Wide field imaging - computational challenges

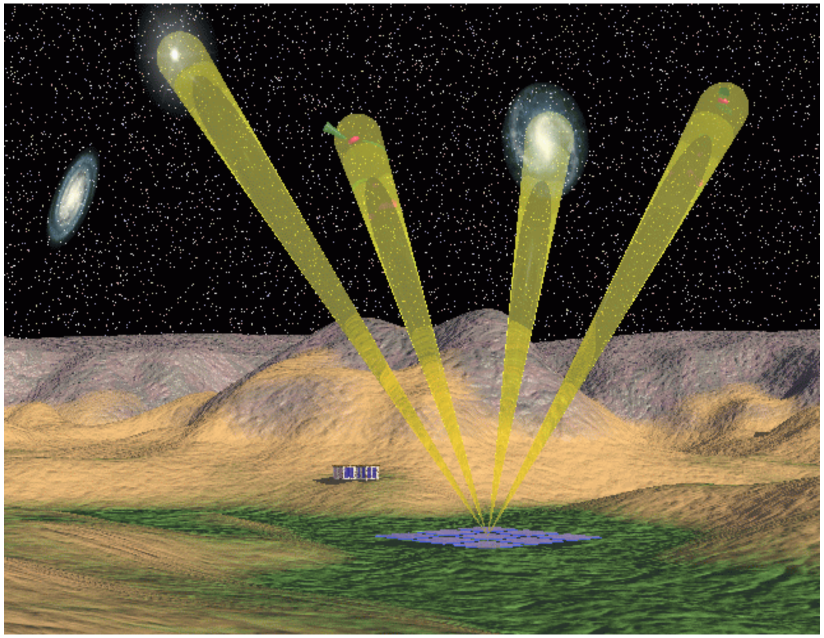
Tim Cornwell

Kumar Golap

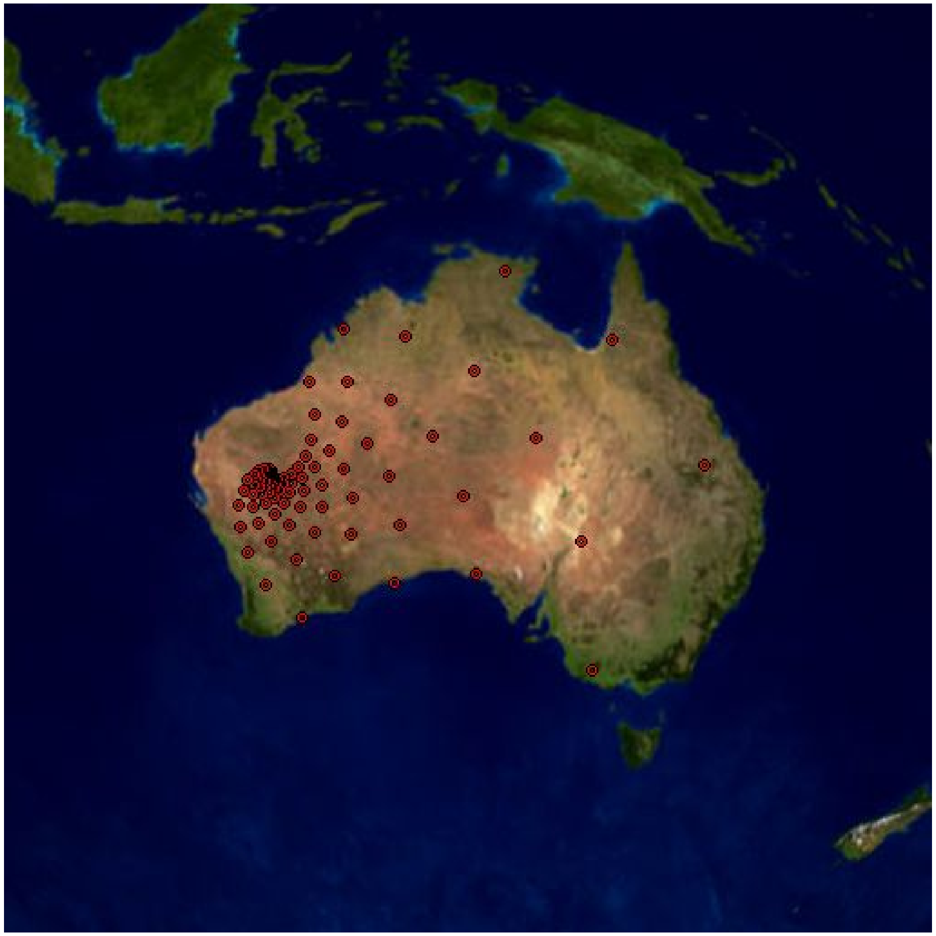
Sanjay Bhatnagar

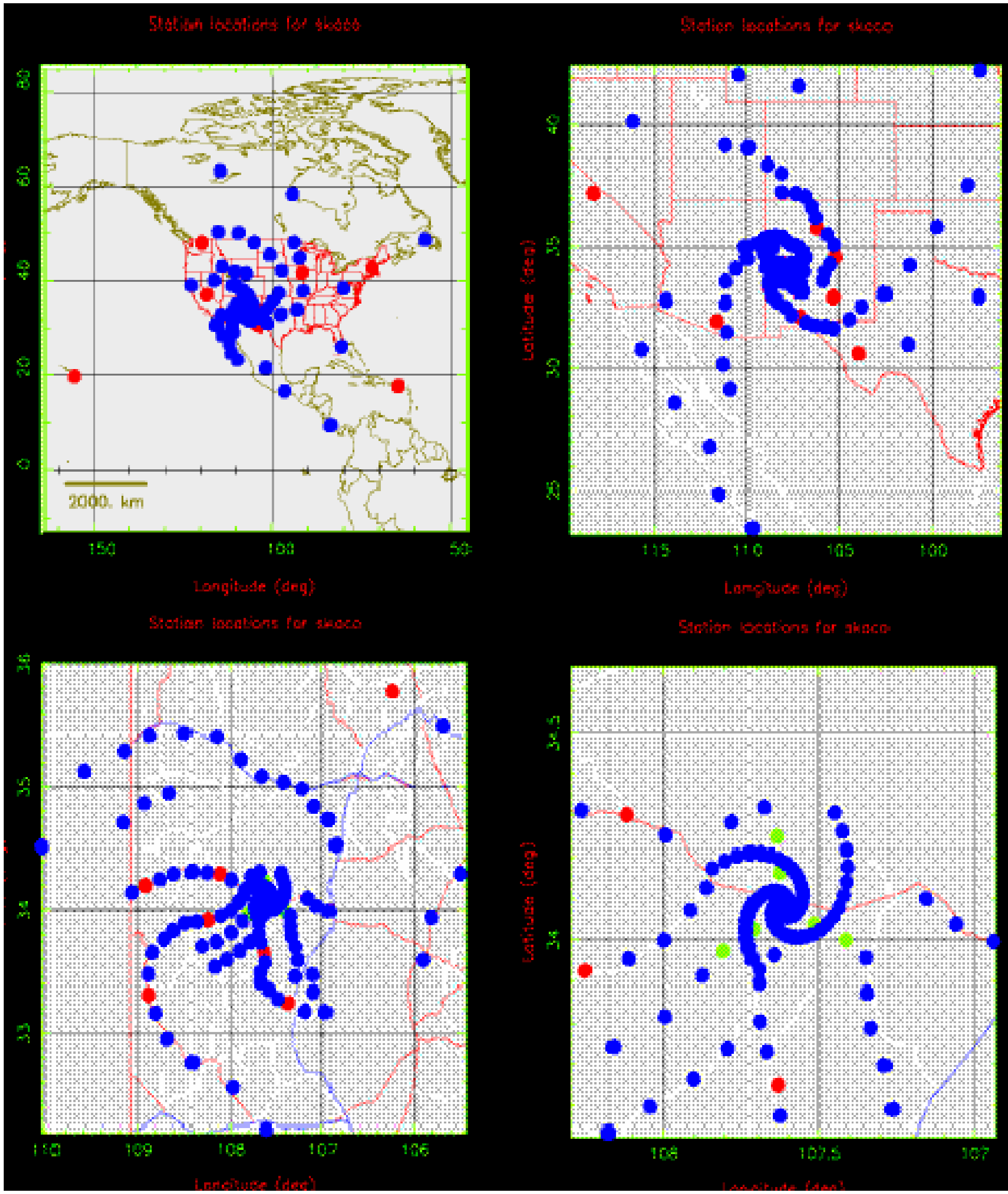
Square Kilometre Array

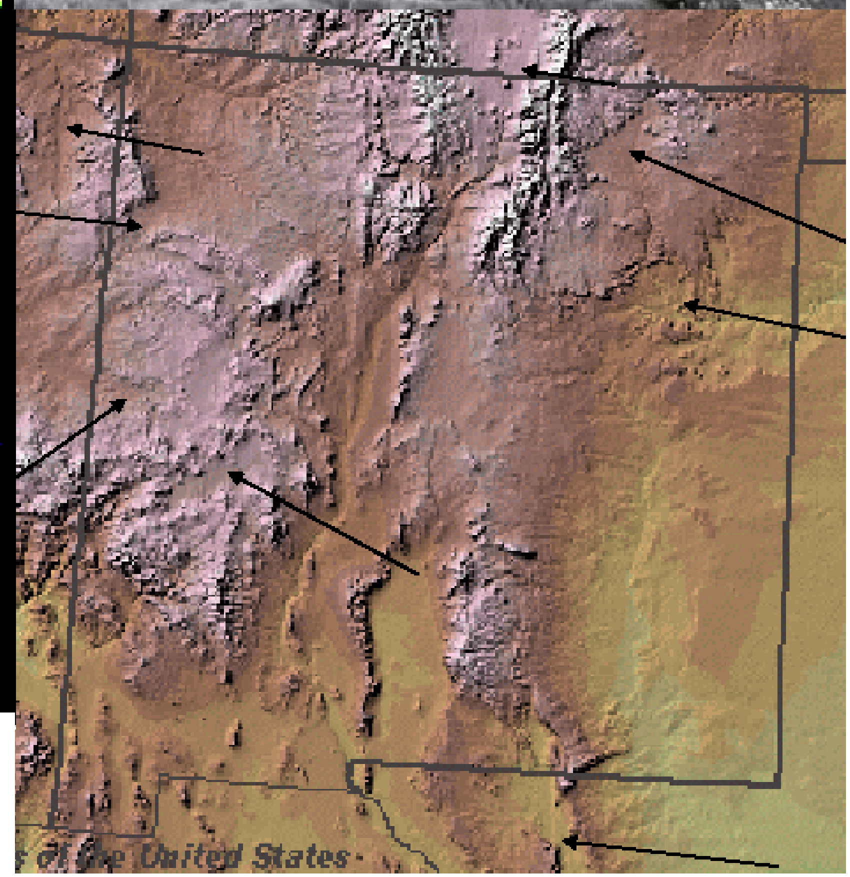
- Next generation radio telescope
  - ~ 50 x EVLA
- Baselines up to 3000km
- Few hundred stations on baselines form 150km to 3000km
- Frequency 0.1 - 25GHz
- Many challenges in calibration and imaging



Possible configuration in Australia

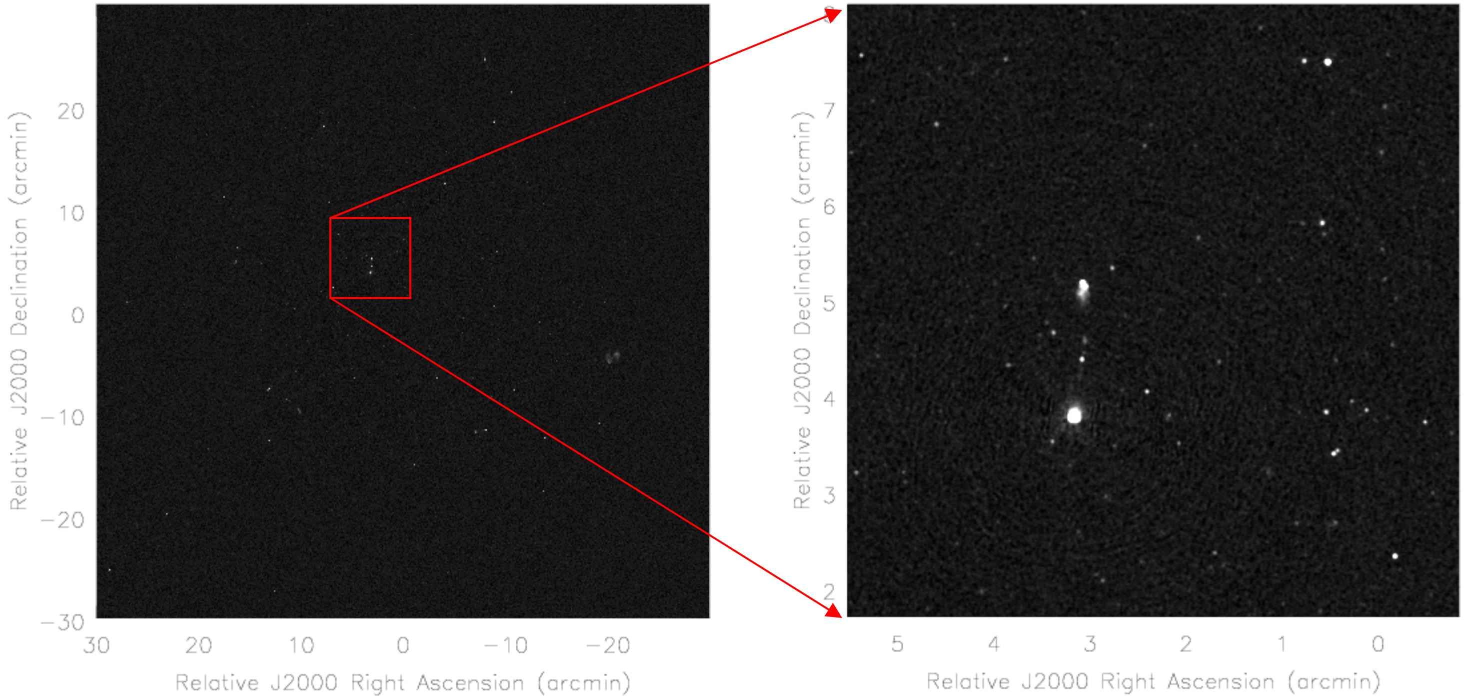






SKA in North America

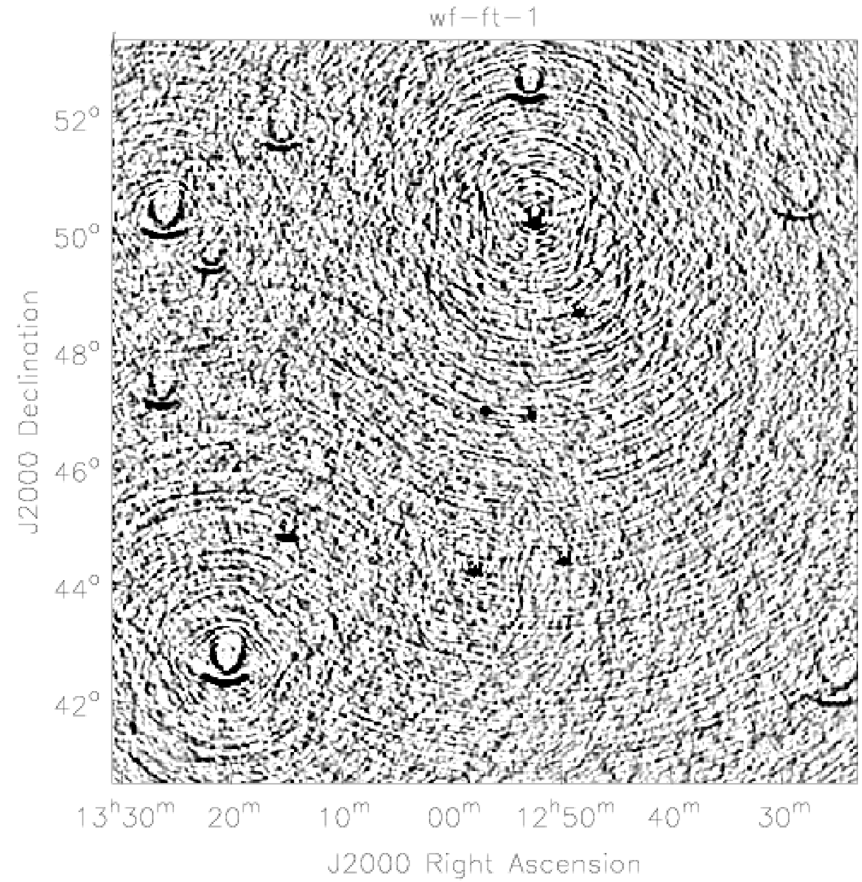
3μJy/beam 1.5arcsec 1.4GHz VLA image



- Square Kilometre Array will be ~ 100 times deeper
- Confusion limits probably require 0.1arcsec resolution
- To reach sensitivity limit, must image accurately all emission over 1 degree FOV
- Image sizes could be up to 80,000 by 80,000 pixels

A problem….

- Point sources away from the phase center of a radio synthesis image are distorted
- Bad for long baselines, large field of view, and long wavelengths
- Algorithms exist and work
  - But slowly….
- Will be a substantial problem for SKA



$$V(u,v,w) = \int I(l,m) e^{j2\pi\left(ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right)} dl dm$$

Faceted approaches

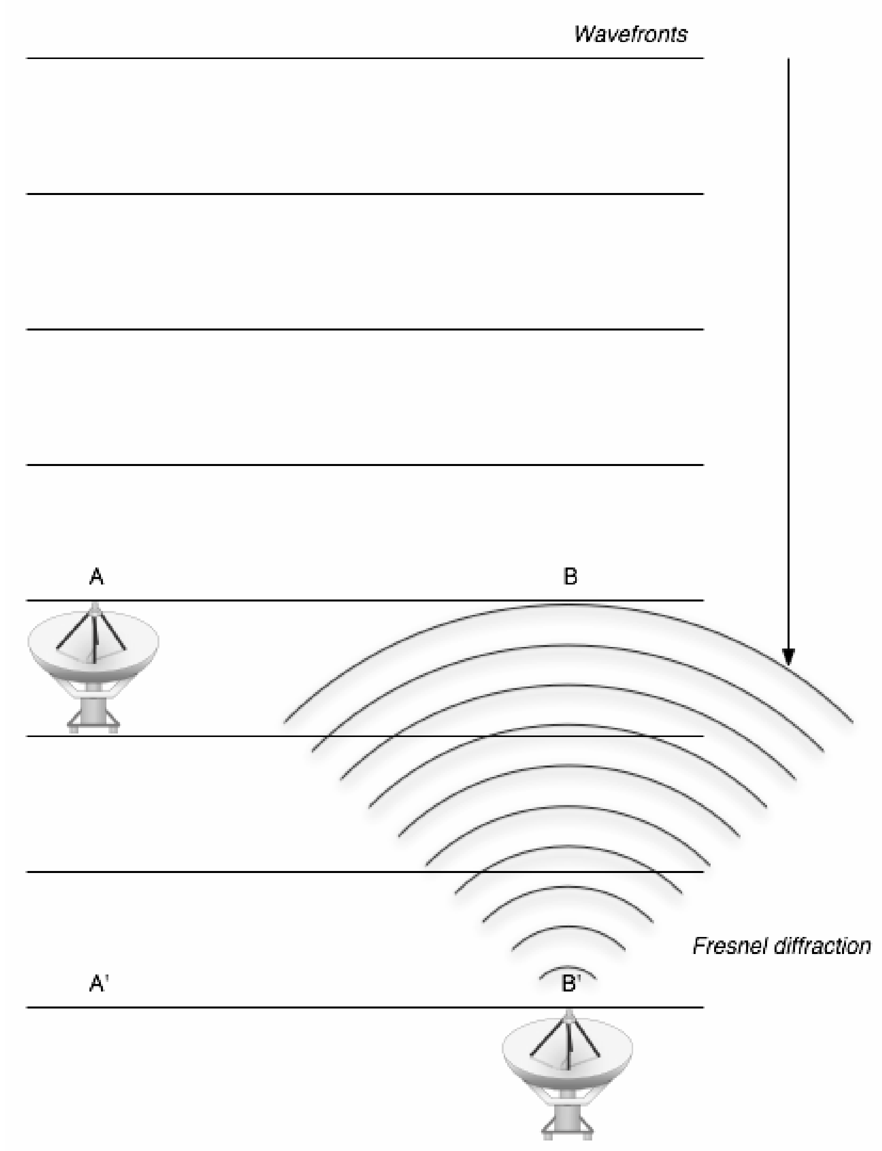
- Approximate integral by summation of 2D Fourier transforms

$$V(u,v,w) = \sum_k e^{j2\pi.\left(ul_k + vm_k + w\left(\sqrt{1-l_k^2-m_k^2}-1\right)\right)} \int I_k(l,m) e^{j2\pi.\left(u(l-l_k)+v(m-m_k)\right)} dl.dm$$

- Can do in image plane (SDE, AIPS) or Fourier plane (AIPS++)
- Fourier plane is better since it minimizes facet edge problems
- Number of facets ~ $\frac{3\lambda B}{D^2}$
- Parallelized via PVM (Cornwell 1992), MPI (Golap *et al.* 1999)

A simple piece of optics...

If we had measured on plane AB then the visibility would be the 2D Fourier transform of the sky brightness



Since we measured on AB', we have to propagate back to plane AB, requiring the use of Fresnel diffraction theory since the antennas are in each others near field

$$\tilde{G}(u,v,w) \approx e^{-j\pi w(u^2+v^2)}$$

The essence of W projection

- Evaluate this integral (and transpose) for regular grid in *(l,m)* and irregularly spaced samples in *(u,v)*

$$V(u,v,w) = \int I(l,m) e^{j2\pi\left(ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right)} dldm$$

- Image space computation

$$V(u,v,w) = \int G(l,m,w) I(l,m) e^{j2\pi(ul+vm)} dldm$$

- Fourier space computation

$$V(u,v,w) = G(u,v,w) \otimes V(u,v,w=0)$$

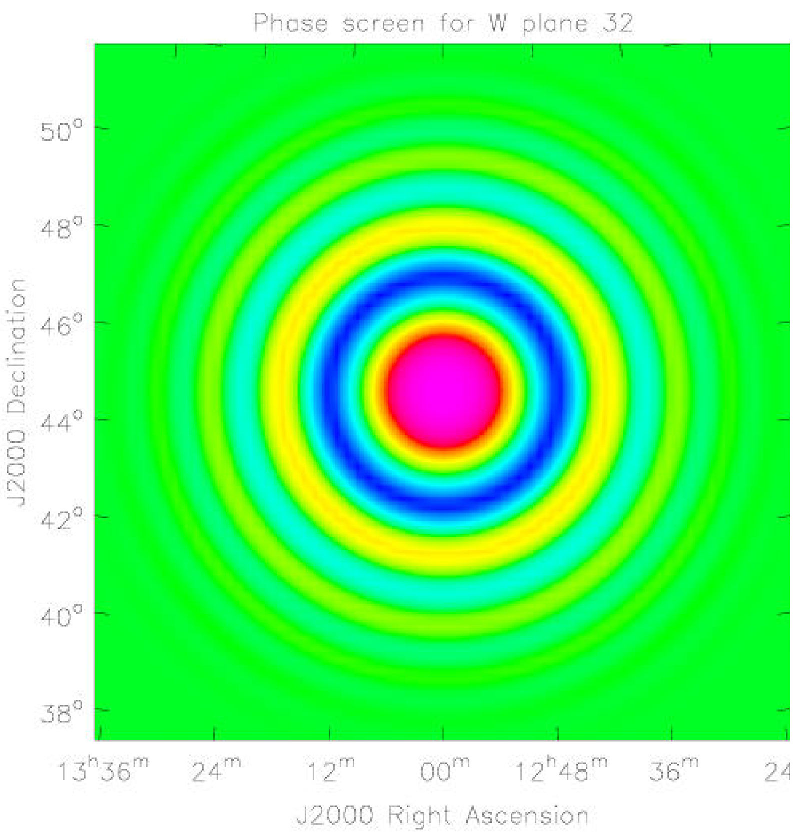
The convolution function

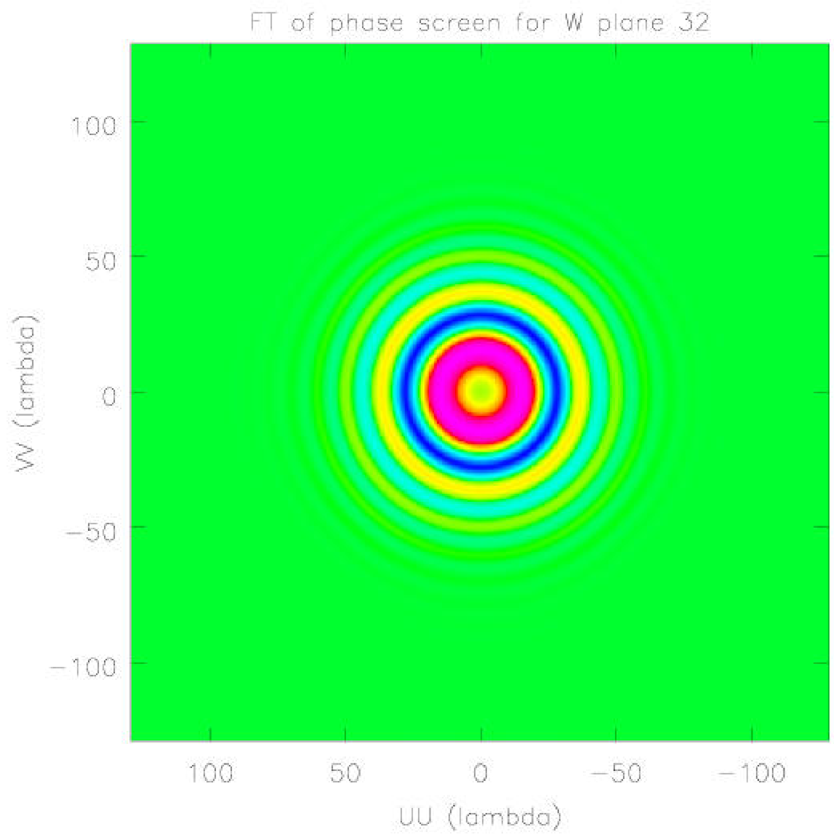
Image plane phase screen

$$e^{j2\pi w\left(\sqrt{1-l^2-m^2}-1\right)}$$

Fourier plane convolution function

$$\approx e^{-j\pi w(u^2+v^2)}$$





From narrow field to wide field

Standard narrow field measurement equation

$$V(u,v) = \int I(l,m) e^{j2\pi(ul+vm)} dldm$$

Fresnel diffraction

$$V(u,v,w) = \tilde{G}(u,v,w) \otimes V(u,v)$$

Standard wide field measurement equation derived using Van Cittert-Zernike theorem

$$V(u,v,w) = \int I(l,m) e^{j2\pi\left(ul+vm-w\left(l^2+m^2\right)/2\right)} dldm$$

$$\approx \int \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{j2\pi\left(ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right)} dldm$$

Wide field imaging = narrow field imaging + convolution

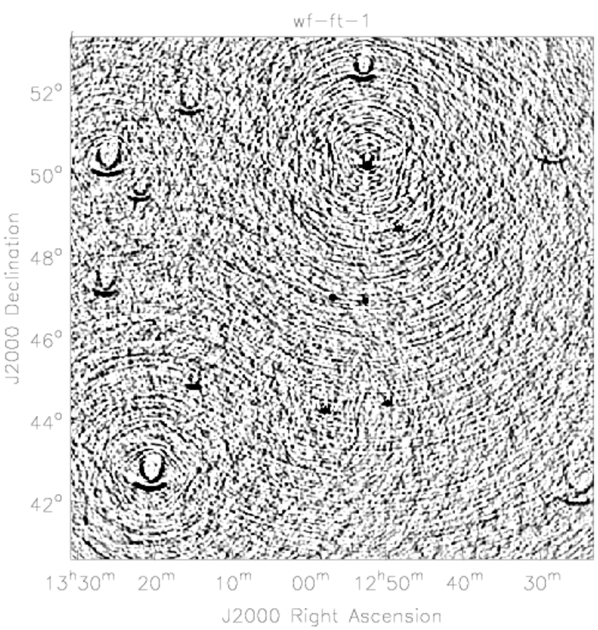
The W projection algorithm

- Calculate gridding kernel for range of values of $\sqrt{w}$
  - Fourier transform phase screens multiplied by spheroidal function (needed to control aliasing)
- Image to Fourier
  - Taper image by spheroidal function
  - Fourier transform
  - Estimate sampled visibilities by convolving gridded values with w dependent kernel
- Fourier to Image
  - Convolve sampled visibilities onto grid using w dependent kernel
  - Inverse Fourier transform
  - Correct for spheroidal function
- Deconvolution
  - Deconvolve in minor cycles using PSF for image center
  - Reconcile to visibility data in major cycles

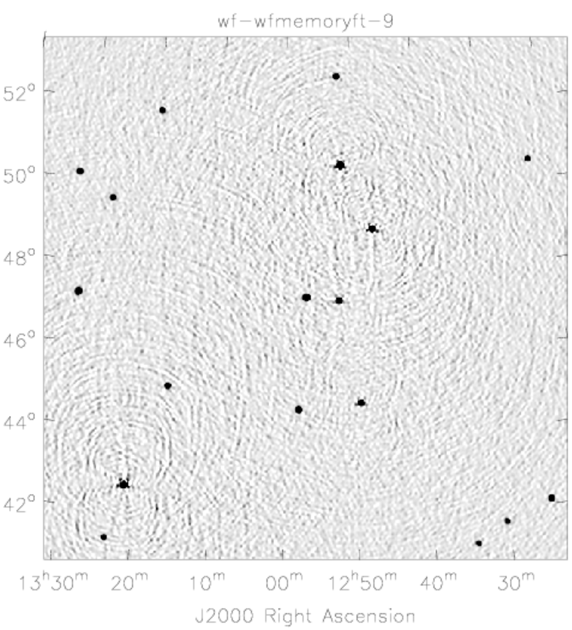
A synthetic example

- Simulation of ~ typical 74MHz field
  - Sources from WENSS
  - Long integration with VLA

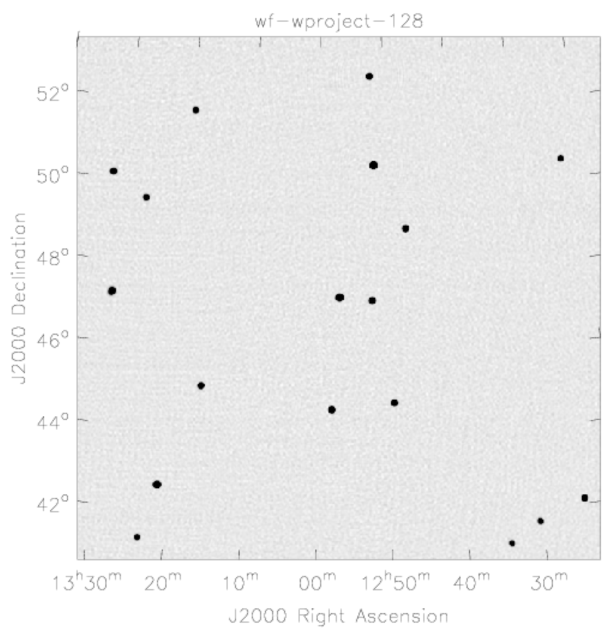
Fourier transform



UVW space facets



W projection



Computing costs for wide-field imaging

- Cleaning is very efficient for “Log N/Log S” fields
  - Cost of minor cycle is negligible
  - Costs are all in gridding, twice for each major cycle
  - Require only ~ 5 - 10 major cycles to reach dynamic range limits
- For extended emission, cost of minor cycles may dominate
  - *e.g.* Multi-Scale CLEAN can be very slow (but effective)
- Large overall penalty for small antennas

Number of antennas	**Time and frequency sampling**	**Non-coplanar baselines**	**Cleaning**	**Total**
N^2	$\frac{B^2}{D^2}$	$\frac{\lambda B}{D^2}$	$\frac{log(\Lambda)}{log(2\sqrt{N_s})}$	$\frac{N^2 B^3 \lambda}{D^4} \frac{log(\Lambda)}{log(2\sqrt{N_s})}$

Is W projection fast enough?

$$C_{SKA} : \$3.5\mathrm{M} \left(\frac{0.1}{\eta}\right)\left(\frac{f}{0.5}\right)^2\left(\frac{B}{5km}\right)^3\left(\frac{D}{12.5m}\right)^{-8}\left(\frac{\lambda}{0.2m}\right)\left(\frac{\Delta\nu}{500MHz}\right)2^{\frac{2(2010-t)}{3}}$$

- $1~30Mflops/s (2010)
- Antenna diameter scaling is horrific!
  - Doubling antenna size saves factor of 256 in computing
- Baseline dependency is tough
- Easy to find hardware costs > SKA cost
- Multi-fielding not included
- Error ~ factor of 3 in each direction
- For 350km baselines with 25m antennas
  - *$120M in 2015*
  - *Without w projection, would be ~ $1B*

Future problems….

- It's only going to get worse
  - Pointing errors
  - Antenna primary beam idiosyncrasies
  - Wide bandwidth
- Computing requirements will increase beyond even W projection

Better approach to parallelization?

- Image space computation

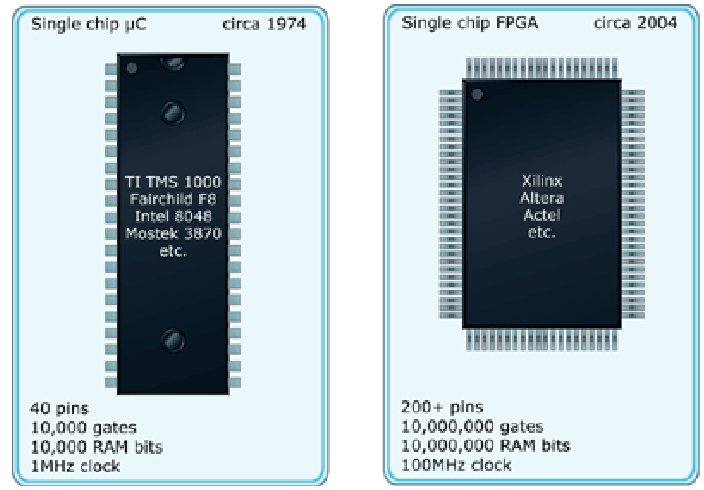
$$V(u,v,w) = \int G(l,m,w) I(l,m) e^{j2\pi(ul+vm)} dl dm$$

- Each processor does:
  - Image plane weighting by function
  - Fourier transform
  - Degridding with limited size convolution function
- Hybrid possible
  - Some image plane, some Fourier plane
  - Can tune division to match machine
- Scaling law will be the same but perhaps with smaller coefficient

Another approach...

- Parallel machine where each node can do large convolutions *quickly*...
  - Implement via FPGAs
  - Cray XD1

- Convolution step done in a few clock cycles
  - ~ 100 times faster
  - Scaling due to non-coplanar baselines vanishes



$$C_{SKA} : \$22\mathrm{K}\left(\frac{0.1}{\eta}\right)\left(\frac{f}{0.5}\right)^{2}\left(\frac{B}{5km}\right)^{2}\left(\frac{D}{12.5m}\right)^{-6}\left(\frac{\Delta\nu}{500MHz}\right)2^{\frac{2(2010-t)}{3}}$$

Summary

- The non-coplanar baselines effect is caused by Differential Fresnel diffraction
- W projection corrects the non-coplanar baselines effect by convolving with Fresnel diffraction kernel in uvw space before Fourier transform
- W projection is an order of magnitude faster than facet based methods
- Non coplanar baselines effect is still a significant obstacle for SKA
- Application-specific acceleration very promising