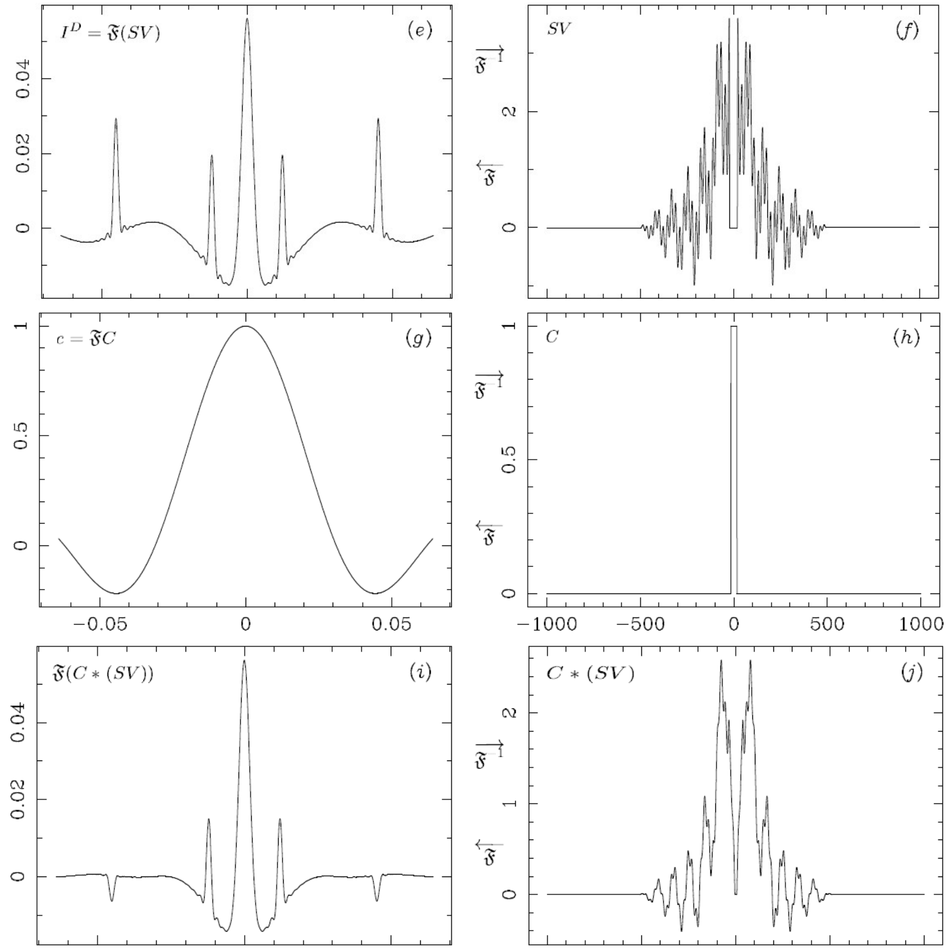
Baseline lengths (wavelengths)

SIRA, Fig. 7-5

Dirty image if a direct FT is computed

FT of the convolution function

Effect in the image domain



Sampled visibilities

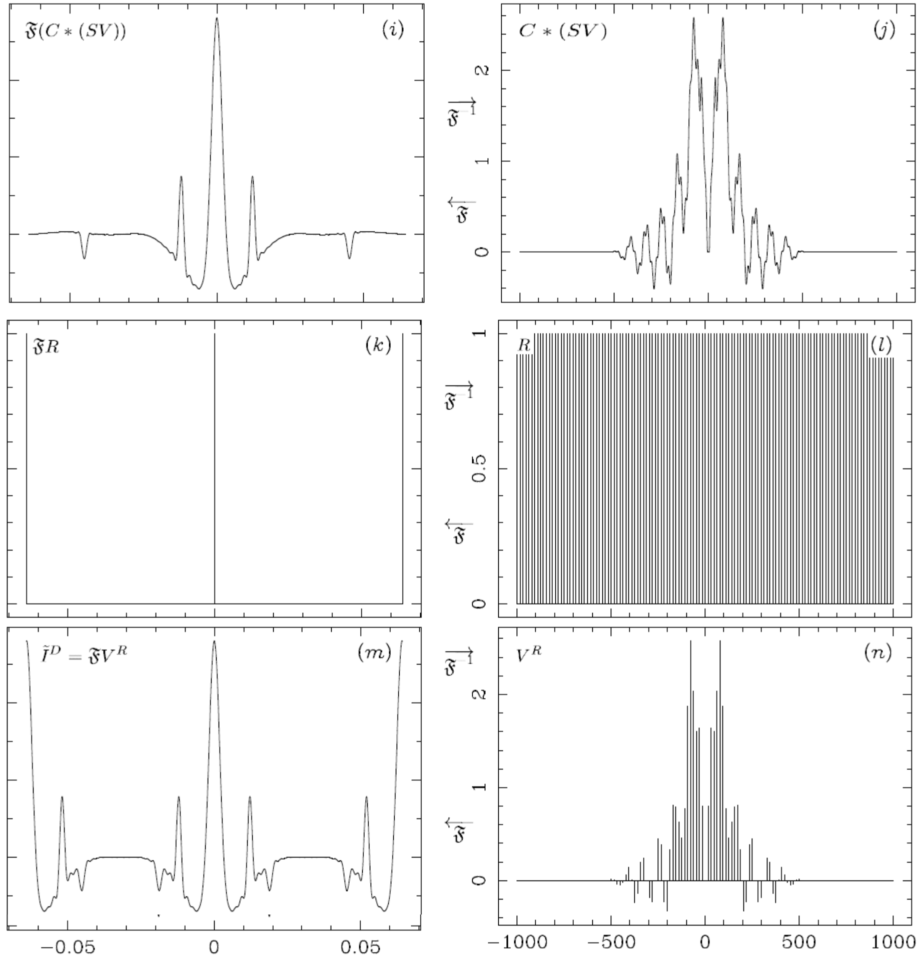
Convolution function

Convolved sampled visibilities

Effect in the image domain

FT of Resampling

Dirty image: aliasing

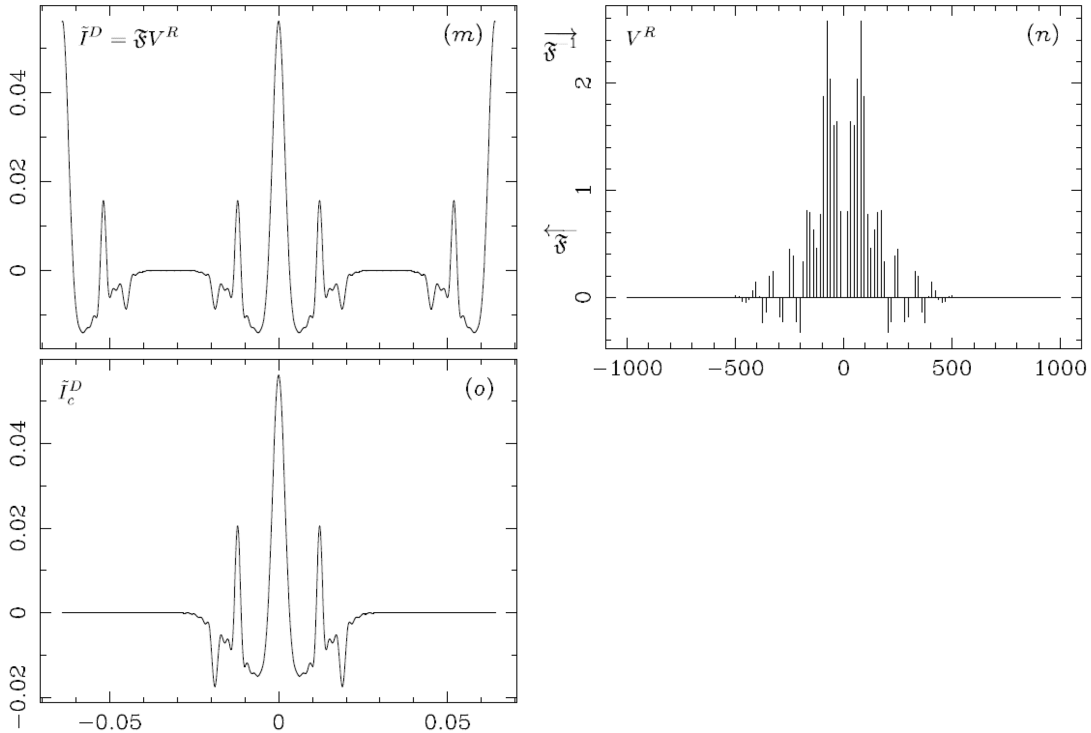


Convolved sampled visibilities

Resampling

Resampled visibility

Dirty image: aliasing



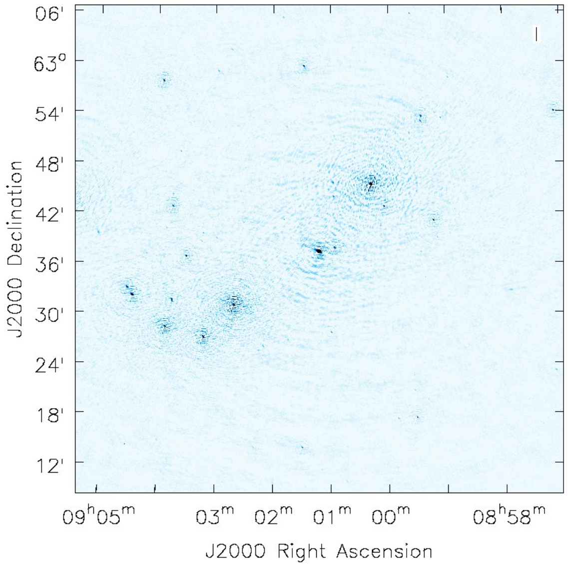
Resampled visibility

Divide by the FT of the convolution function

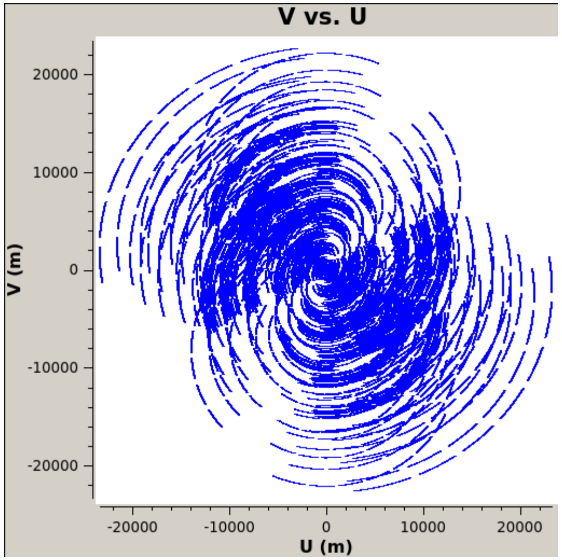
This image is far from satisfactory representation of the actual distribution: can do better than this by **deconvolution.**

Imaging: FT of gridded visibilities

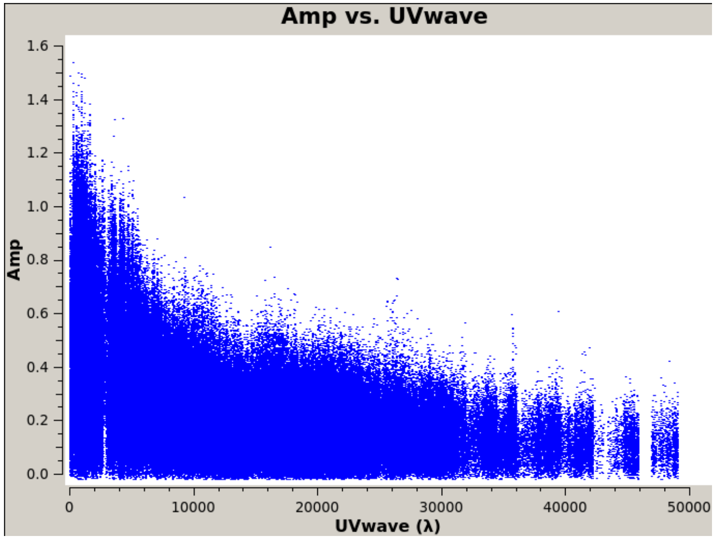
$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)V'(u,v)e^{2\pi i(ul+vm)}\,du\,dv$$



"Dirty" image



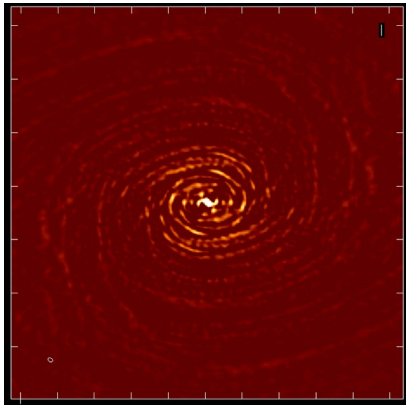
Sampling



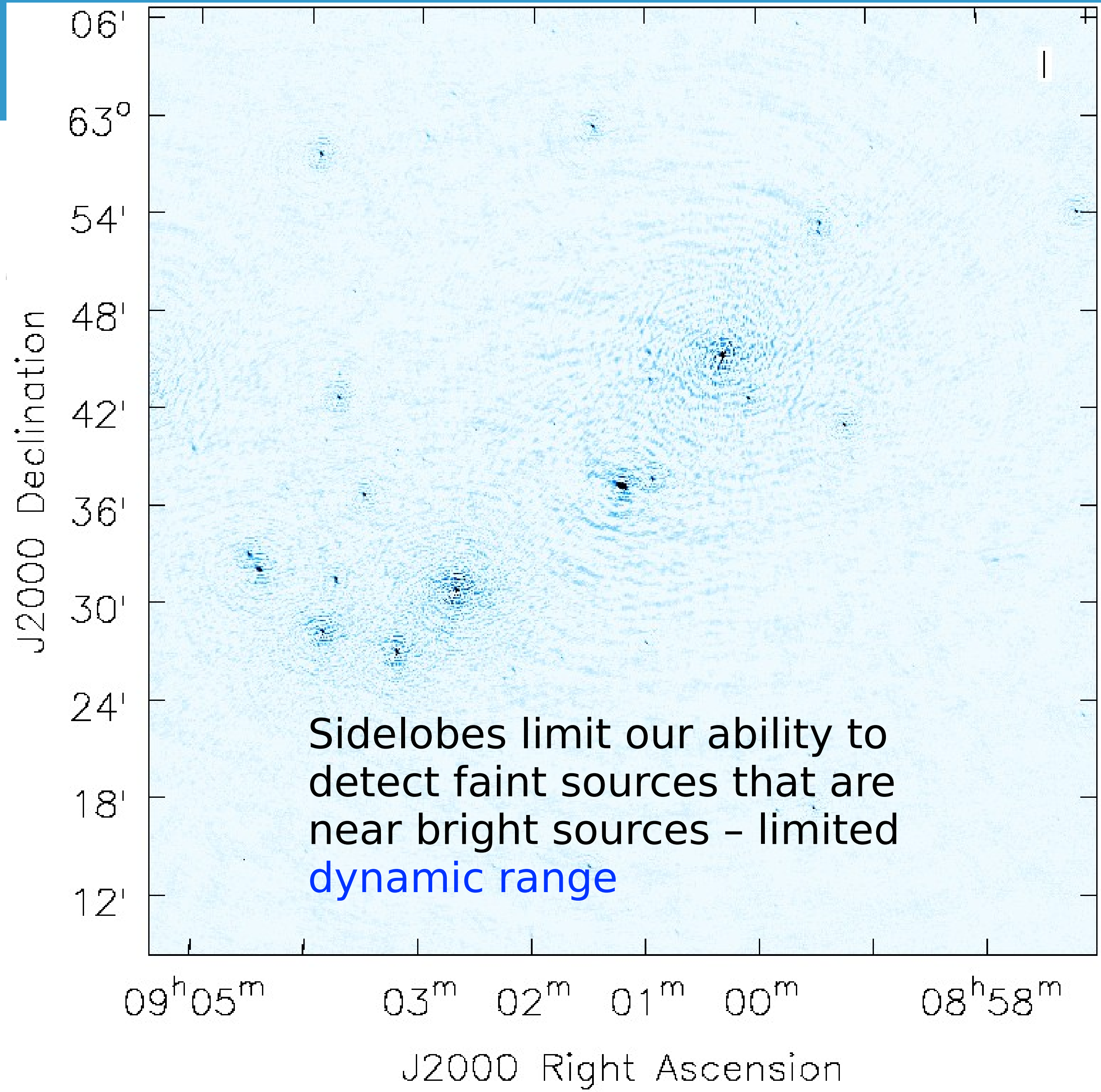
Observed visibilities (complex numbers) Only amp. Shown.

Imaging

$$I^D(l,m) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}$$



Point source response
or synthesized beam

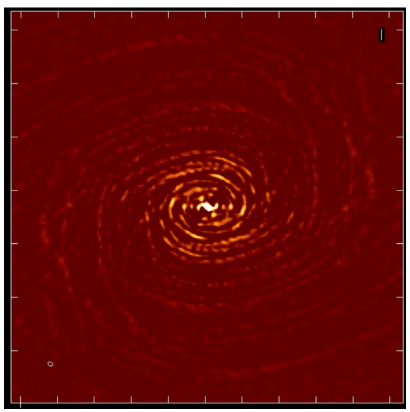


Sidelobes limit our ability to detect faint sources that are near bright sources – limited dynamic range

“Dirty” image

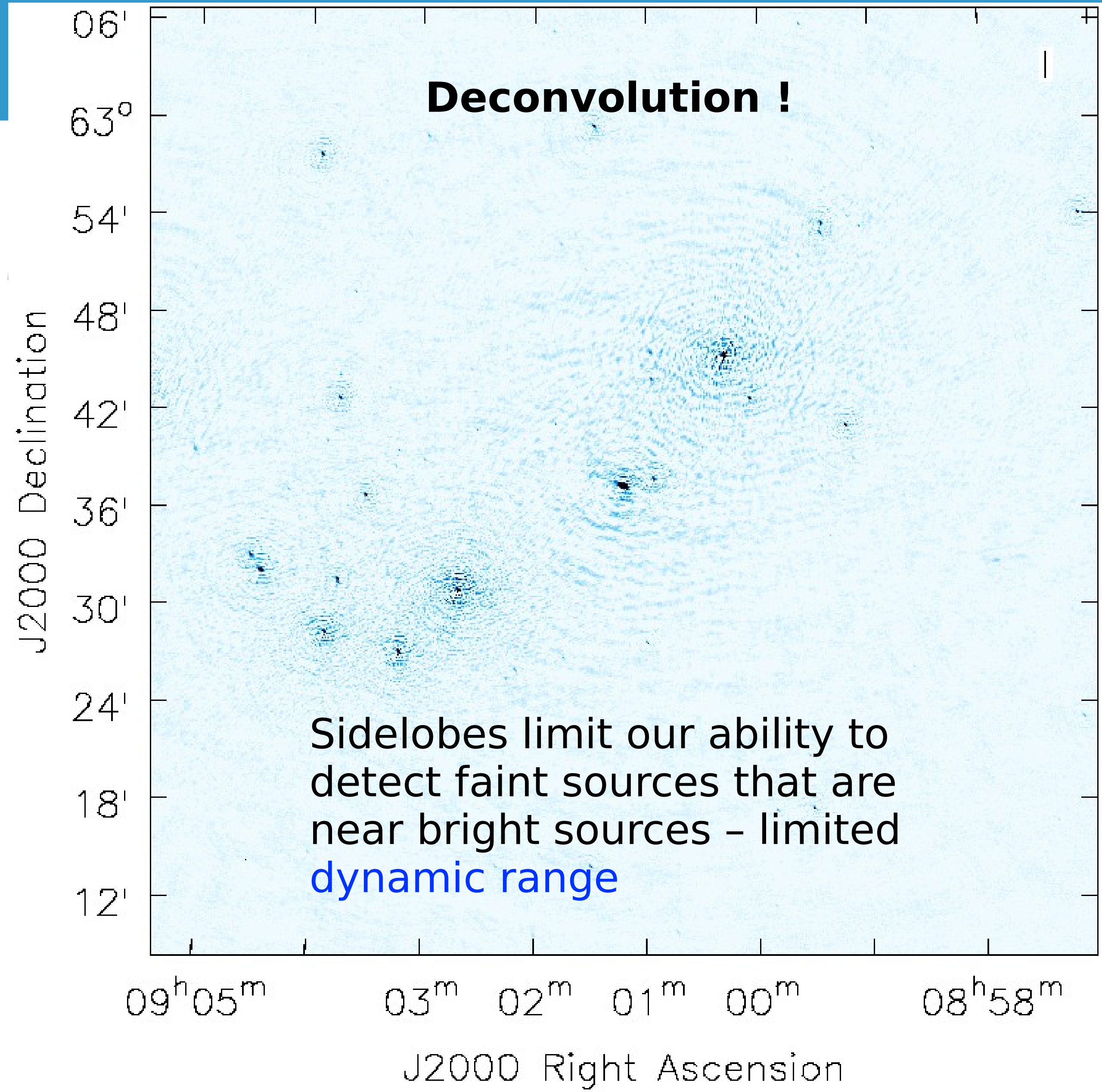
Imaging

$$I^D(l,m) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty}$$



Point source response
or synthesized beam

“Dirty” image



Summary

- Imaging: FT of the sampled visibilities
- FFT: fast computation but requires the visibilities on a regular grid
- Visibility gridding: achieved typically by “convolution”
- Synthesized beam: FT of the sampling. Can be modified based on the choice of weights.
- Natural and uniform weights
- Tapering
- Graphical representation of the steps involved in the process of obtaining an image using gridded visibilities.