Imaging and Deconvolution

ATNF Radio Astronomy School 2017

Mark Wieringa

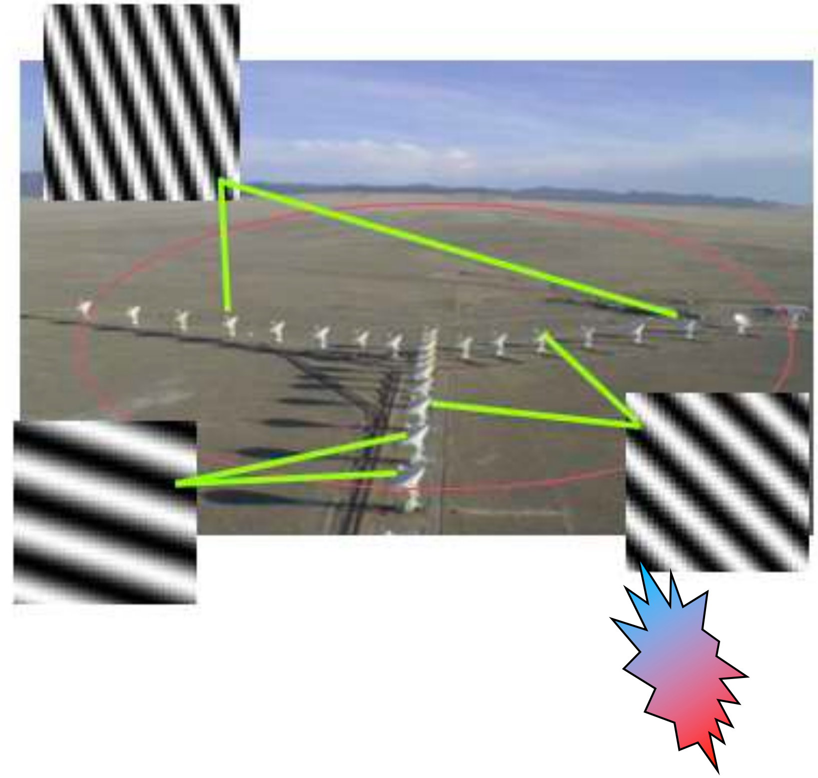
26 September 2017

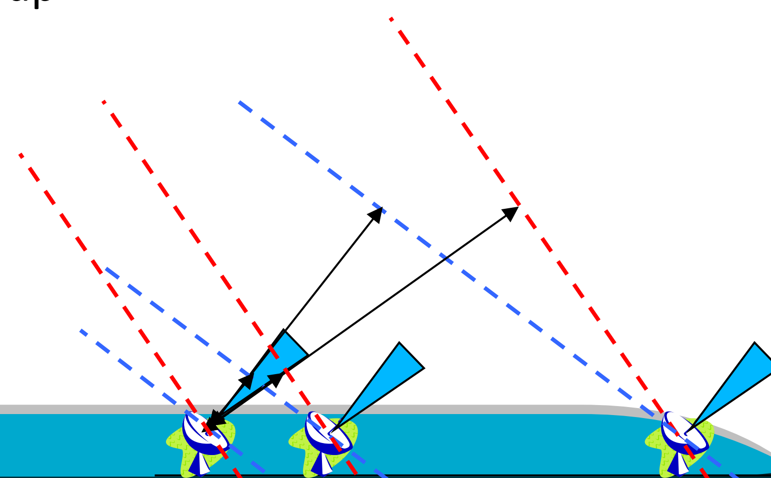
Topics

- Interferometry and Fourier basics
- Visibilities and the uv-plane
- Creating Images
  - gridding, weighting, resolution, fov, mfs, cubes
- Deconvolution Methods
  - clean, mem

Interferometry

- Long baseline
  - Delay variation of many wavelengths across field,
  - narrow fringe pattern,
  - extended sources average out
  - high resolution
- Short baseline
  - Small delay variation across field
  - wide fringe pattern,
  - flux from extended sources adds up
  - low resolution





Response of single interferometer

$$\langle E_i E_j^* \rangle \propto V_{ij}(u,v) = \iint I^{sky}(l,m)\, e^{2\pi i (ul+vm)}\, dl dm$$

V(u,v) – complex visibility = coherence between average electric fields across FOV of two antennas (i and j) separated by a baseline (u,v) = Fourier transform of sky brightness

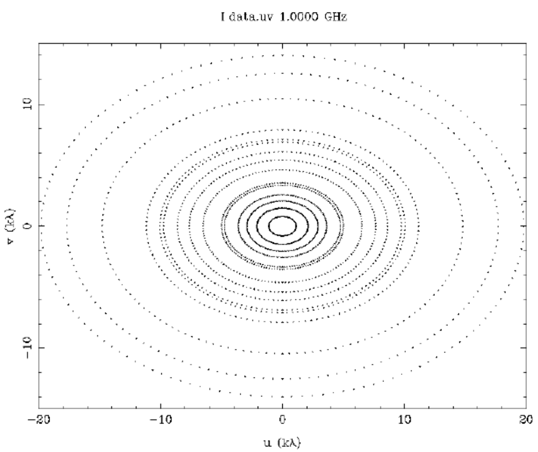
- Baseline coordinates (u,v) – also called **spatial frequencies**, in E-W and N-S direction
  - attached to the earth: rotates with time relative to sky
  - measured in wavelengths: changes with frequency
- ***I***(l,m) – sky brightness distribution
  - l,m coordinates on the sky, tangent plane projections of RA and Dec

Single u,v point measures response at one spatial frequency in one direction

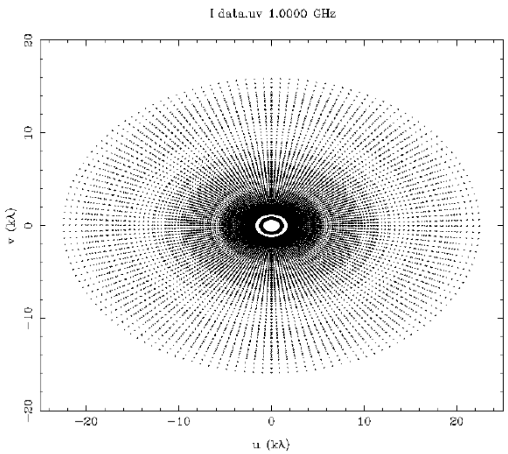
For a good image of the sky we try to fill the **uv plane** with measurements:

- Use as many antennas, times and frequencies as we can to achieve this

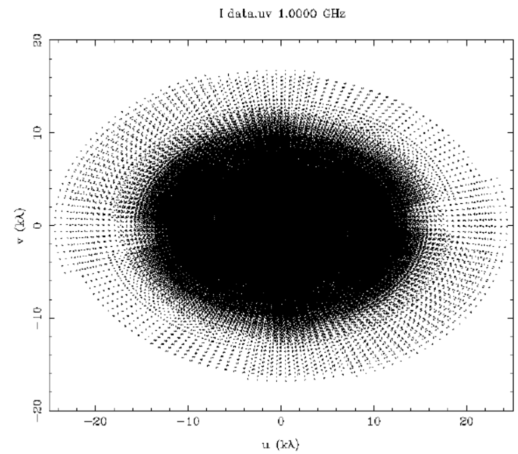
Filling the uv plane



Earth rotation synthesis
ATCA 1 to 12 h
6km array

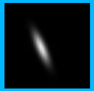


Multi frequency synthesis
ATCA 1 to 10 channels
30% bandwidth
6km array



Adding more antennas
ASKAP 2 to 36 antennas
6h, 30% bandwidth

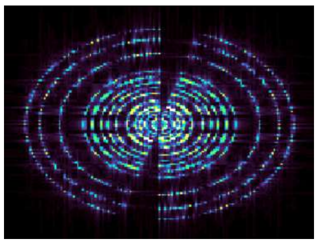
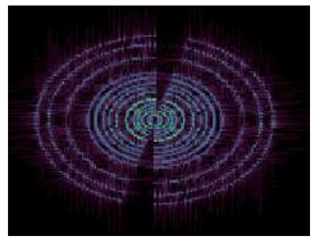
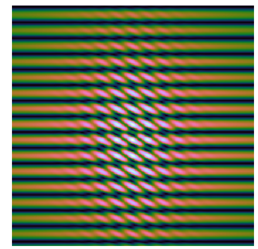
Fourier Basics

Image	FT	uv plane
$f(l,m)$		$F(u,v)$
Large		Small (brighter)
Constant		$\delta(u,v)$ "zero spacing"
Real		Hermitian $F(u,v)=F(-u,-v)^*$
Multiply $f\,g$		Convolve $F*G$
Shift $f(l+a,m)$		Phase Gradient $F(u,v)e^{2\pi iau}$
Add $f+g$		Add $(F+G)$
Rotate		Rotate
Gaussian 		Gaussian 

Imaging

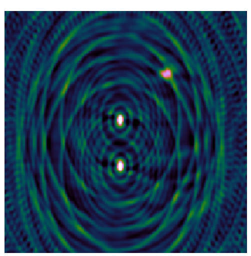
We observe the visibility V for the positions (u,v) where we have measurements:
the uv coverage is a sampling function

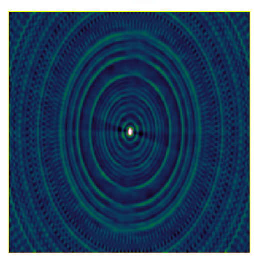
$$V^{obs}(u,v) = S(u,v) \cdot V(u,v)$$

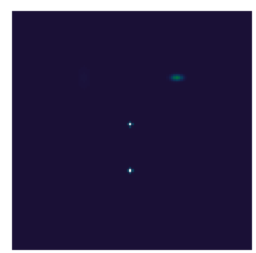




Taking the Fourier transform, this corresponds to:

$$I^{obs}(l,m) = I^{PSF}(l,m) * I^{sky}(l,m)$$

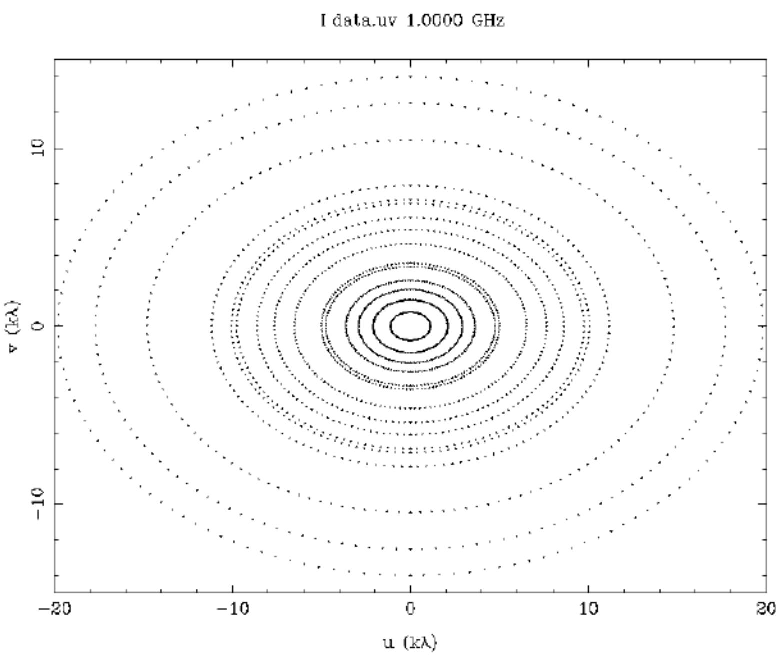

'Dirty image'

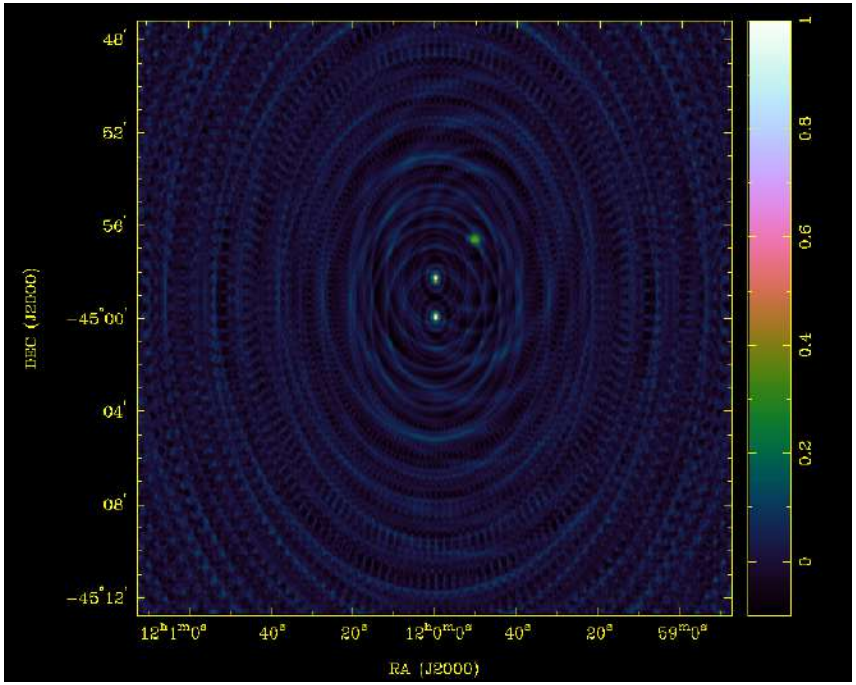

'Dirty beam'



Earth rotation synthesis

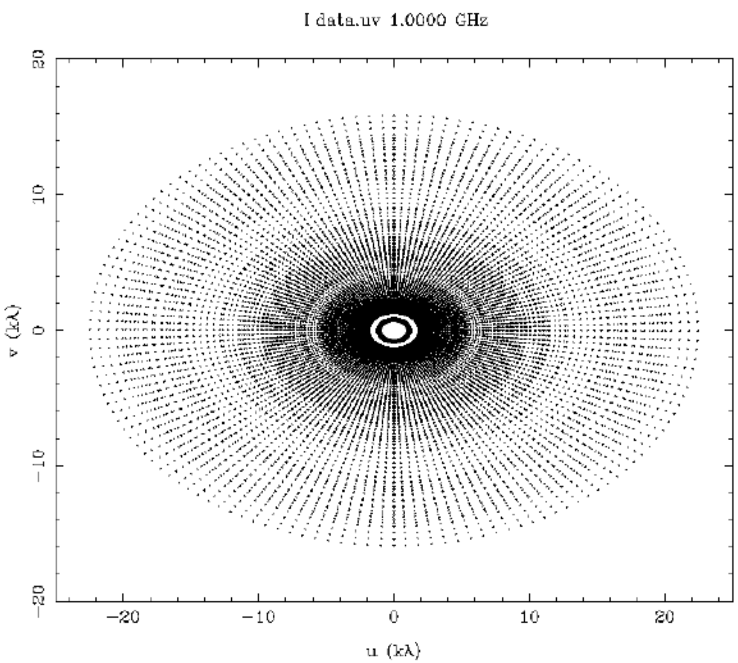
integrate in time: 1 to 12 h, ATCA 6km array

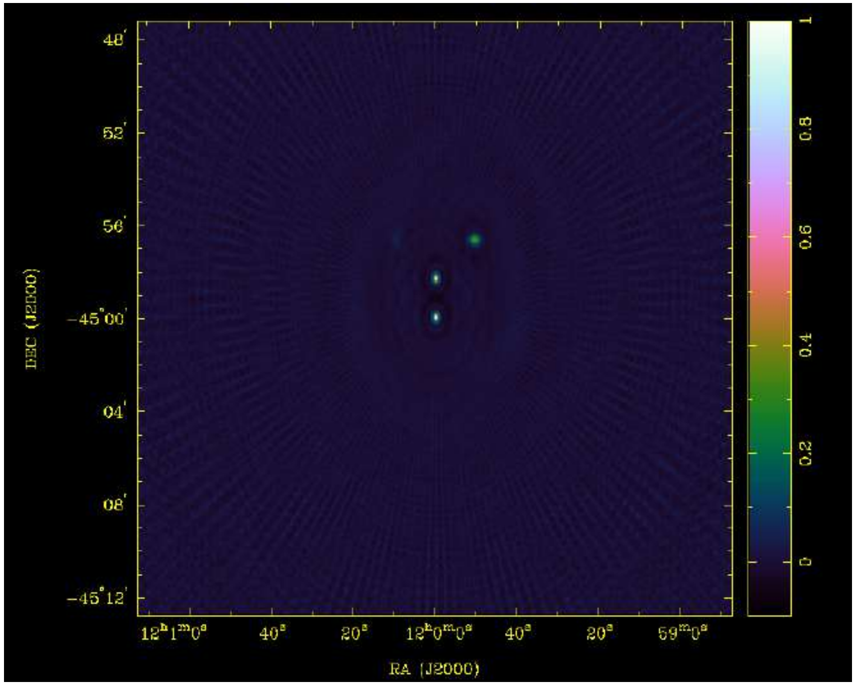




Multi-frequency synthesis

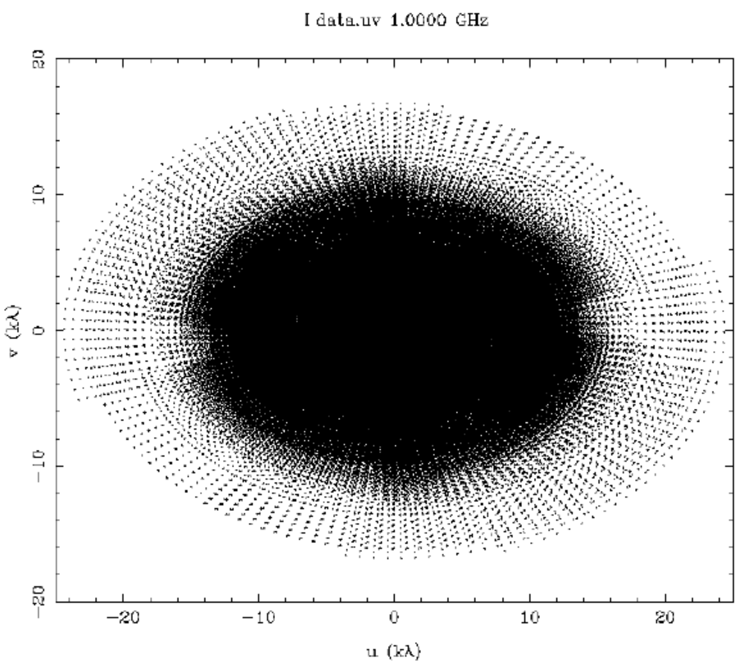
integrate in frequency: 1 to 10 channels over 30% bandwidth, 12h

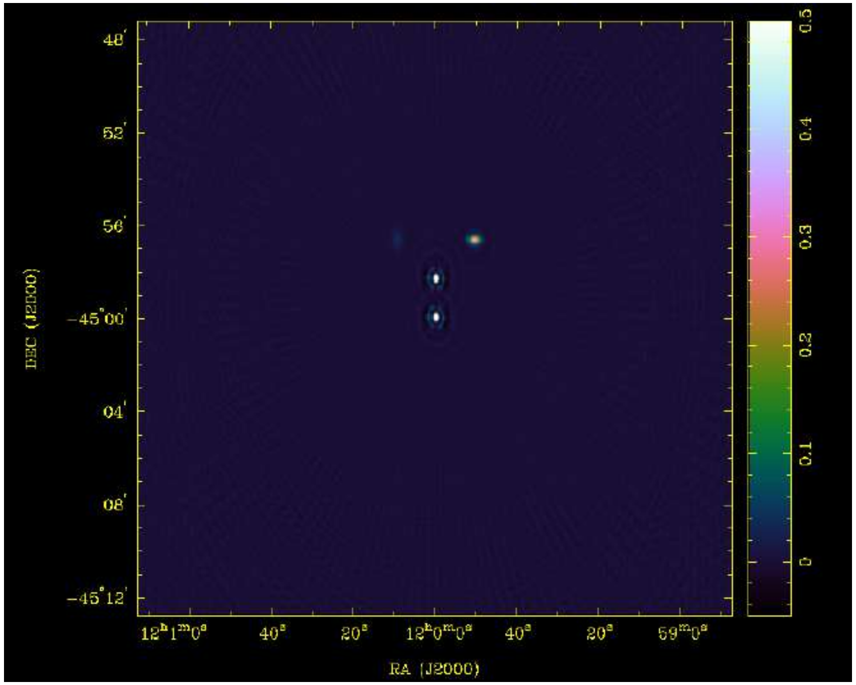




Adding more antennas

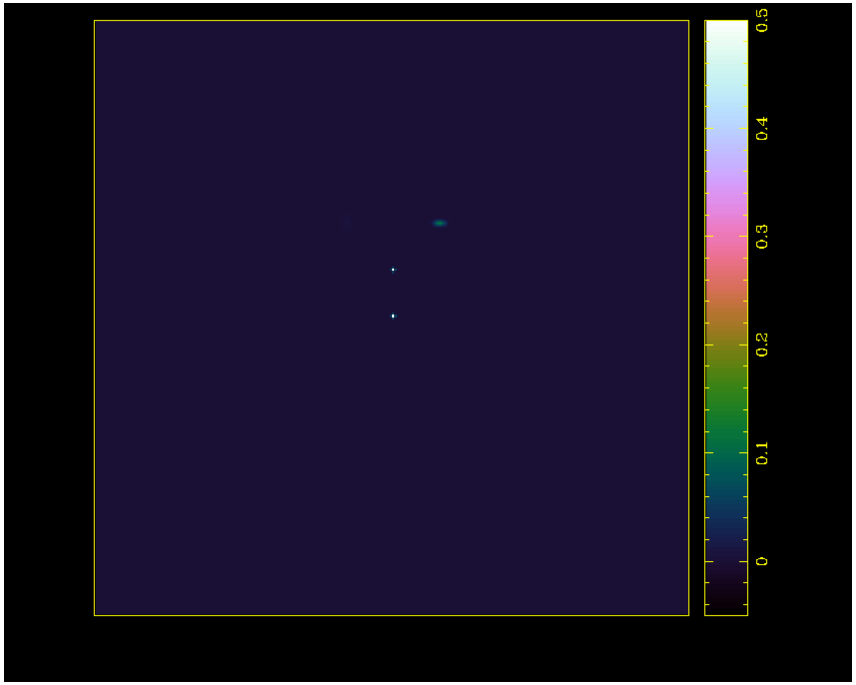
ASKAP 2→36 antennas, 6h, 10 freq (30% bw)



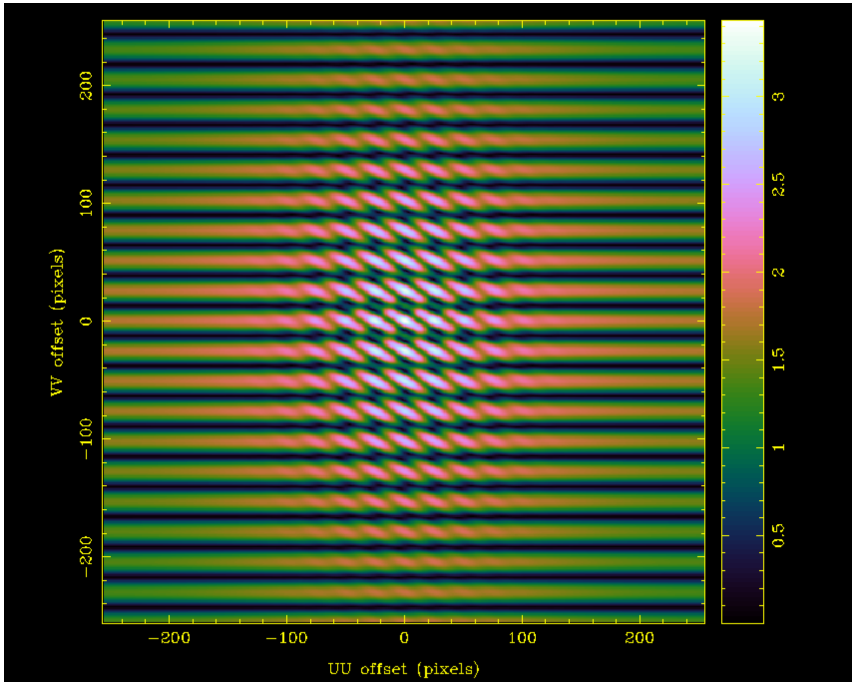


Fourier pairs in radio astronomy

Sky

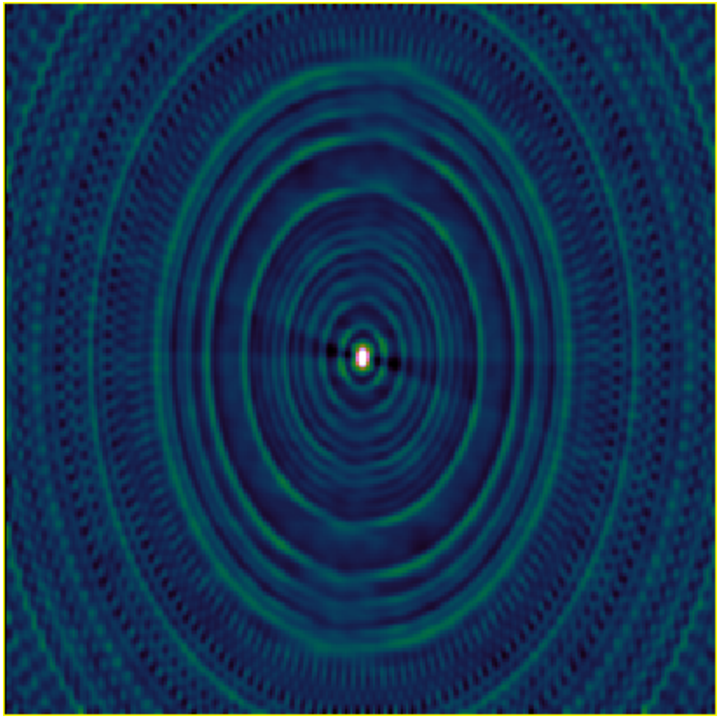


uv-plane



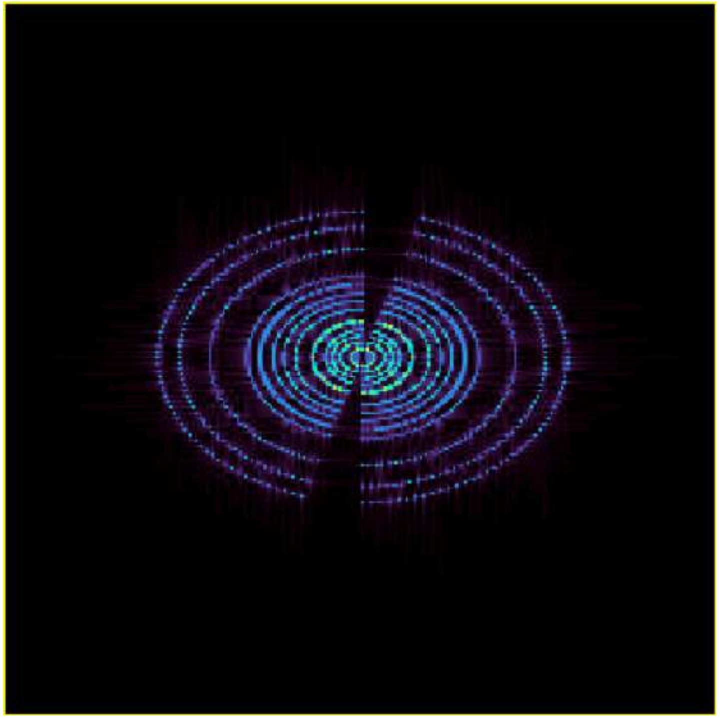
Fourier pairs in radio astronomy

Dirty beam, PSF



Peak of PSF is 1
Surrounded by positive and negative sidelobes

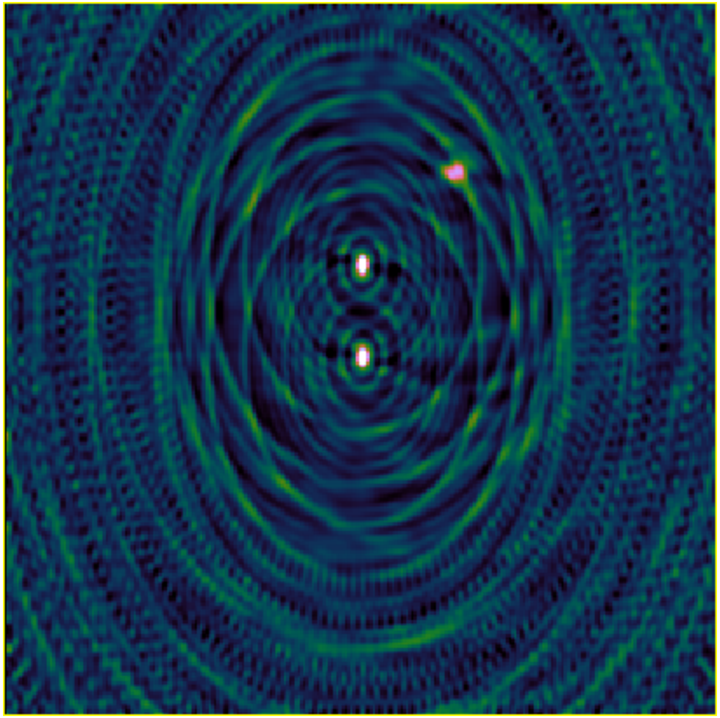
Gridded uv-coverage



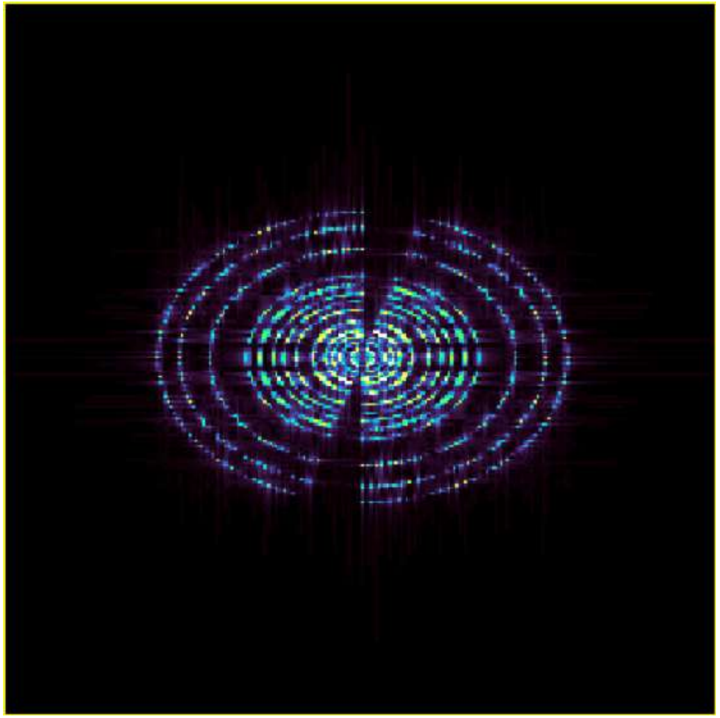
Points in uv plane gridded to nearest pixels using convolution with ‘prolate spheroidal’ function – suppresses aliasing

Fourier pairs in radio astronomy

Dirty Image



Gridded visibilities

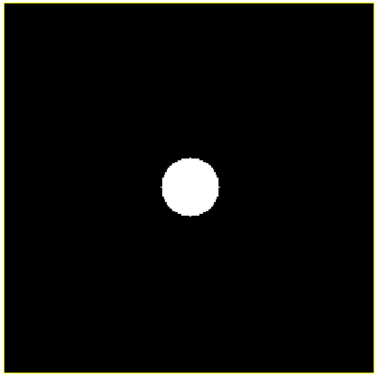


Each source surrounded by positive and negative sidelobes

Points in uv plane gridded to nearest pixels using convolution with ‘prolate spheroidal’ function

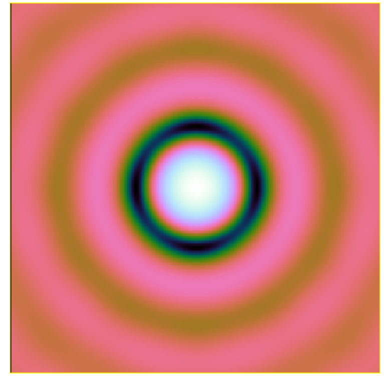
Fourier pairs in radio astronomy

Dish illumination pattern

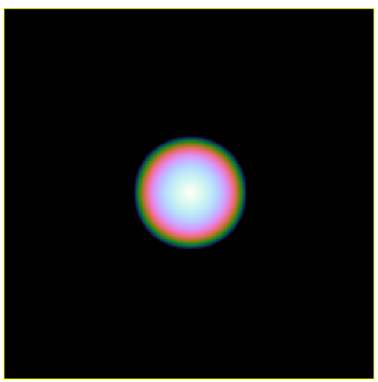


FT

Voltage beam

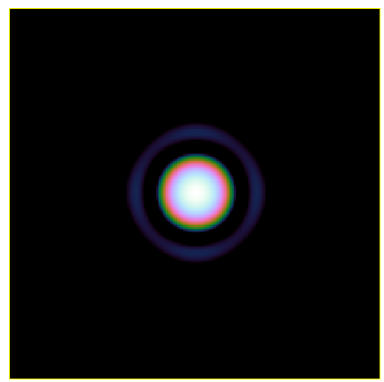


Cross corr. of ill. pattern



FT

Primary Beam



PB=VBxVB*

UV plane features

- Dense tracks
  - Baselines – tracks in *(u,v)* plane
- Low level in between tracks
  - Gaps in coverage, missing information
- Hole in center
  - No information on low spatial frequencies, i.e., largest scale structures filtered out
- Outer boundary
  - No info on small scale structure – resolution limit
- Ways to Fill the UV Plane
  - Centre: add single dish data, mosaicing
  - Gaps:
    - multiple configurations (750m,1.5km,3km)
    - Multi-frequency synthesis
    - Time/Earth rotation

ATCA

ASKAP

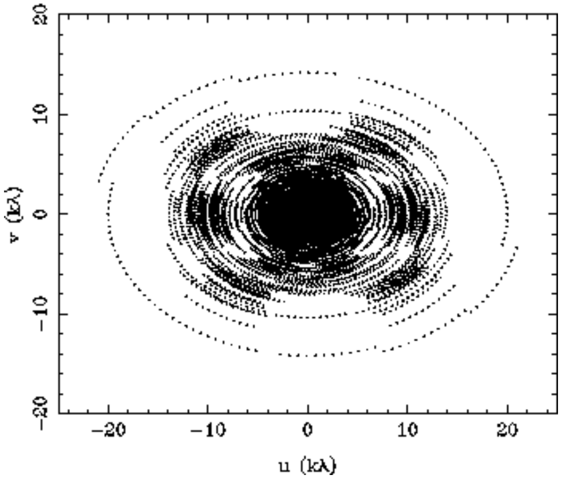


Image size and resolution

Some decisions best made at proposal or observing stage, some at imaging stage

Field of view (FOV, λ/D) – Image size

- Determined by primary beam size, more with mosaicing (multiple fields)
  - ATCA 21cm – 33' beam, 3mm – 30'' beam
  - ASKAP ~1 degree FOV (x36 beams: ~25-30 sq deg for combined image)
  - may need to image larger field to remove sidelobes from distant sources
    - image to first or second sidelobe at 1 GHz (or whole sky for MWA)
- Shortest baseline determines largest structure we can image well (λ/Bmin)

Resolution (λ/Bmax), pixel (or cell) size

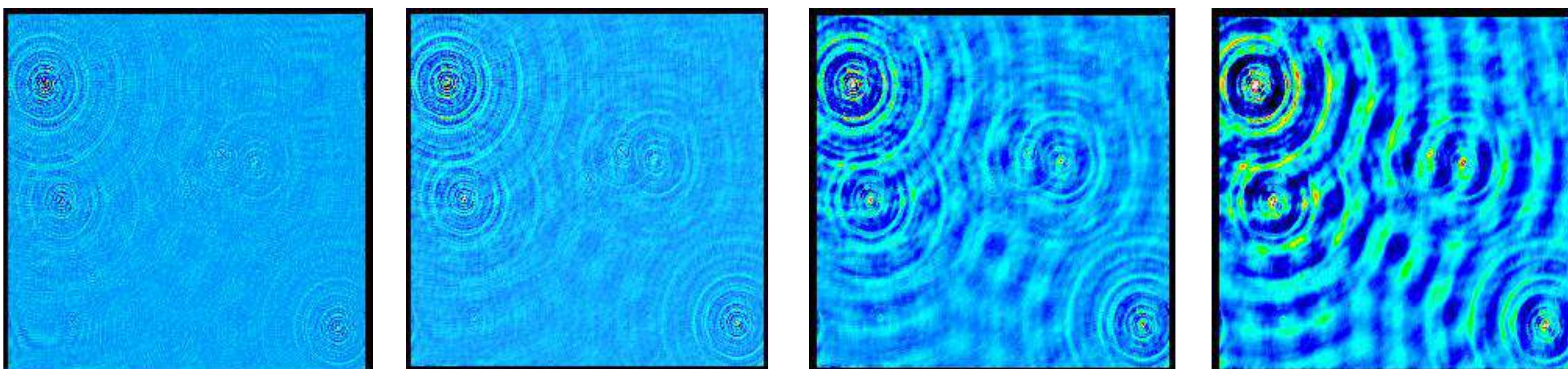
- Need >2 pixels/beam for Nyquist sampling – use more (3-5) for better deconvolution
- Longest baseline determines limit to resolution (ATCA 21cm, ASKAP36: 6km– 7”)
- For small sources use high resolution configurations – need many pixels for large FOV
- For extended, low surface brightness sources use compact configurations
  - e.g., EW352, H75 and exclude long baselines (CA06)
  - This lets you use larger pixels for the same FOV and greatly improve surface brightness sensitivity

Weighting schemes

- Uniform (lowest sidelobe level) – equal weight for all filled uv cells
- Natural (lowest noise level) – equal weight for all visibilities
- Robust/Briggs - optimal combination of above two ← use these
- Taper (decrease resolution) – weight=gaussian(u,v)
- S/N Weighting – weight=$1/\sigma^2$ –σ can depend on time/baseline/freq

Tapering and S/N weighting can be combined with the others

	Uniform R=-2	Robust=0.5	Robust=1.0	Natural R=2
Beam:	7x5”	8x6.5''	9.6x7.5''	12x8''
Sens.:	1.45	1.16	1.06	1.0



Weighting schemes

Choose weighting based on science goals

- Point source detection experiment: Natural with S/N weighting
- Easy imaging of bright object: Uniform, no S/N weighting
- Low surface brightness, extended: Tapering, Robust
- Busy field: Robust (0-1) – good sensitivity with low sidelobes

	Robust/Uniform	Natural	Taper
resolution	higher	medium	lower
sidelobes	lower	higher	depends
point source sensitivity	lower	maximum	lower
extended source sensitivity	lower	medium	higher

Continuum vs Spectral line imaging

- Continuum
  - Create single output image at 'average' frequency
  - wide bandwidth – 2 GHz (ATCA) , 300 MHz (ASKAP)
  - combine channels (ATCA continuum obs. Has 2048 x 1MHz channels)
  - possibly combine multiple centre frequencies (MFS)
    - E.g., for max. sensitivity at 4cm use 5.5 & 7.5 GHz, each with 2048 channels over 2GHz.

- Line – image one or more spectral lines
  - ATCA can do 32 simultaneous 'zooms' – 1 or 16 MHz bands with 2048 channels
  - ASKAP has 16000 channels over 300 MHz
  - Creates output image cube (2048^3=8 Gpixel)– image for each frequency channel (RA,Dec,Vel)
  - Check velocity frame(lsr/bary), Doppler correction
  - Specify spectral resolution & velocity range
  - May need to remove continuum emission first (uvlin, uvmodel)
  - Cannot use MFS to improve uv coverage
    - uv coverage poorer and noise higher (small bw)
    - results in lower DR images

Beyond the dirty image…

- Calibration, weighting and Fourier transform get us to the best possible dirty image
- Infinite number of ‘invisible distributions’ between our uv points
- To improve things further we want to:
  - Remove the sidelobes of the dirty beam from our image (clean…)
- Dirty Image = Sky convolved with dirty beam
  - Need a deconvolution procedure
    - Linear?
    - Non-linear?

Linear deconvolution …

$$I_{\mathrm{D}} = B * I$$

$$\mathcal{F}[I] = \mathcal{F}[I_{\mathrm{D}}] / \mathcal{F}[B]$$

- Noise properties are well understood
- Generally non-iterative and computationally cheap

But

- It does a very poor job in interferometry
  - basically the same as uniform weighting!
- Linear deconvolution is fundamentally unable to extrapolate to unmeasured spatial frequencies.

Non Linear Deconvolution

- A good non-linear deconvolution algorithm is one that picks plausible ('invisible') distributions to fill the unmeasured parts of the uv plane.
  - Avoids too much extrapolation (stick to resolution limit)
- Need to make assumptions to get a realistic estimate
  - Main assumption: Real sky does NOT look like typical dirty beam
    - Rings, spokes, negative regions, infinite extent: all very unlikely
- Different algorithms make different assumptions:
  - CLEAN (pixel based), Point Source Fitting
    - Sky is mostly empty, with occasional peaks
  - Scale-sensitive algorithms: multi-scale clean, Gaussian Source fitting
    - Sky consists of bounded, overlapping, regions of emission
  - MEM (Maximum Entropy Method) (pixel based)
    - Sky is uniform (& positive) – favours smooth distributions of flux
  - MORESANE/IUWT – Combination of multi-scale and MEM type approaches
    - Sky decomposed into wavelets, with conj. gradient deconvolution
    - Available in wsclean

CLEAN

- Original version by Högbom (1974)
  - Purely image based – subtract scaled beam from image (minor iterations only)
- Clark
  - Use small patch of beam, subtract model from gridded visibilities using FFT (major/minor iterations)
- Cotton-Schwab
  - Like Clark, but subtract from ungridded visibilities and re-image residuals
- SDI (Steer/Dewdney/Ito)
  - try to cope with extended emission by subtracting patch of pixels above ~90% of peak (avoids 'corrugations')
- Basic assumption: Highest peak in dirty image is real

Restoring the image

- After deconvolution we are left with a residual image
  - Noise
  - Weak source structure below the CLEAN cutoff limit
    – 5σ (large area), 2-3σ (clean boxes)
  - Sidelobes of faint and extended sources
- Restored Image
  - Take residual image
  - Add point components convolved with gaussian fit to central peak of dirty beam
  - Resulting image is best guess of real sky with measurement noise
  - Avoids ‘super-resolution’ of component model

CLEAN Algorithms

1. Make image
   a. grid/weight visibilities ➔ gridded visibilities
   b. fft gridded visibilities ➔ (residual) dirty image
2. Find position of highest peak in image
3. Subtract a fraction of this peak (‘gain factor’) using a scaled dirty beam(patch) at this position (usually gain~0.1)
4. Add model component to list or build up model image
5. Go to 2, unless prescribed flux limit or iteration limit reached
6. If doing major iterations
   a. subtract current model from gridded or ungridded visibilities and
   b. goto 1 (Clark 1b, CS 1a) unless final flux or iteration limit reached
7. Restore model convolved with gaussian beam to residual image

major

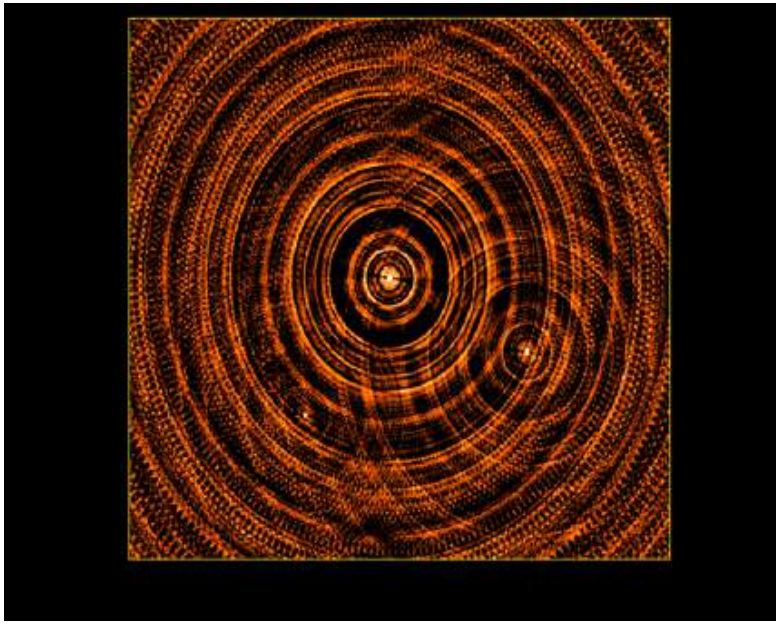
minor

CLEAN example 1

Model: 5 point sources + 1 gaussian

Point: 1,0.5,0.25,0.1,0.01 Jy

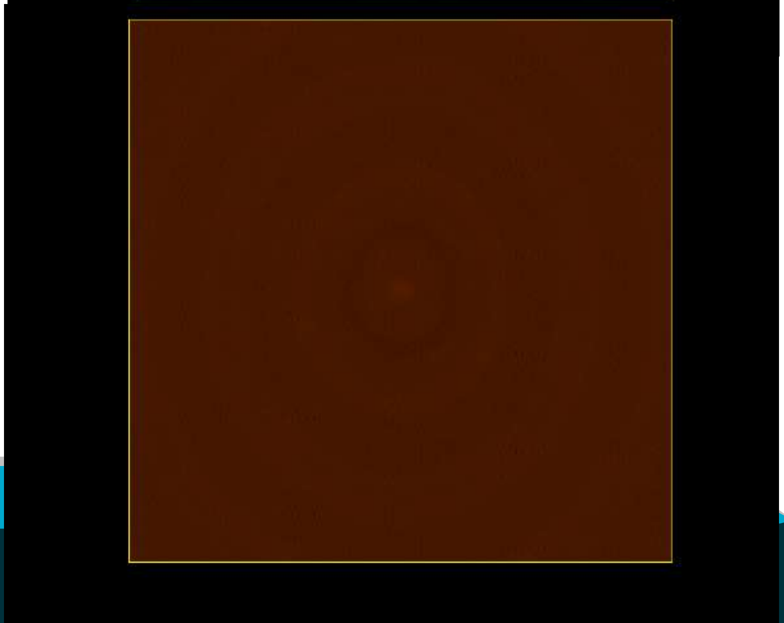
Gaussian: 0.1Jy, 10”x10”



Dirty image

Restored

1,5,10,20,50, 100,200,500, 1000 CC’s

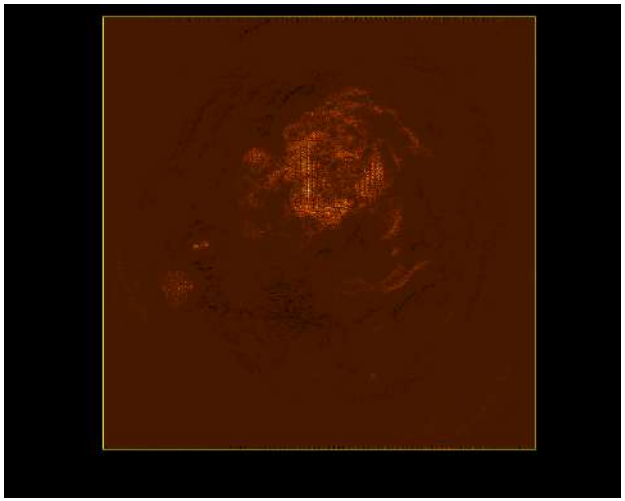


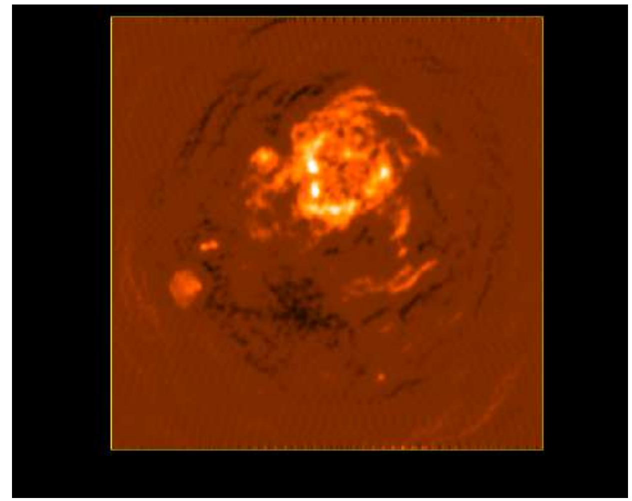
Residual

CLEAN example 2

Clean Model

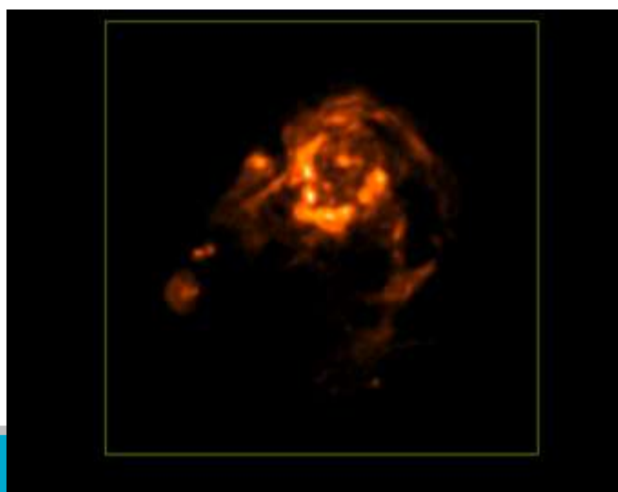
1, 10, 100, 1000, 10000, 100000 CC's

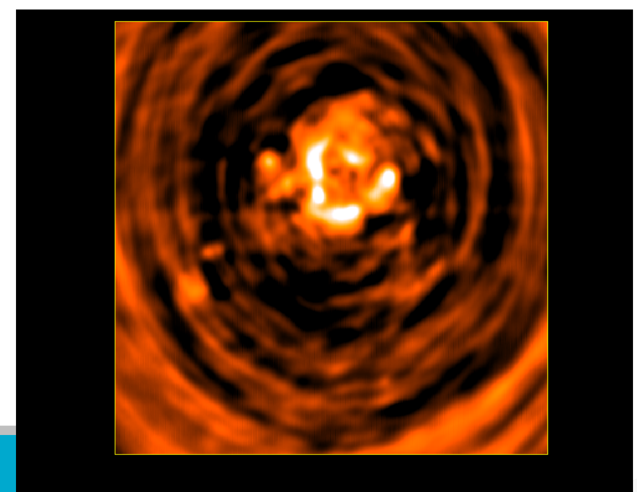




Restored Image

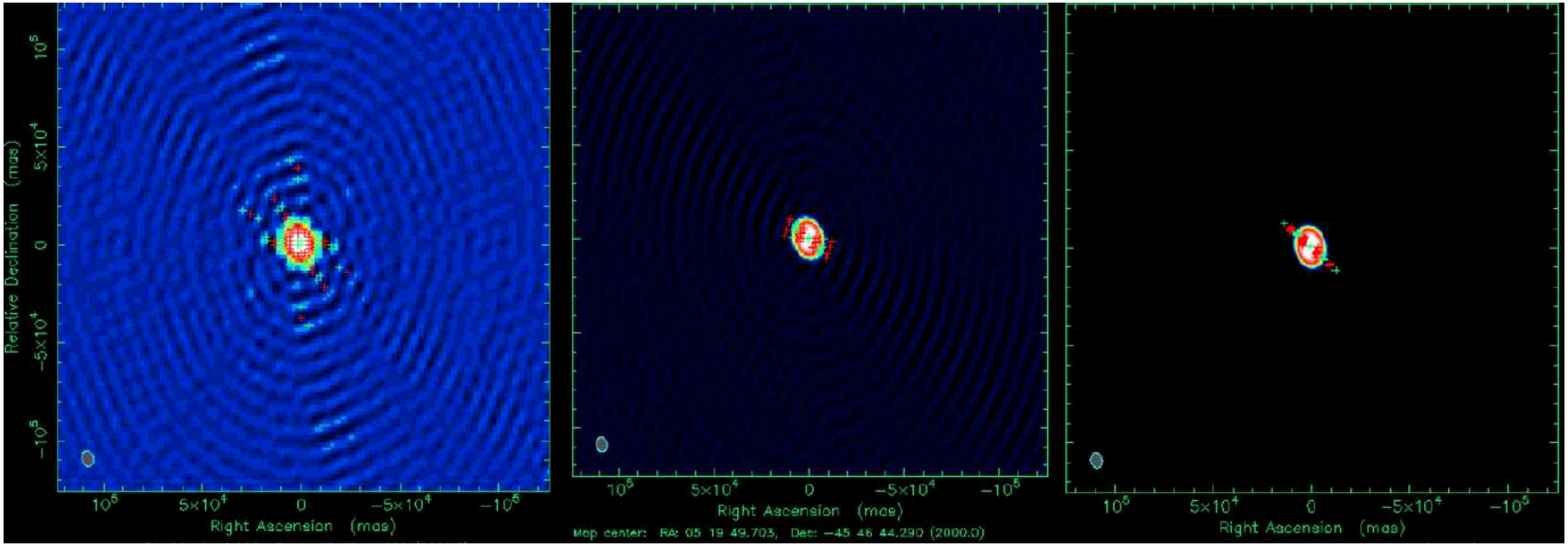
True Sky





Dirty Image

Clean – effect of cell size



3 pixels/beam | 6 pixels/beam | 12 pixels/beam

A strong point source between pixels takes a lot of positive and negative components to clean and causes Dynamic Range problems (1000:1)
Fix: put the source on a pixel or (if >2) use smaller pixels to limit size of errors

CLEAN Strengths & weaknesses

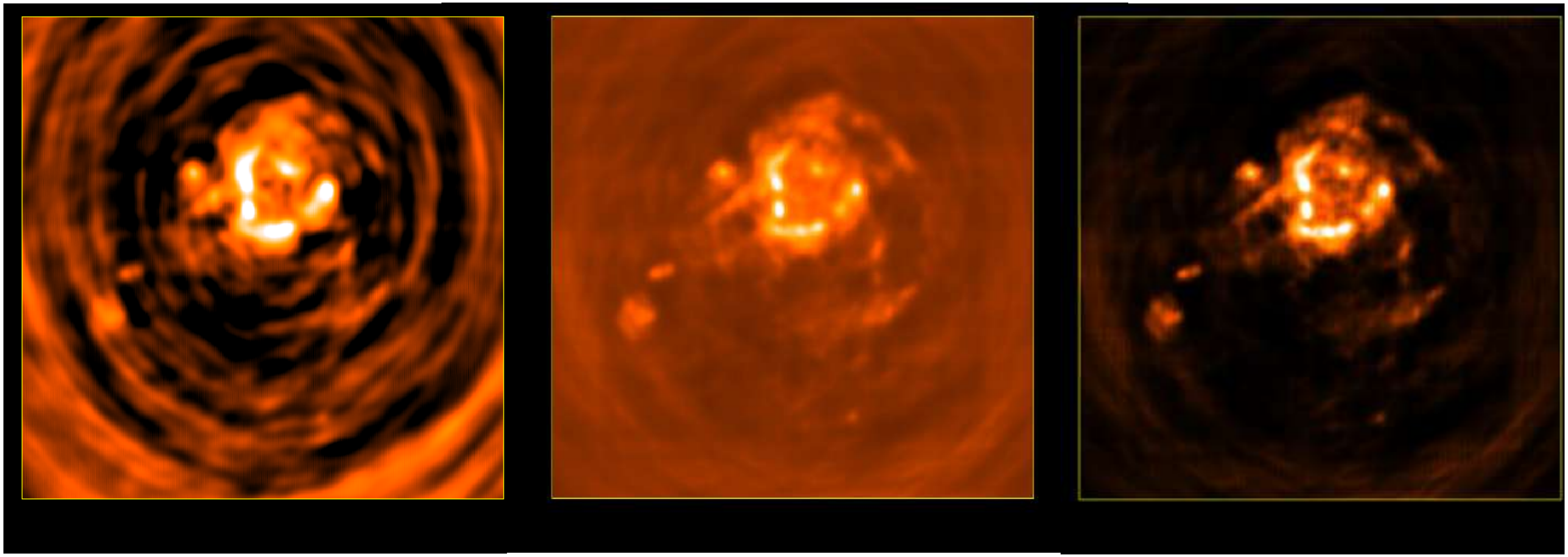
- CLEAN works well for fields with many compact sources
- Effect of defects (RFI, source variability, gain errors) is generally very local (unlike MEM)
- Algorithm can be generalized
  - Deal with wide bandwidth
    - MFS Taylor Term imaging
  - Deal with extended emission
    - using Multiple Scales

- Standard CLEAN works poorly for very extended objects:
  - Slow (too many faint point components needed)
  - Corrugation instability
  - Negative bowl effect – large source surrounded by negative pixel
- CLEAN Bias (`overcleaning')
  - artificially low noise, wrong fluxes
  - Affects snapshot/mosaic/line observations with sparse uv coverage
  - Fix: clean boxes, fewer iterations, joint deconvolution
- Clark CLEAN can diverge
  - Poor beam, large # iterations
  - Fix: extend size of beam patch, use Högbom Clean instead or do major iteration sooner

MEM

- MEM – Maximum Entropy Method
  - Tries to find the 'smoothest' image that is consistent with the data
  - Image with lowest 'information content' for given total flux
  - No data → flat image
- Define smoothness via the 'entropy' H
  - $H = -\sum_k I_k \ln(I_k/M_k e)$, I_k=pixel k in the image
  - Use of logarithm enforces positivity constraint
    - negative sidelobe suppression
- M_k is the prior image - a flat default image can be used, but a good low resolution image, if available, is better
- Data constraints are added via $\chi^2=\sum_i(V_i^{obs} - V_i^{mod})^2/\sigma_i^2$

MEM



Dirty Image MEM restored image Final MEM model

25 iterations of MEM

Final image = Converged solution

MEM Strengths/Weaknesses

- In MEM it is easy to add multiple constraints
  - E.g. information from overlapping fields – mosaicing
  - Single dish image added to interferometer data
- Works well for extended images
  - Can be much faster than CLEAN for extended structure
  - Tends to fail for point sources embedded in extended emission
- More sensitive to data defects (calibration problems etc)
  - Effect of errors not localized, may affect convergence

Scale sensitive methods

- Both CLEAN and MEM work on single pixels
  - No inherent notion of source size
  - When we look at an image we identify a collection of sources of different sizes – makes physical sense too
  - Adjacent pixels in an image are NOT independent
    - Resolution limit
    - Intrinsic source size
    - E.g., a gaussian source covering 1000 pixels can be represented by only 6 parameters instead of 1000.
- Scale sensitive algorithms try to capture this extra information about a ‘plausible sky’
  - Reduces number of degrees of freedom in solution
  - Separation of signal and noise easier

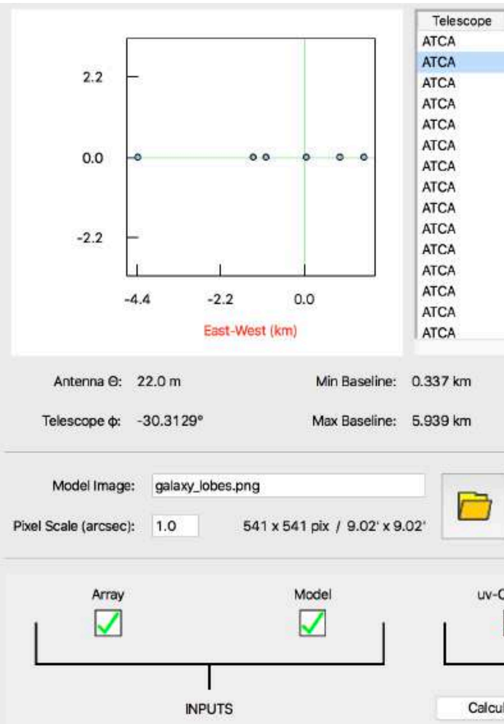
Scale sensitive methods

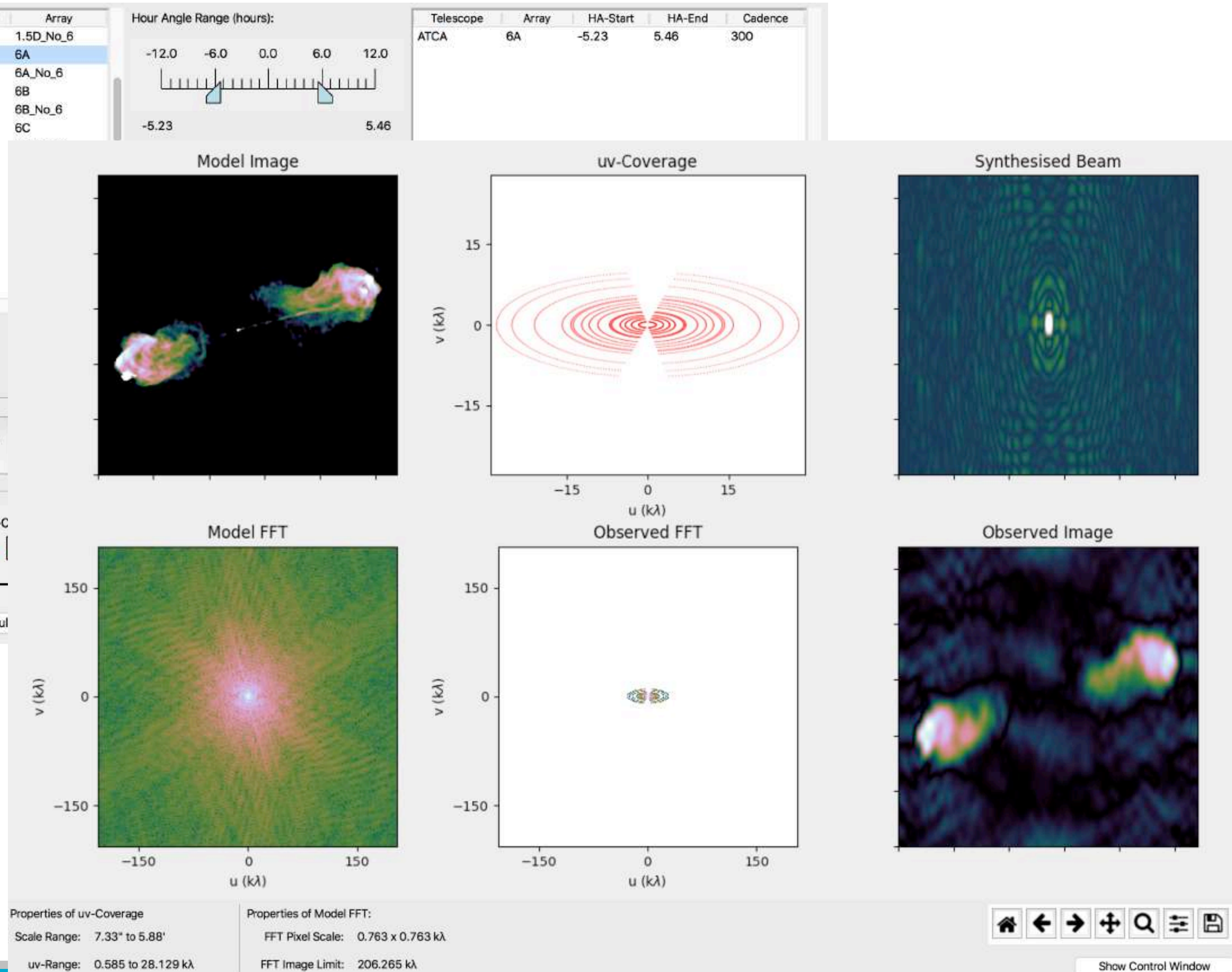
- SDI Clean (Steer-Dewdney-Ito)
  - One of the early attempts to make CLEAN cope with extended structure
  - Subtracts scaled beam from a patch of pixels around each peak found
- Multi-resolution CLEAN (Wakker-Schwarz)
  - Make images at 2-4 different resolutions, clean lowest resolution first, then clean residuals at higher resolution
  - Combined model sensitive to all scales with greatly decreased number of iterations
- Multi-scale CLEAN (Cornwell-Holdaway)
  - Similar, but cleans all scales simultaneously – more robust
  - Find peak across all images
  - Remove fraction of peak at that scale from all images
  - Add corresponding ‘blob’ to model
  - Iterate until we reach the noise level in all images

Miriad & CASA

	Miriad	CASA – tclean task
Clean algorithms	clean	tclean (newer) and clean task
Joint Pol. clean	cclean (Q&U)	deconvolver='clark' (IQUV)
Multi-frequency Clean	mfclean (I, Iα)	nterms=2,3,4..
Maximum entropy	maxen	deconvolver='mem'
Mosaicing (max entropy, clean)	mosmem mossdi	tclean gridder='mosaic'
Joint polarimetric MEM (single pointing or mosaicing)	pmosmem	-
W-term or 2d beam	No	gridder='wproject'/'awproject'

Friendly Virtual Radio Interferometer





Great tool to get familiar with imaging and Fourier Transforms

See Tutorial (Cormac Purcell)

Acknowledgements

- Previous Imaging talks by Emil Lenc, Urvashi Rau, David Wilner

References

- Thompson, A.R., Moran, J.M. & Swenson, G. W. 2017, "Interferometry and Synthesis in Radio Astronomy" 3rd edition (OPEN ACCESS ebook)
- NRAO lectures: http://www.aoc.nrao.edu/events/synthesis/