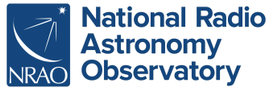
Widefield Imaging

21st Synthesis Imaging Workshop

Preshanth Jagannathan, NRAO

2026-05-29



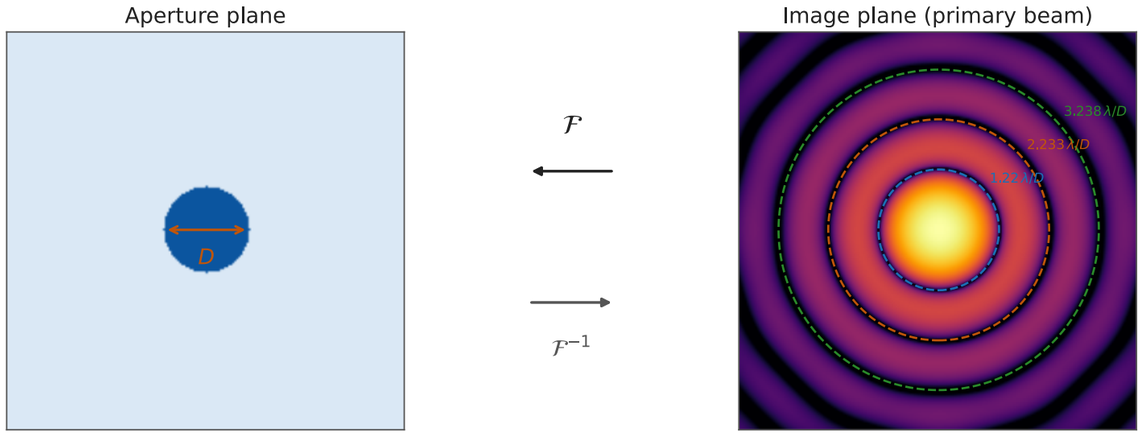
What came before

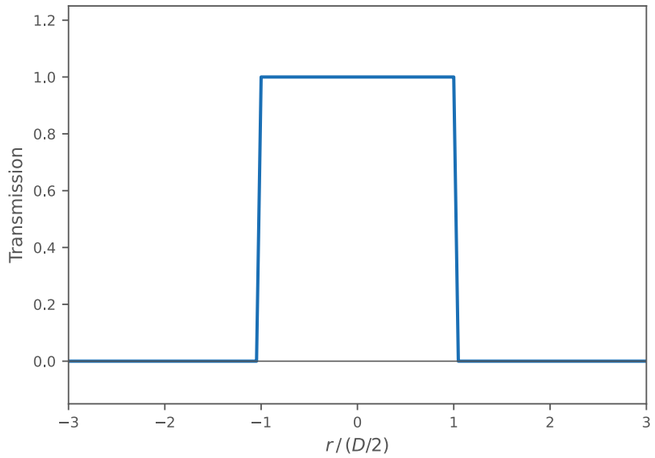
- **Urvashi Rau** – introduction to radio astronomy and interferometry
- **Josh Marvil** – calibration and imaging

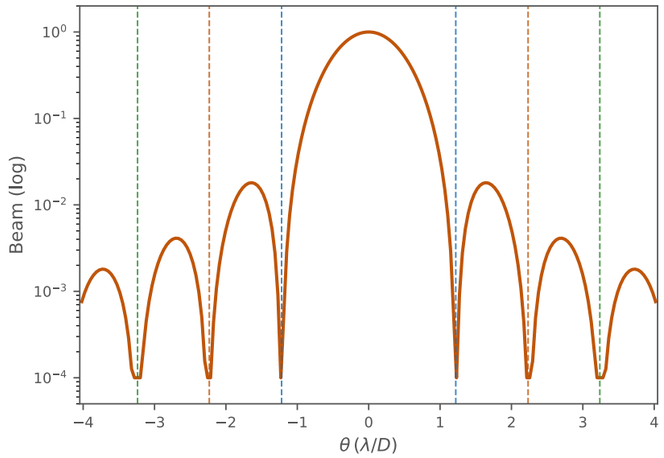
Today: the standard imaging equation breaks for wide fields. We look at what breaks and what to do about it.

Wideband effects – PB spectral index, wideband beam errors – are covered by Urvashi.

Fourier pairs







Two Fourier pairs that recur throughout this talk:

- $P(l,m) \xrightarrow{\mathcal{F}} A(u,v), \quad A(u,v) \xrightarrow{\mathcal{F}^{-1}} P(l,m)$

  Primary beam to aperture illumination. The primary beam (PB) is the field of view of a homogeneous array.

- $S(u,v) \xrightarrow{\mathcal{F}} \mathrm{PSF}(l,m)$,
  $\mathrm{PSF}(l,m) \xrightarrow{\mathcal{F}^{-1}} S(u,v)$

  uv sampling function to point spread function. Incomplete coverage gives an imperfect PSF.

Convolution theorem

$$f(x) * g(x) \quad \longleftrightarrow \quad F(\xi) \cdot G(\xi)$$

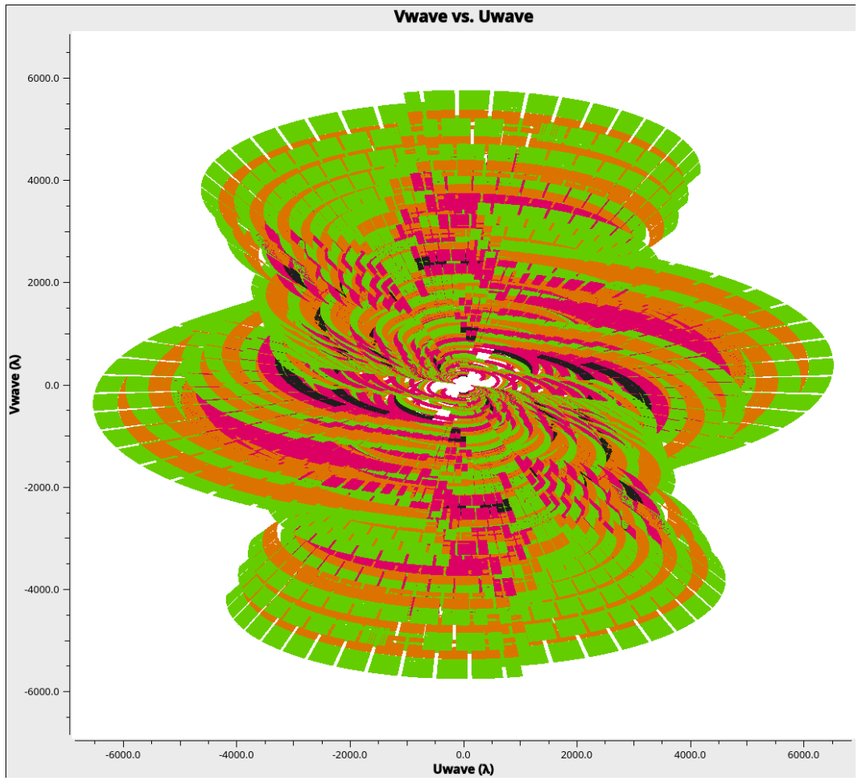
Convolution in one domain is multiplication in the other.

In radio imaging: the dirty image is the true sky convolved with the PSF:

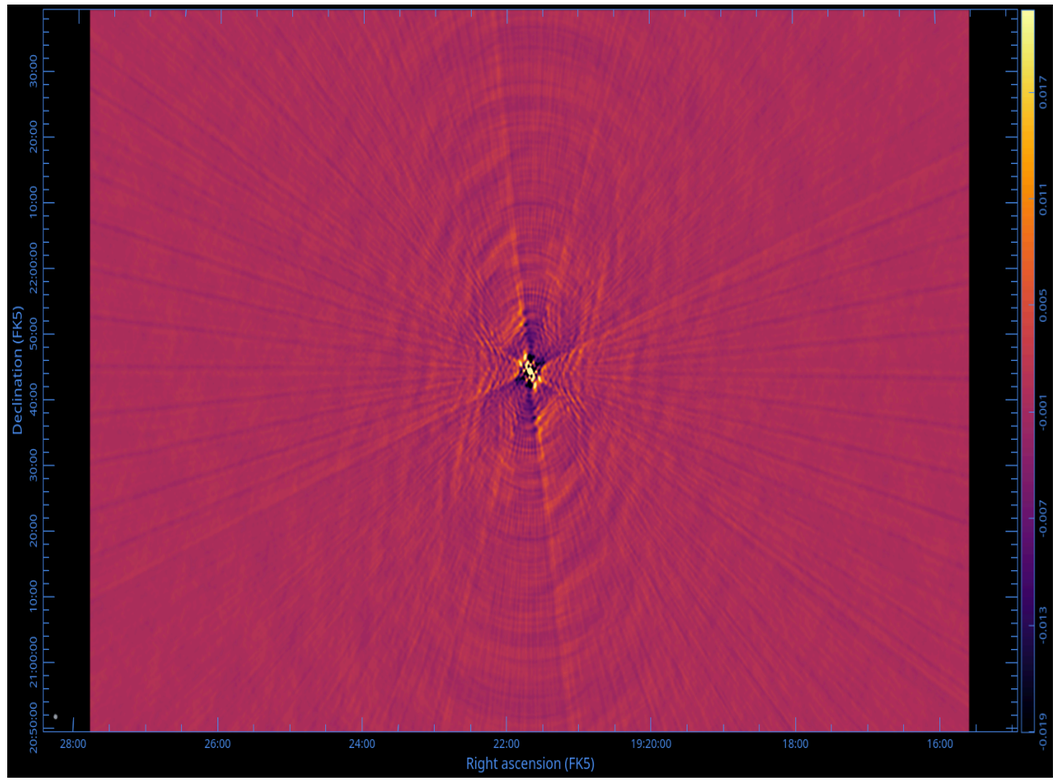
$$I^{\text{dirty}}(l, m) = \text{PSF}(l, m) * I^{\text{sky}}(l, m)$$

because multiplying by $S(u, v)$ in the uv plane is convolution with the PSF in the image plane. Deconvolution (CLEAN) undoes this.

uv coverage and the PSF



$S(u, v)$: VLA uv coverage.



$\mathcal{F}\{S\} = \mathrm{PSF}$. Gaps in coverage become sidelobes.

The Measurement Equation

The standard imaging equation

An interferometer measures the Fourier transform of the sky brightness:

$$V_{ij}^{\mathrm{obs}}(\nu, t) \approx M_{ij}(\nu, t)\, S_{ij}(\nu, t) \iint I(l, m)\, e^{-2\pi i(ul+vm)}\, dl\, dm$$

- M_{ij} – direction-independent gains (calibration, outside the integral)
- S_{ij} – uv sampling function
- $I(l, m)$ – sky brightness
- $e^{-2\pi i(ul+vm)}$ – 2D Fourier kernel

This is what Josh covered. It works when the field is small. What changes when it is not?

The full measurement equation

Standard form:

$$V_{ij}^{\mathrm{obs}}(\nu,t) \approx M_{ij}(\nu,t)\, S_{ij}(\nu,t) \iint I(l,m)\, e^{-2\pi i(ul+vm)}\, dl\, dm$$

Full widefield form – new terms in blue:

$$V_{ij}^{\mathrm{obs}}(\nu,t) \approx M_{ij}(\nu,t)\, S_{ij}(\nu,t) \iint M_{ij}^{S}(l,m,\nu,t)\, I(l,m,\nu,t)\, e^{-2\pi i(ul+vm+w(n-1))}\, \frac{dl\, dm}{n}$$

- $M_{ij}^{S}(l,m,\nu,t)$ – direction-dependent effects: primary beam (this talk), ionosphere (not covered here)
- ν, t in the sky – frequency and time variability
- $w(n-1)$ – geometric phase from sky curvature

Two problems, one equation

$$M_{ij}(\nu,t)\, S_{ij}(\nu,t) \iint M^S_{ij}(l,m,\nu,t)\, I(l,m,\nu,t)\, e^{-2\pi i(ul+vm+w(n-1))}\, \frac{dl\, dm}{n}$$

$\underbrace{M_{ij}(\nu,t)}$
direction-independent gains (outside the integral)

$\underbrace{M^S_{ij}(l,m,\nu,t)}$
primary beam
direction, time, frequency, polarization

$\underbrace{w(n-1)}$
direction-dependent effect
w-term

M_{ij} outside the integral means calibration is separable. Both blue terms are per-baseline. M^S_{ij} is a full 4×4 Mueller matrix.

The equivalent matrix form

The same equation written in terms of Jones matrices and the aperture:

$$\vec{V}_{pq}^{\,\mathrm{obs}} = \mathcal{F}_{pq}\, S^{ab}\, (\vec{A}_i \circledast \vec{A}_j^{*}) \star \vec{V}_{ab}$$

where $\mathcal{F}^{\dagger}(\vec{A}_{ij}) = \vec{M}_{ij}$ and the Jones chain is:

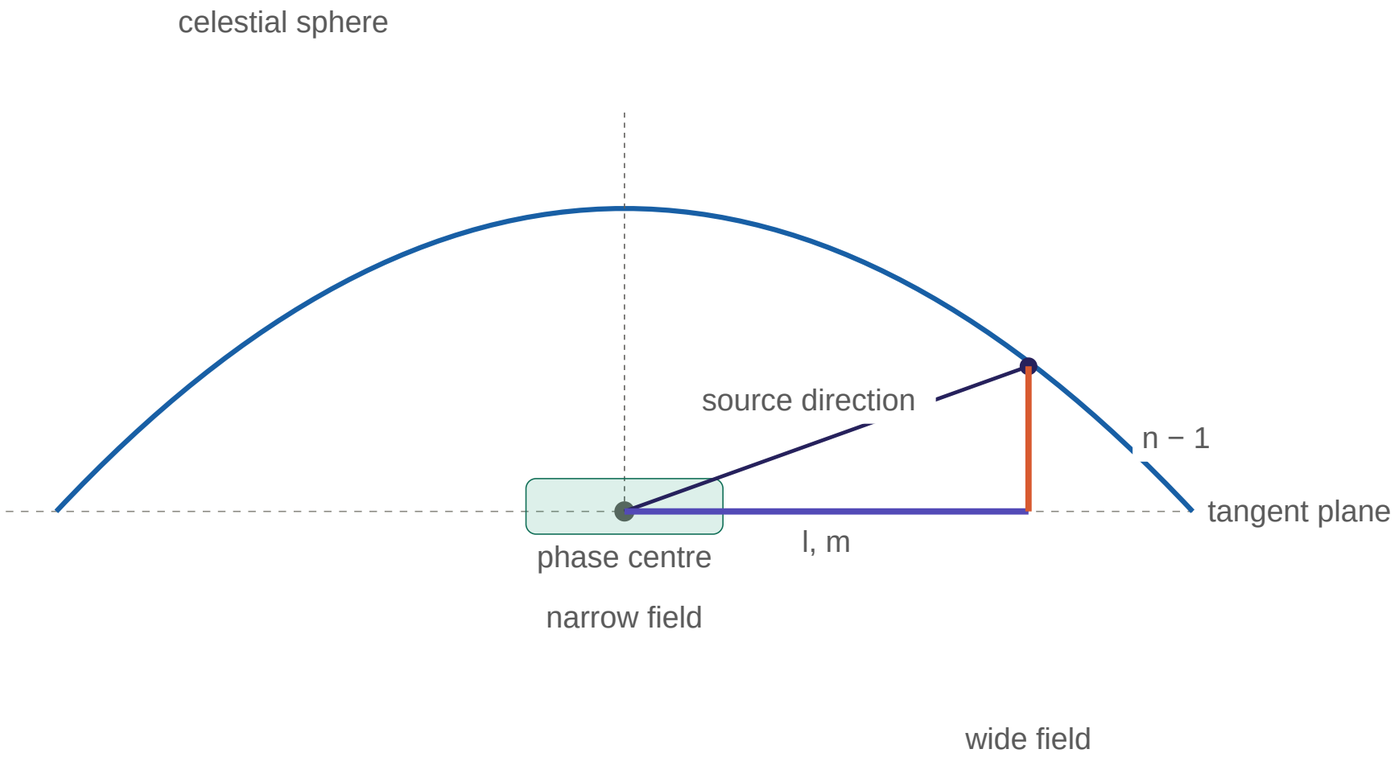
$$\vec{e}_a = J_i \cdot \vec{\epsilon} \qquad \vec{e}_b = J_j \cdot \vec{\epsilon} \qquad \vec{I}_{ab} = \vec{e}_a \otimes \vec{e}_b$$

These are completely equivalent descriptions of the same physics.

- The integral form is useful for understanding what the measurement computes
- The matrix form is useful for algorithm and computational design

Direction Dependent Effect : W Term

Sky curvature and the w-term



l, m

$n = \sqrt{(1 - l^2 - m^2)}$

When does the w-term matter?

Deriving the threshold

W-term phase at sky angle θ, for baseline component $w = B/\lambda$:

$$\Delta\phi_w = -2\pi\, w(n-1) \approx \pi\,\frac{B}{\lambda}\,\theta^2$$

using $n - 1 \approx -\theta^2/2$ for small angles. At the FoV edge $\theta \sim \lambda/D$:

$$|\Delta\phi_w| \approx \pi\,\frac{\lambda B}{D^2}$$

Set $|\Delta\phi_w| = 1\,\mathrm{rad}$. The critical baseline is:

$$B_{\mathrm{thresh}} = \frac{D^2}{\pi\,\lambda}$$

Below this the w-term is negligible. Above it, correction is needed.

VLA at L-band ($\lambda = 0.21\,\mathrm{m}$, $D = 25\,\mathrm{m}$)

$$B_{\mathrm{thresh}} = \frac{625}{\pi \times 0.21} \approx 950\,\mathrm{m}$$

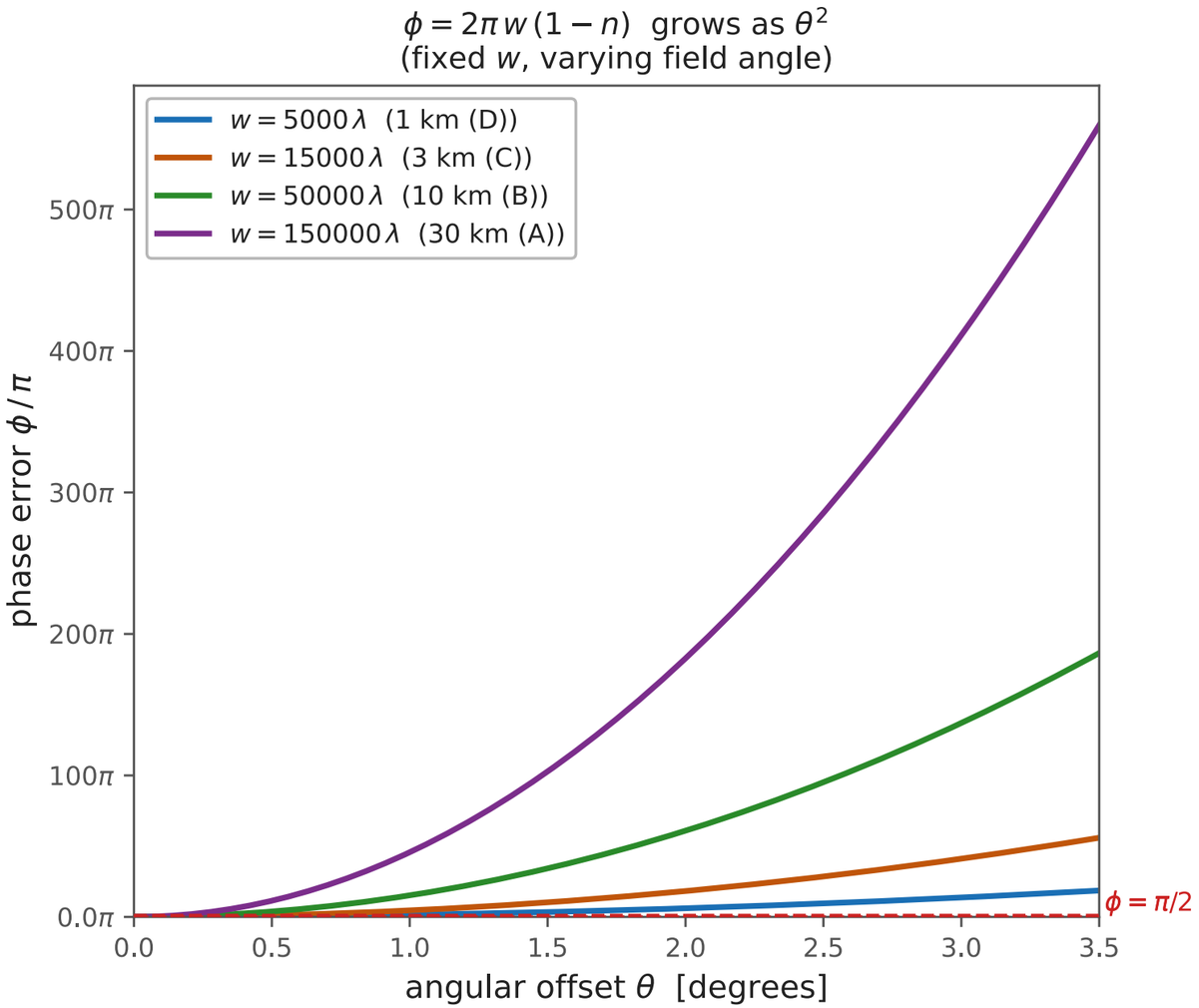
D config extends to 1 km; A config to 35 km. The w-term matters for any configuration beyond C.

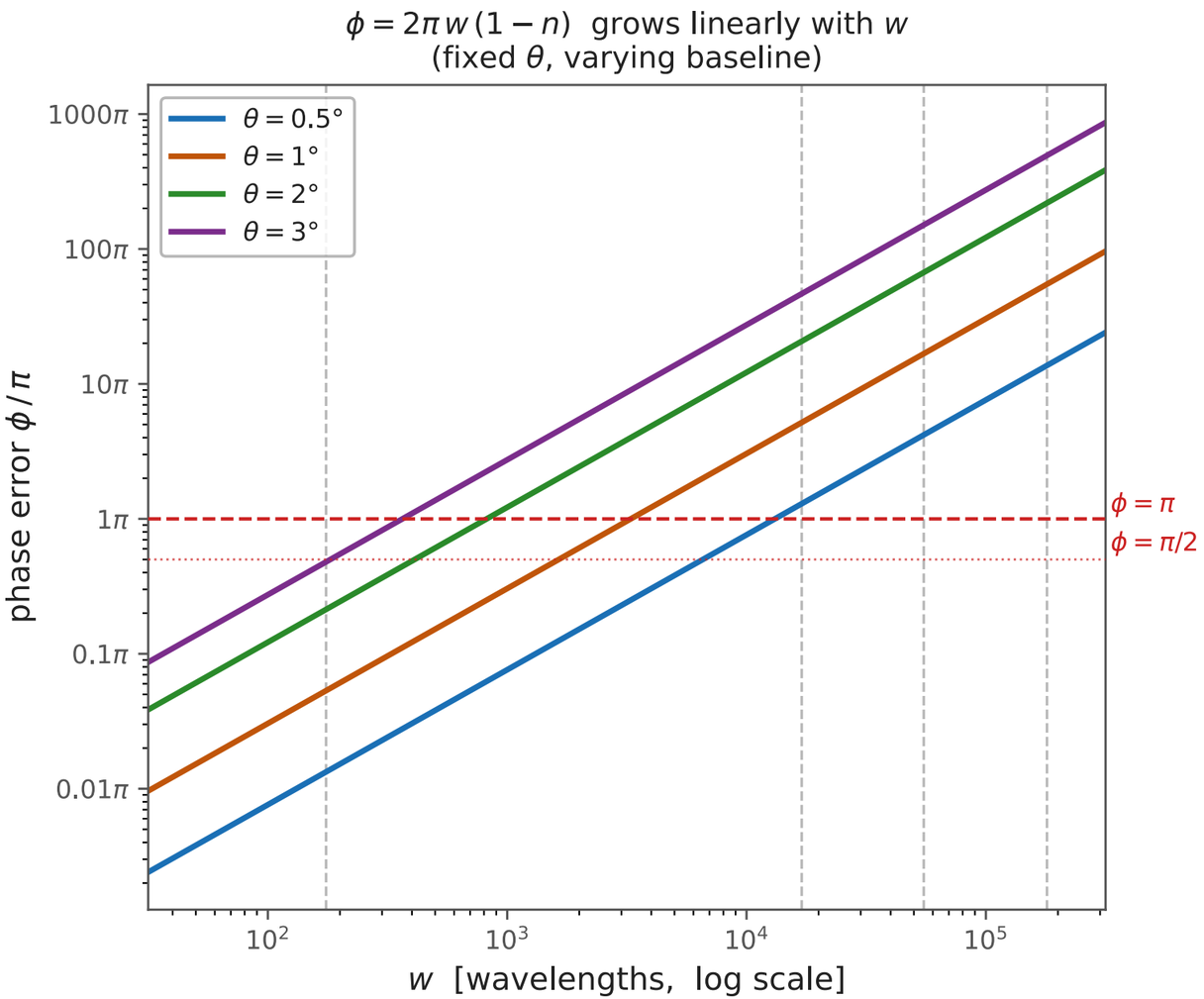
MeerKAT at L-band ($\lambda = 0.21\,\mathrm{m}$, $D = 13.5\,\mathrm{m}$)

$$B_{\mathrm{thresh}} = \frac{182}{\pi \times 0.21} \approx 276\,\mathrm{m}$$

The smaller dish shrinks B_{thresh} dramatically. Nearly every MeerKAT baseline exceeds it.

The w-term effect



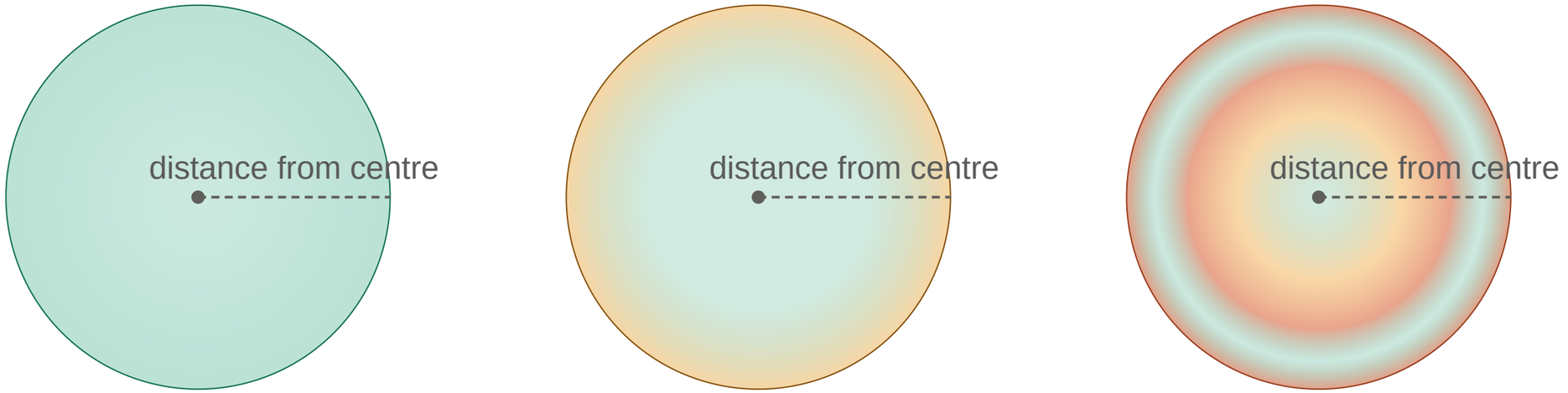


Left: for a fixed baseline, phase error grows as θ^2 – sources far from the phase centre are most affected.

Right: for a fixed sky position, phase error grows linearly with w – longer baselines and non-coplanar arrays pay the highest cost.

The Fresnel phase across the field

w-term phase error across the field: $2\pi\, w\, (n - 1)$



narrow field

medium field

wide field

Correcting the w-term

Four algorithms, same physics, different cost models:

Faceting (Cornwell and Perley 1992): tile the sky into tangent-plane patches small enough that $n \approx 1$ per facet. Standard 2D FFT per facet. Cost scales with number of facets.

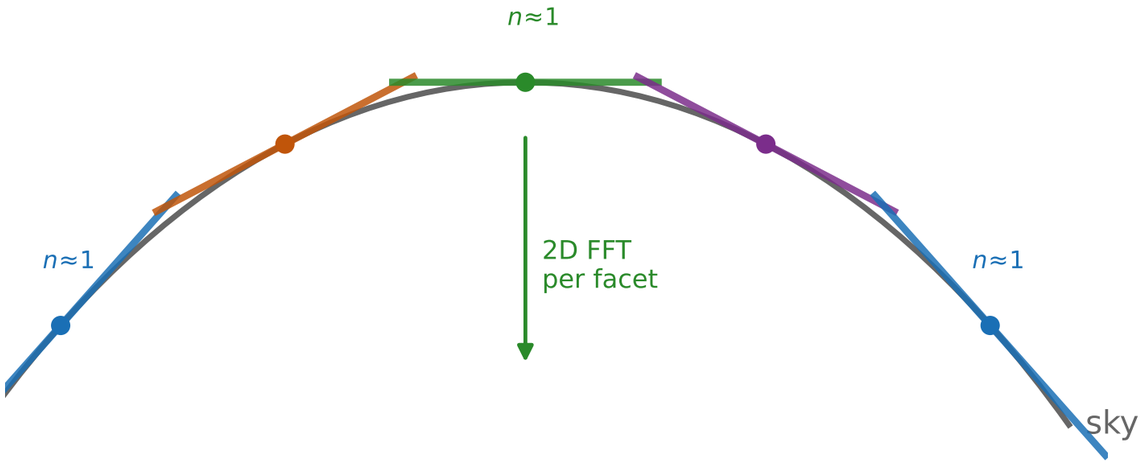
W-Projection (Cornwell et al. 2008): fold the Fresnel phase into the gridding kernel $\tilde{G}(u, v, w) \approx e^{j\pi(u^2+v^2)/w}/jw$. One grid, one FFT. Kernel support scales with $w_{\max}$.

W-snapshot (Cornwell et al. 2012): image each short integration separately where the array is co-planar ($w \approx au + bv$, a linear combination of u and v), then **regrid** (l, m) to a common tangent plane before stacking. The regridding step is the main cost.

W-stacking (Cornwell et al. 2012; Offringa et al. 2014): bin visibilities by w, FFT each layer, multiply by $e^{2\pi i w(n-1)}$ in the image plane, stack. No regridding – correction is a multiplication. Used by WSClean.

Faceting and W-Projection

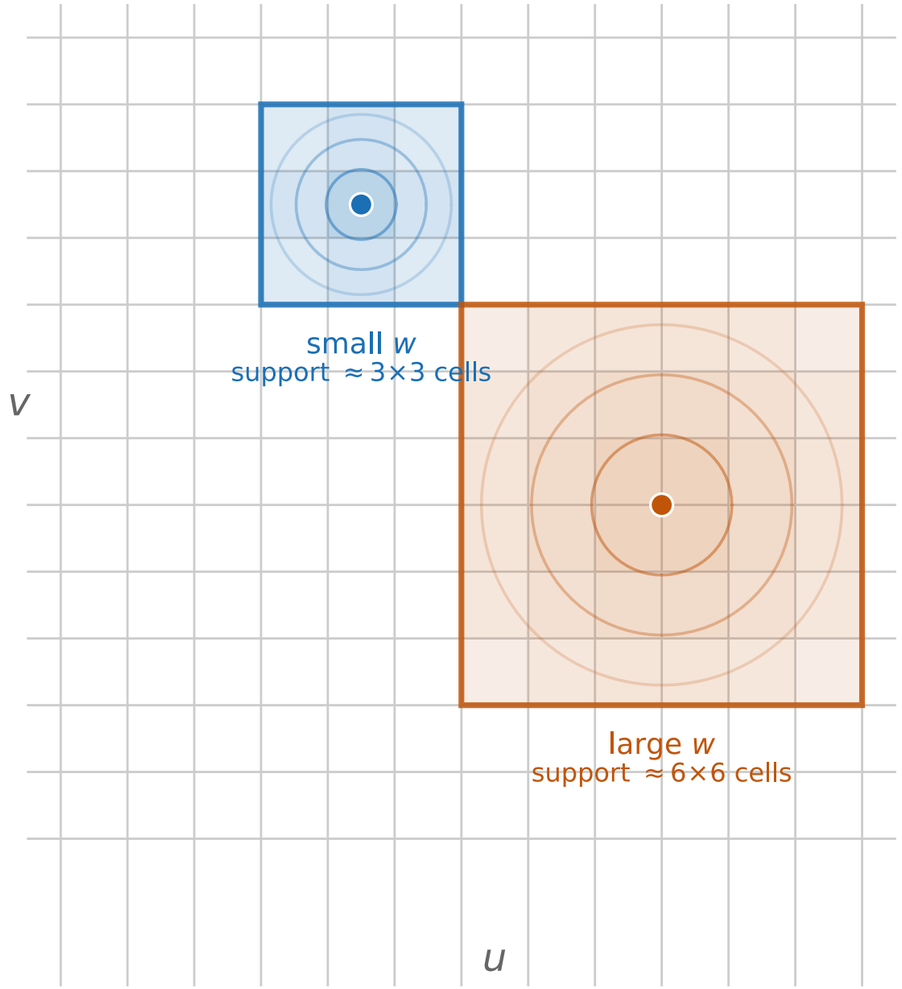
Faceting



tile sky into flat patches — ignore *w* within each

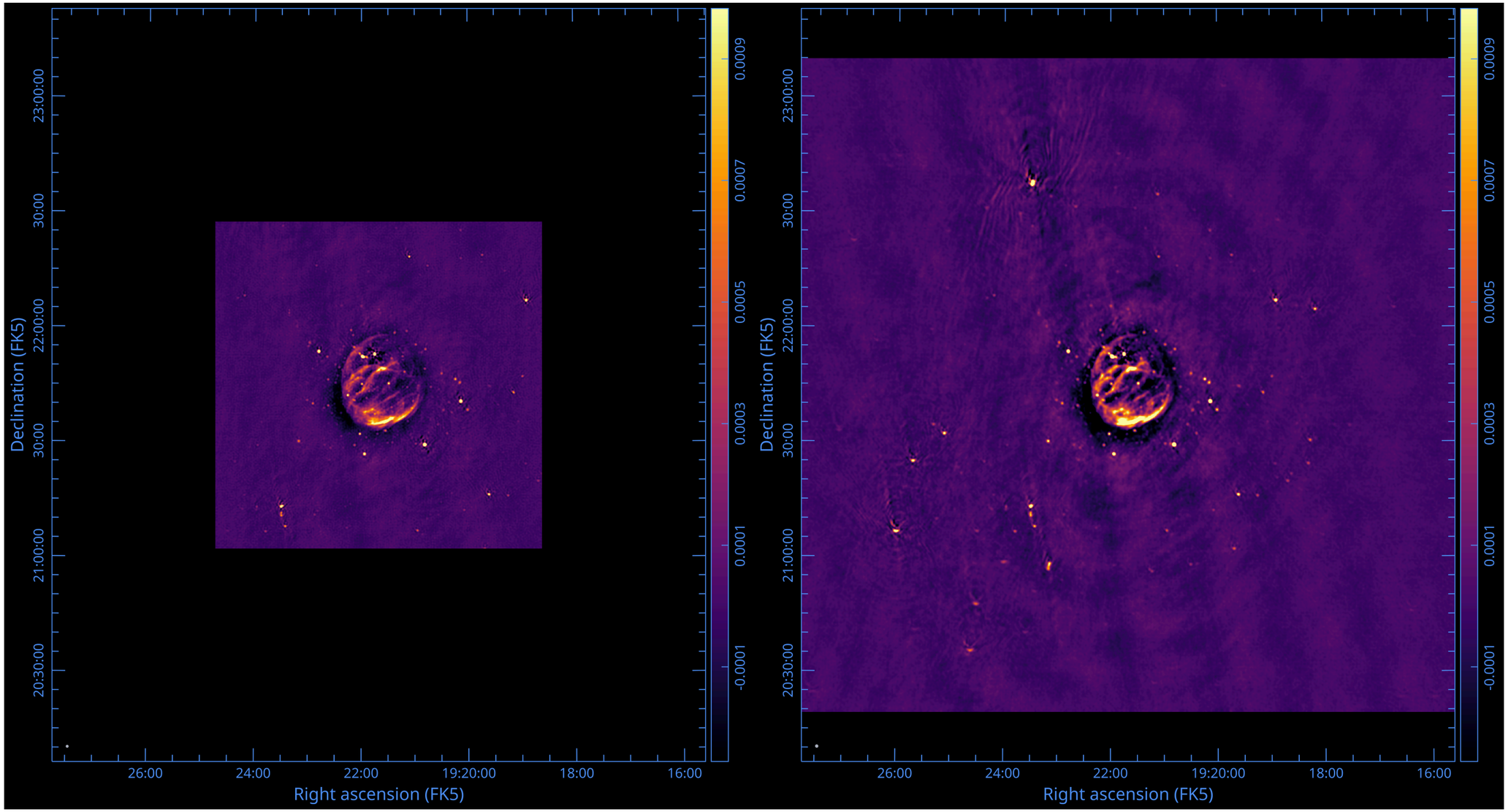
W-Projection

one grid ⟶ one FFT



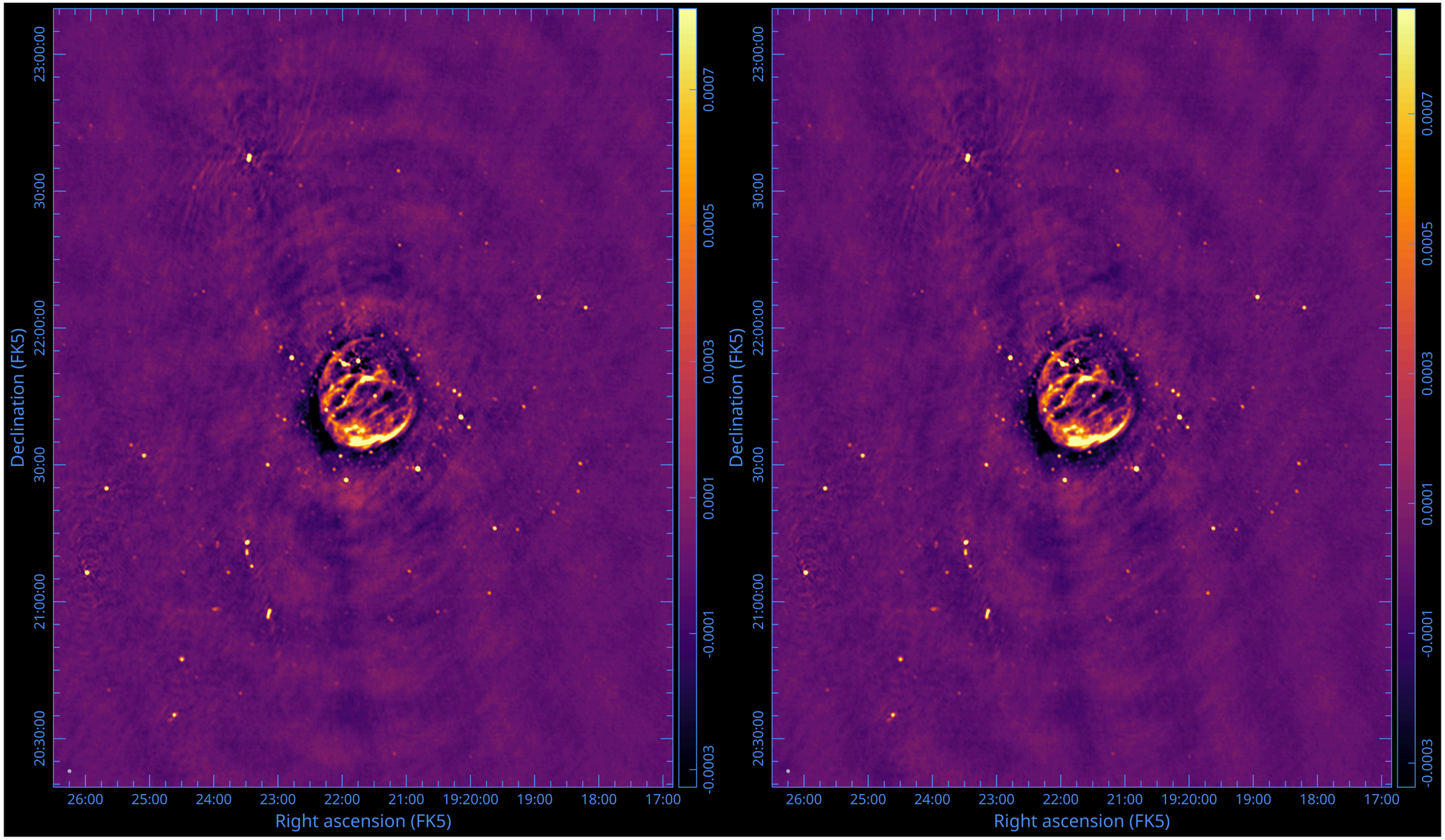
grid each visibility with its Fresnel kernel — kernel support grows with *w*

Image what your science needs



Left: outlier field gridder – smaller image, targeted at the science target, best image quality. Right: standard prolate spheroidal gridder – full wide-field image, lower quality across the field.

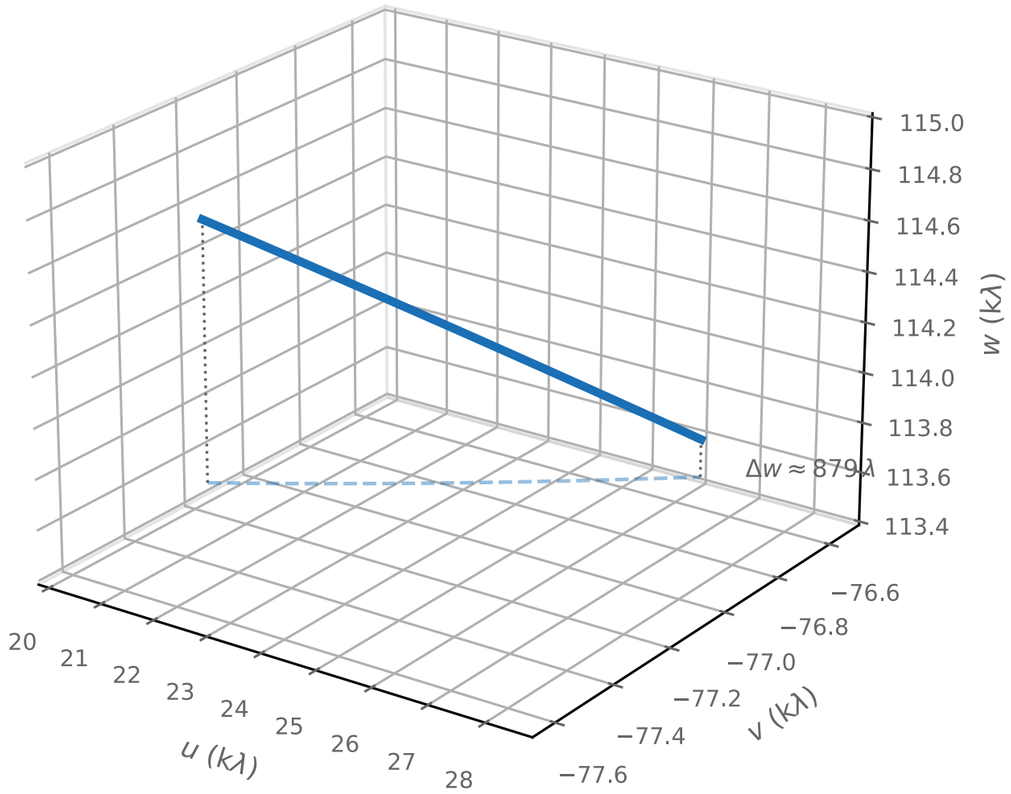
Faceting vs W-Projection – real data



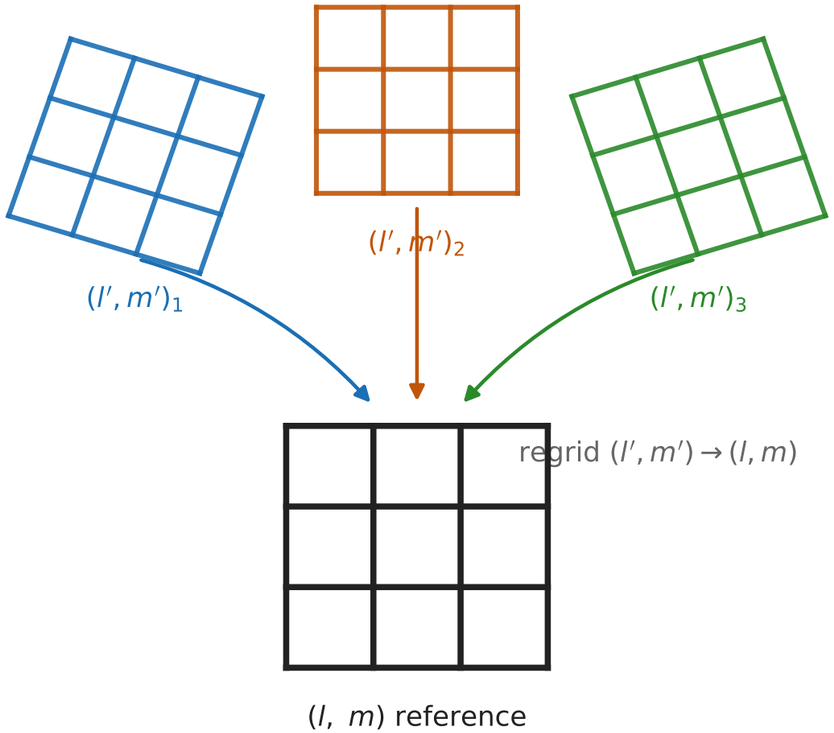
Left: faceting. Right: W-Projection. Same data, same field.

W-Snapshot

10-min snapshot: 30 km, 1.4 GHz, $\delta = 34°$
$\varphi_{\max} = \pi/4 \quad \Rightarrow \quad \Delta t_{\text{snap}} \approx 9$ min



Each snapshot images in its own tangent plane



Coordinate regridding from (l', m') to (l, m) per snapshot is the dominant cost. Δt_{snap} shortens at low declination – more snapshots, higher cost. Tangent-plane offsets are exaggerated for clarity.

W-stacking

Bin visibilities by w into layers. Grid each layer with a standard kernel, FFT to the image plane, multiply by $e^{2\pi i w_k (n-1)}$, stack.

$$I_{\text{dirty}}(l,m) = \sum_k e^{2\pi i w_k (n-1)} \; \mathcal{F}^{-1}\left[\sum_{j:\, w_j \in \text{layer } k} V_j(u,v) \right]$$

- k – w-layer index; w_k – representative w value for layer k
- $n = \sqrt{1 - l^2 - m^2}$ – direction cosine along the line of sight
- $e^{2\pi i w_k (n-1)}$ – Fresnel phase correction applied per layer in the image plane
- $\mathcal{F}^{-1}[\cdots]$ – inverse FFT of visibilities gridded within layer k

The correction is an **image-plane multiplication** – not a convolution kernel, not a coordinate regrid. Memory cost is one image per w-layer.

WSClean uses w-stacking, choosing the number of layers to keep the residual phase below a set threshold.

Direction-dependent effects: the primary beam

The primary beam in the ME

$$\vec{V}_{ij}^{\,\mathrm{obs}}(\nu,t) \approx M_{ij}(\nu,t)\, S_{ij}(\nu,t) \iint M_{ij}^{S}(l,m,\nu,t)\, \vec{I}(l,m,\nu,t)\, e^{-2\pi i(ul+vm+w(n-1))}\, \frac{dl\, dm}{n}$$

$\vec{V}_{ij}$ is a **4-vector** – (pp, pq, qp, qq) or equivalently (I, Q, U, V). M_{ij}^{S} is therefore a 4×4 Mueller matrix, not a scalar. The primary beam couples polarization states: even an unpolarized sky produces apparent polarization if the beam is not corrected.

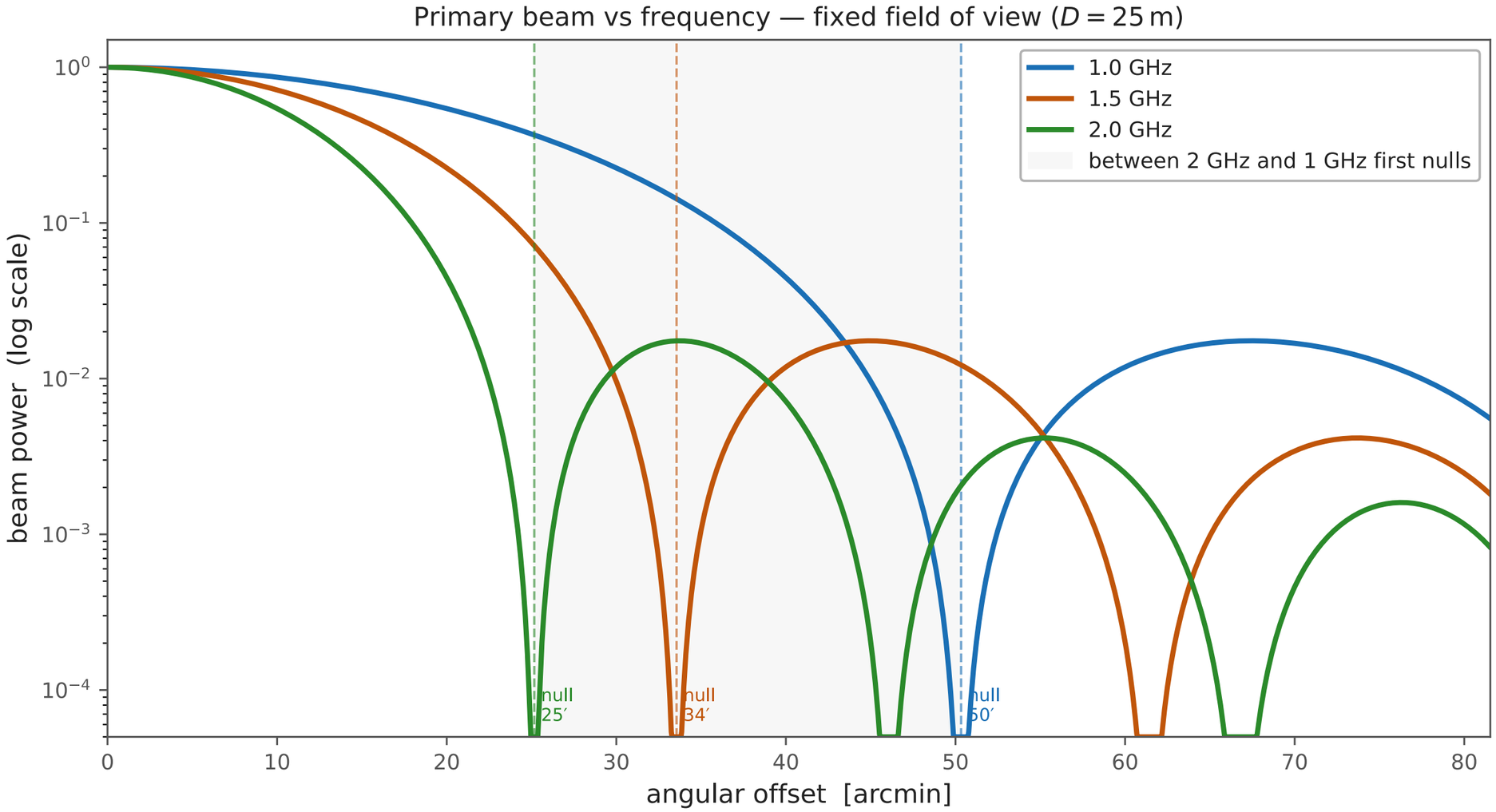
The standard assumption: one static scalar beam divided out of the image:

$$I^{\mathrm{obs}}(l,m) = \sum_{ij,t,\nu} I_{ij}^{\mathrm{PSF}}(l,m,t,\nu) * \Big[P_{ij}(l,m,t,\nu)\, I_{ij}^{\mathrm{sky}}(l,m,t,\nu) \Big]$$

P_{ij} changes with **direction** (l,m), **time** t (parallactic angle), and **frequency** ν (beam scales as λ).

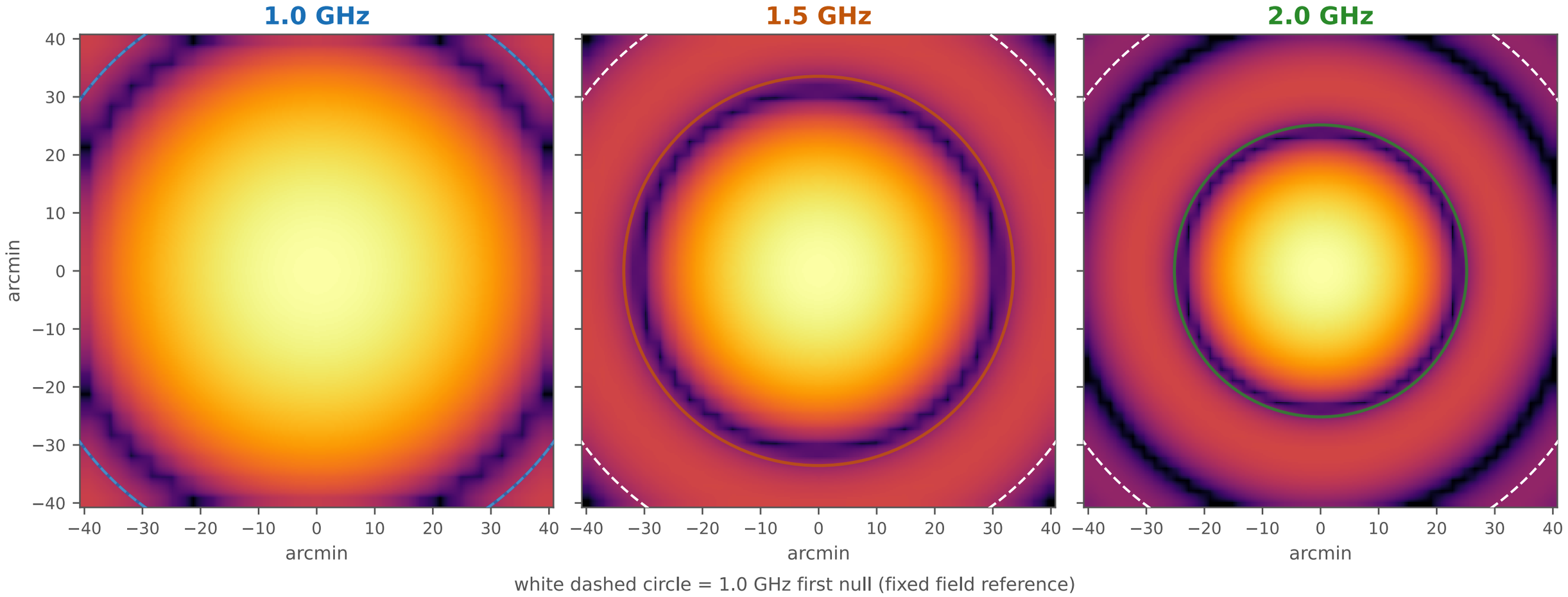
For PB spectral index effects see Urvashi's talk.

Frequency-dependent beam



At fixed field of view the sidelobes of the 2.0 GHz beam fall inside the field where the 1.0 GHz main lobe was. A single static beam model cannot correct this.

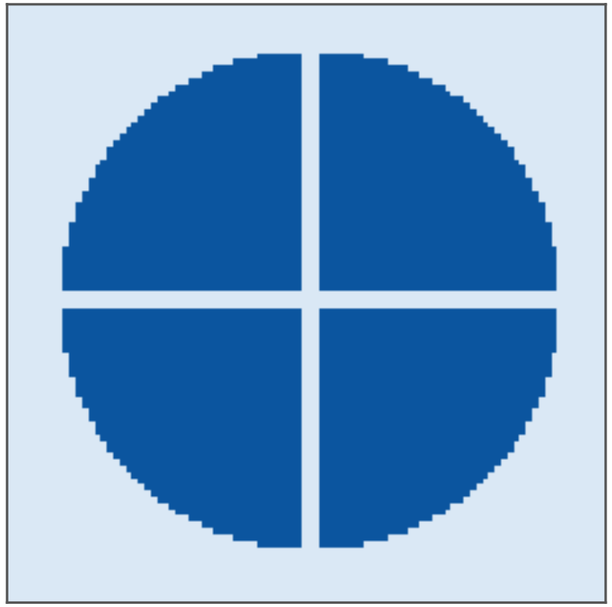
Frequency-dependent beam – 2D



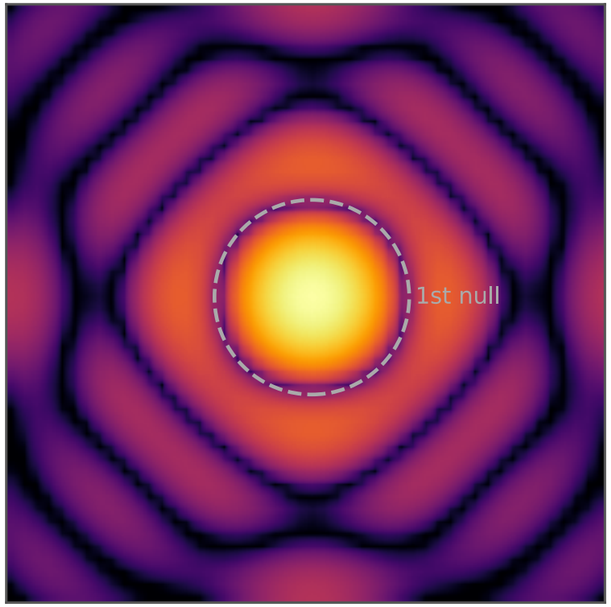
White dashed circle is the 1.0 GHz first null, fixed on all three panels. The coloured ring is the first null of each frequency.

Time-dependent beam: parallactic angle

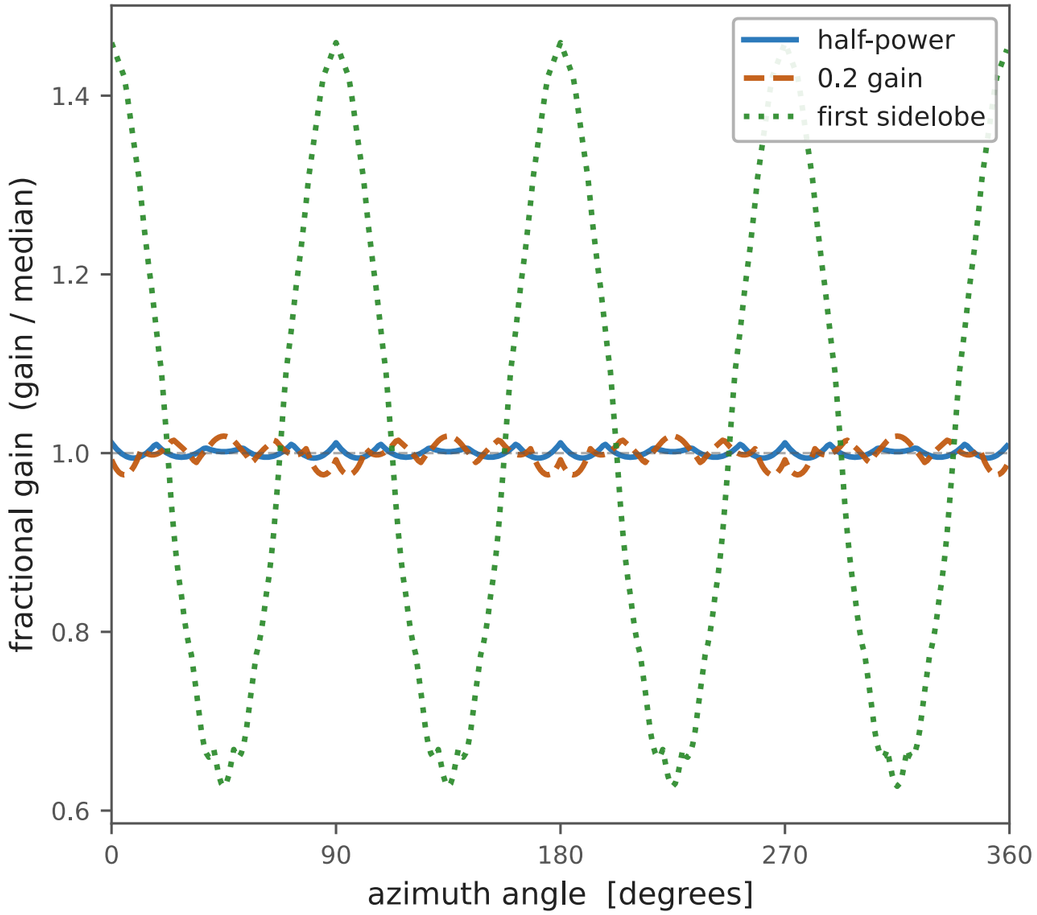
Aperture with cross shadow



Beam (PA = 0°) non-circular sidelobes



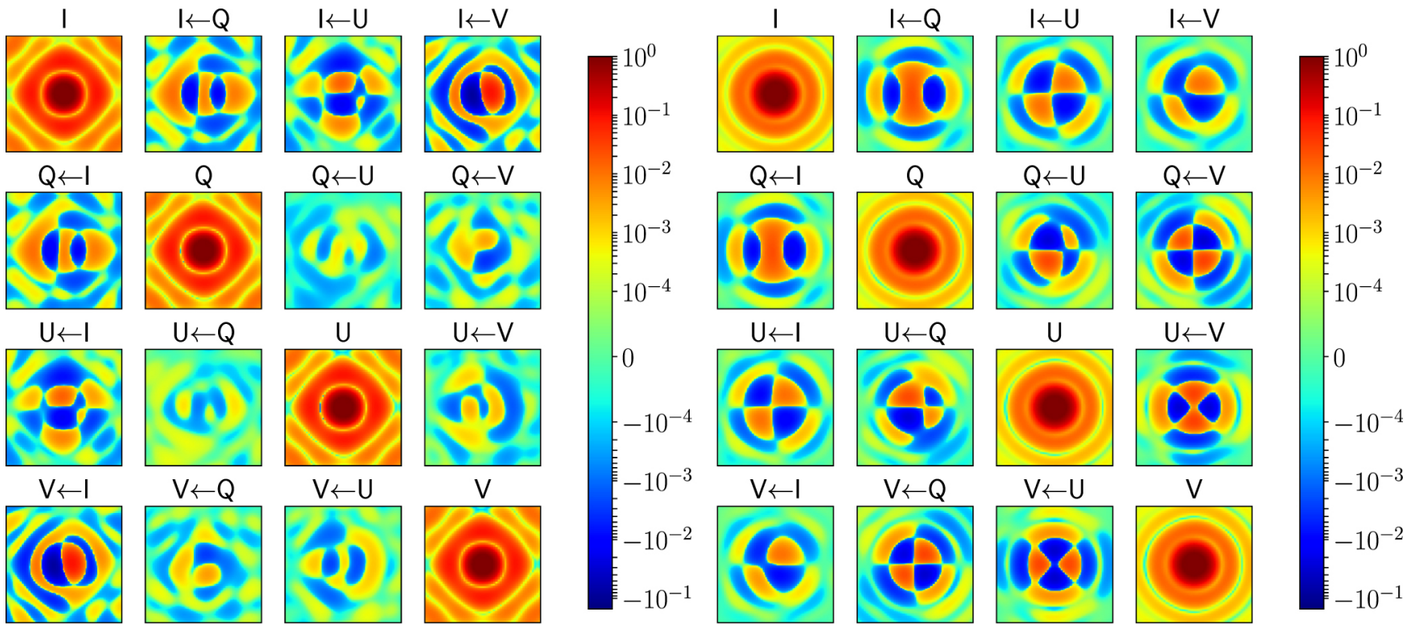
Fractional gain vs azimuth at three field radii (PA = 0°)



The feed support casts a cross shadow on the aperture. The FT gives structured, non-circular sidelobes. As Earth rotation changes the parallactic angle, the cross rotates – the gain at any fixed sky position traces the oscillating curves on the right.

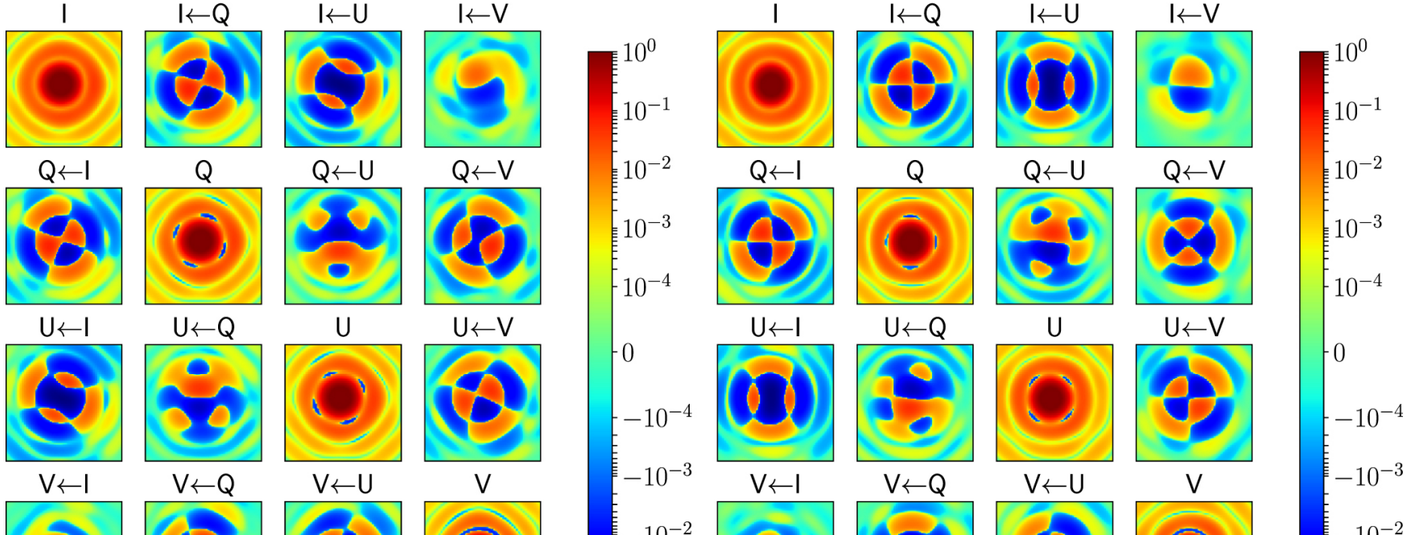
A-Projection

Polarimetric beam structure



(a) EVLA

(b) MeerKAT



The beam is a 4×4 Mueller matrix. Off-diagonal terms couple Stokes parameters: an unpolarized sky produces apparent polarization if the beam is not corrected.

Deriving the Mueller correction

The aperture conjugate is the matched filter for the beam; gridding with it deconvolves the beam during the gridding step itself.

The full polarimetric correction requires the Mueller form of the A-term. For antenna pair ij:

$$M_{ij}(l, m; \nu, t) = J_i(l, m; \nu, t) \otimes J_j^*(l, m; \nu, t)$$

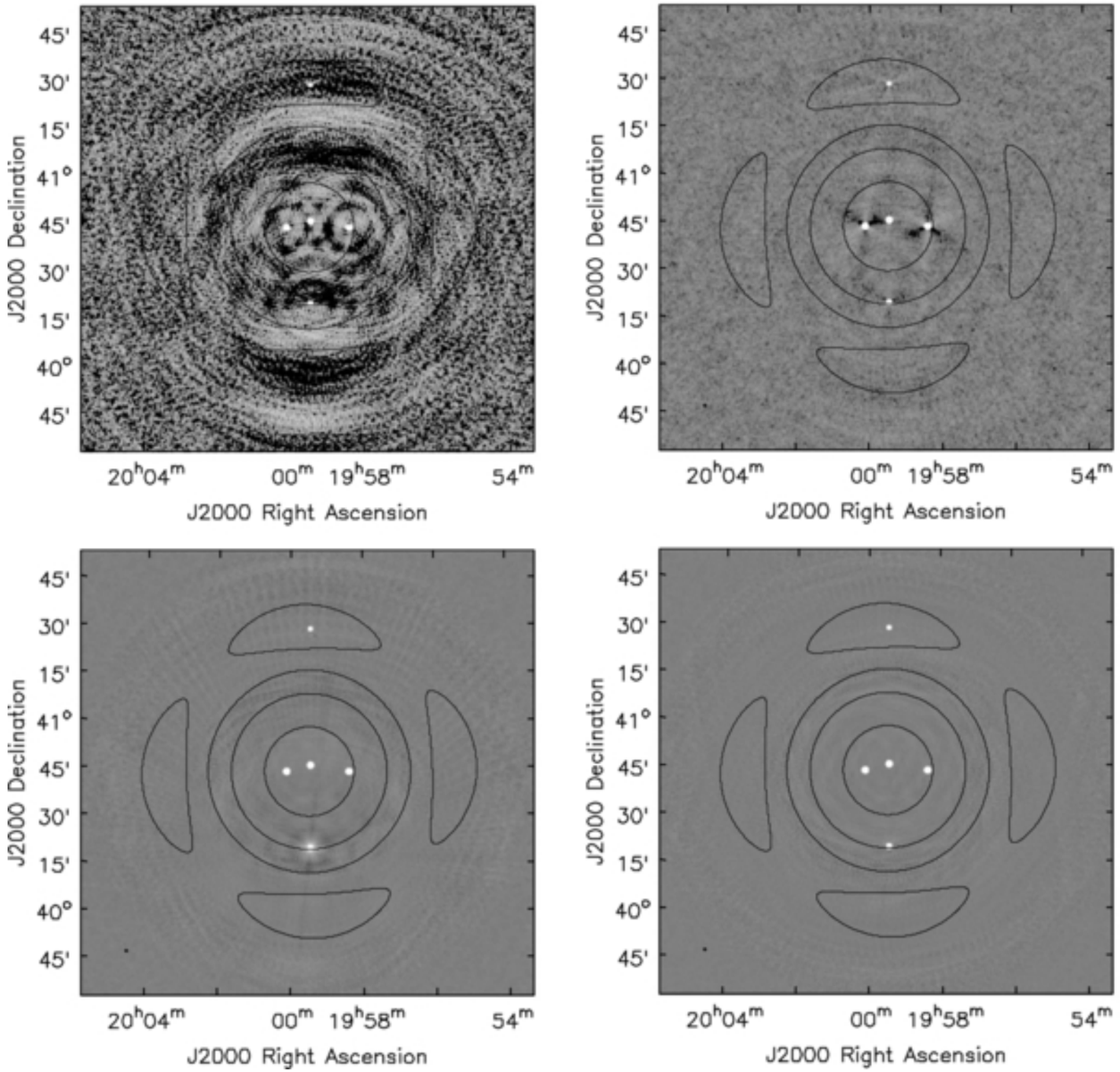
A-Projection (Bhatnagar et al. 2008): grid each visibility with the Hermitian conjugate aperture Mueller matrix. After the FFT the result is weighted by $|M_{ij}|^2$:

$$\frac{\mathcal{F}^\dagger \left[M_{ij}^{M\dagger} \star \mathbf{V}_{ij}^{\text{obs}} \right]}{|M_{ij}|^2} = \mathbf{I}$$

Two image-plane normalisation choices:

- Divide by one M^S: **flat noise** image
- Divide by $(M^S)^2$: **flat sky** image

A-Projection in action – Stokes I



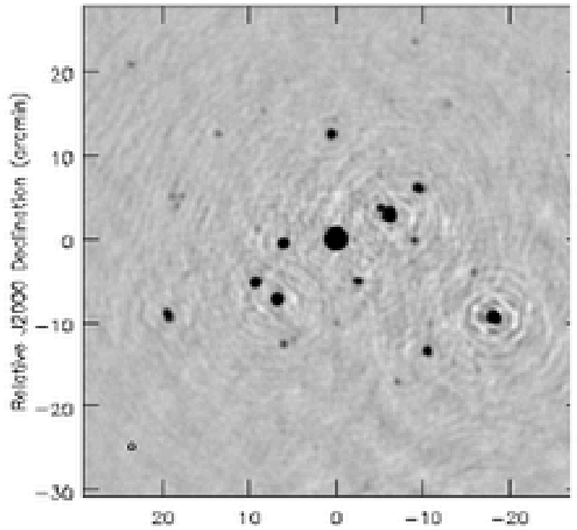
A-Projection in action – Stokes V

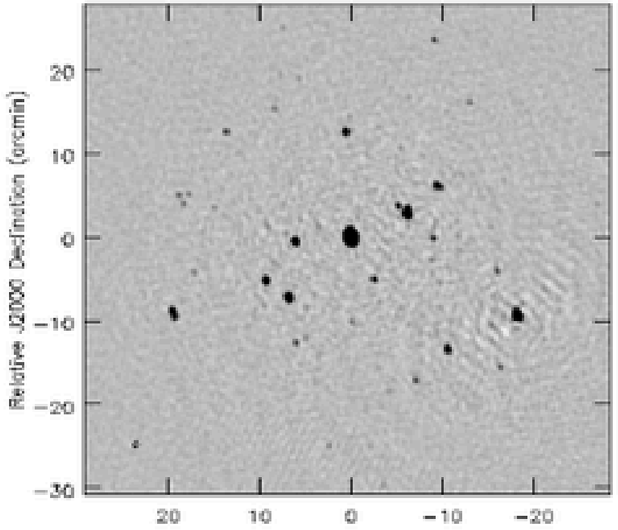
Before

After

Stokes I

Artifacts around all sources away from the pointing center



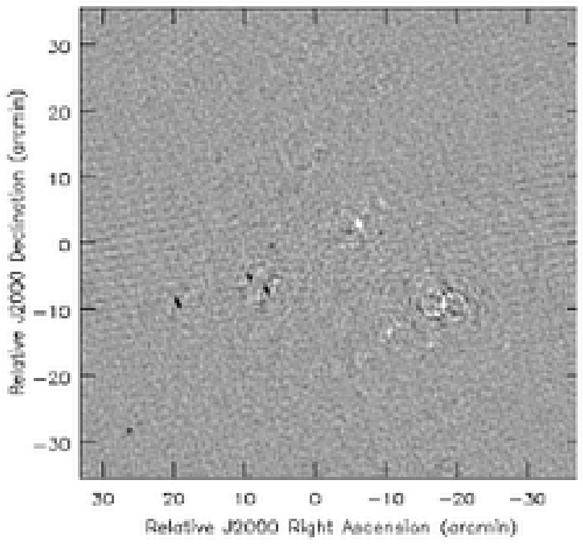


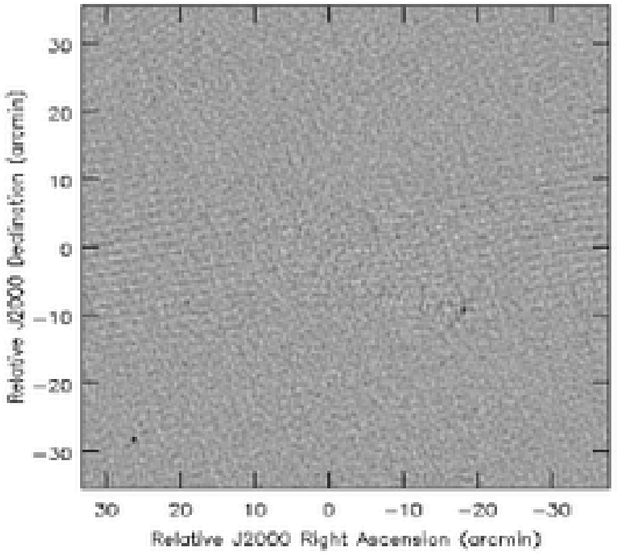
Stokes I

Artifacts removed or reduced within the main lobe

Stokes V

Artificial signals around bright sources due to beam squint



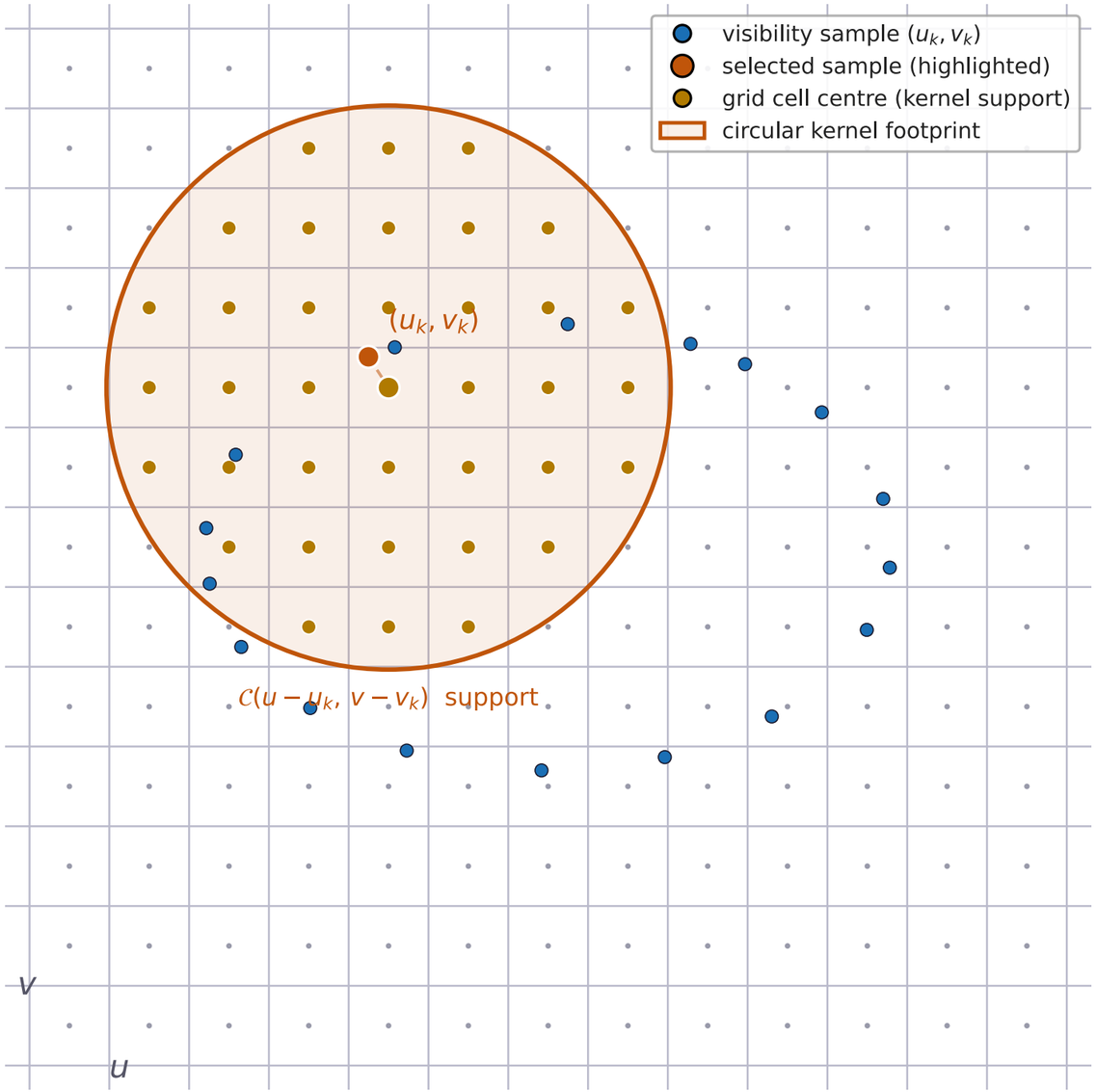


Stokes V

Instrumental Stokes V removed within the main lobe

Gridding and computational cost

Gridding: convolutional resampling



Visibilities live at arbitrary (u_k, v_k). The FFT requires a regular grid. Each visibility is distributed by convolution with a kernel:

$$\tilde{V}(u,v) = \sum_k V_k \cdot \mathcal{C}(u-u_k,\ v-v_k)$$

The kernel $\mathcal{C}$ spreads each sample across neighbouring grid cells. Its support width N_{sup} sets the cost:

$$\text{cost} \propto N_{\text{vis}} \times N_{\text{sup}}^2$$

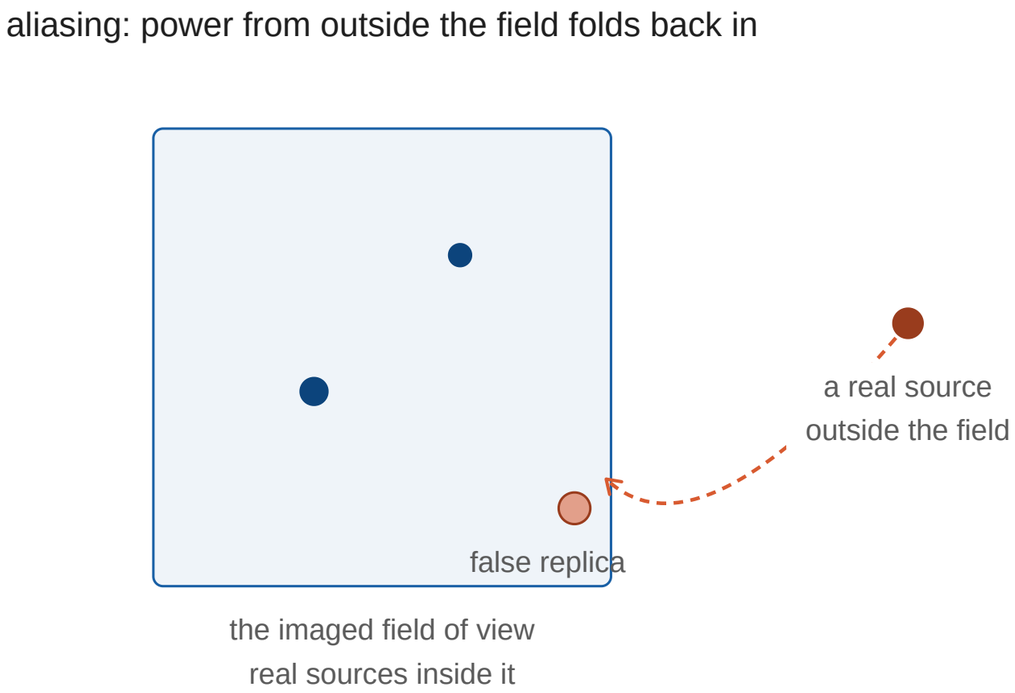
The kernel effect is divided out in the image plane after the FFT.

Different kernels address different problems:

- **Prolate spheroidal** – suppresses aliasing
- **Fresnel kernel** – corrects the w-term (W-Projection)
- **Aperture conjugate** – corrects the primary beam (A-Projection)

Aliasing

A finite grid aliases: power from outside the imaged field folds back in.

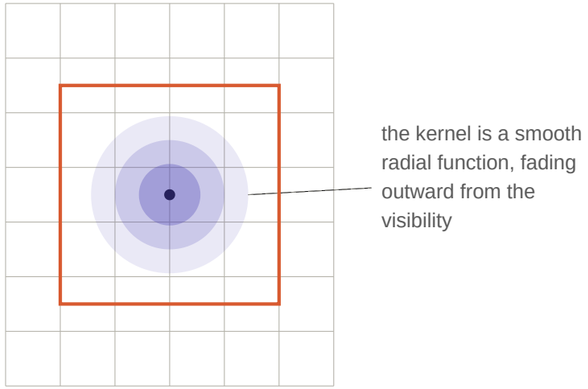


The prolate spheroidal kernel suppresses this – its Fourier response is maximally concentrated inside the field. Support approximately 3x3 pixels.

Kernel support and field of view

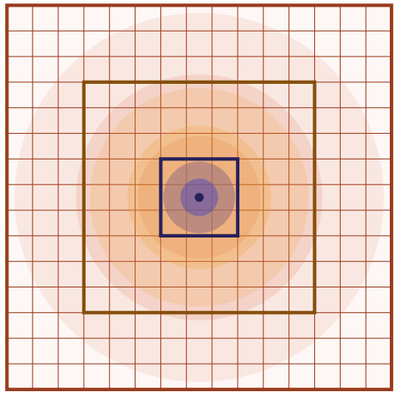
Larger kernel support lets you image further from the phase centre but costs more per visibility. Support size is set by how far out in the primary beam you need to go.

kernel and support: the cells one visibility writes to



the support is the square block of cells it writes to

here 4 x 4 = 16 cells — that cell count is the cost per visibility



W-term kernel
support 15 x 15
= 225 cells

A-term kernel
support 9 x 9 = 81 cells

spheroidal kernel
support 3 x 3 = 9 cells

Computational cost

Gridding cost scales as $N_{\text{vis}} \times N_{\text{sup}}^2$. For an interferometer like the VLA, $N_{\text{vis}} \sim 10^{10}$.

Correction	Support	N_{sup}^2	Relative cost
Prolate spheroidal	$\sim 3 \times 3$	9	$1\times$
A-term	$\sim 9 \times 9$	81	$\sim 9\times$
W-term	$\sim 200 \times 200$	40 000	$\sim 4000\times$
Faceting	$\sim 3 \times 3$ per facet	9 per facet	depends on number of facets

With 10^{10} visibilities and a 200-pixel w-kernel: $10^{10} \times 4 \times 10^4 = 4 \times 10^{14}$ kernel-pixel operations per gridding pass. Gridding is approximately 80% of total imaging cost.

Only choose the level of algorithmic complexity your science requires.

Mosaicking

When one pointing is not enough

If the source is larger than the primary beam, or sources are scattered over a wider area, a single pointing cannot cover the field.

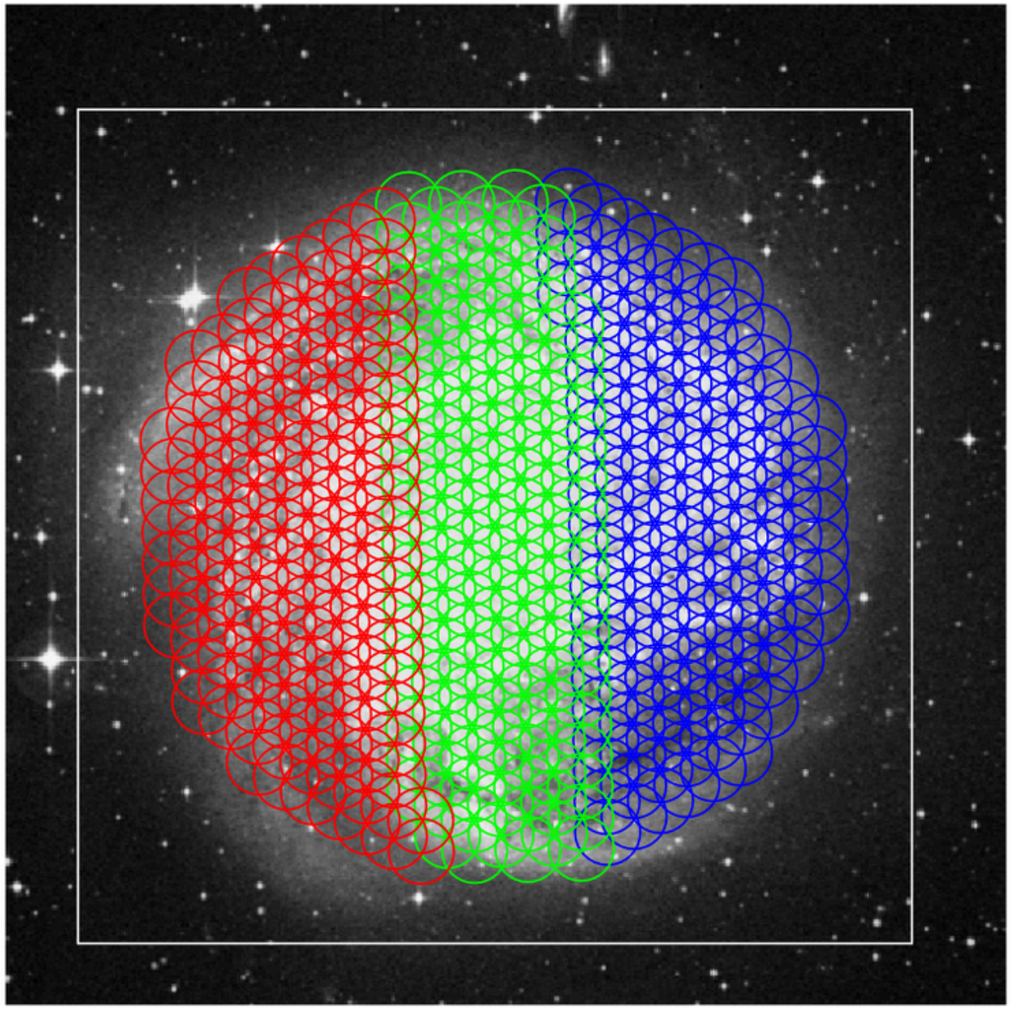
Mosaicking: combine K pointings, each centred at $\hat{\mathbf{p}}_k$, to image a field larger than the primary beam.

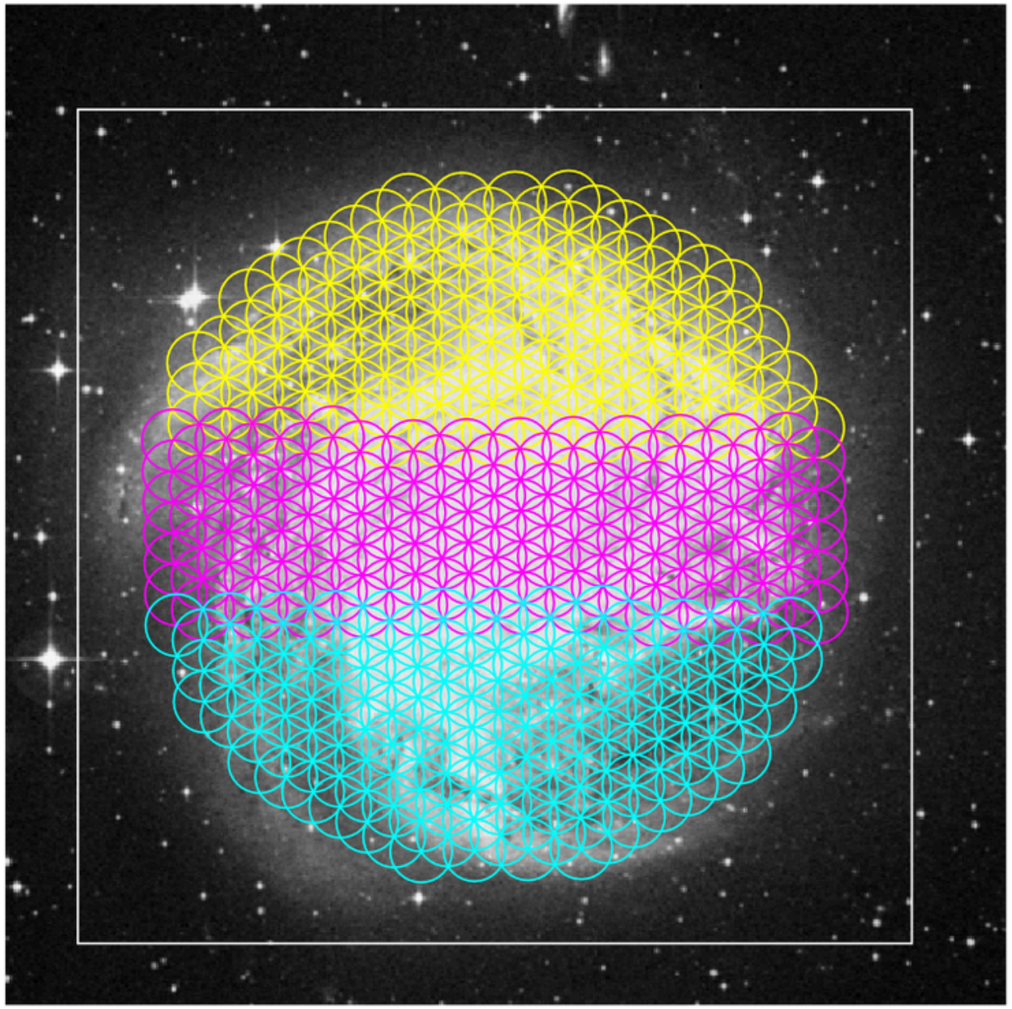
Each pointing multiplies the sky by its own primary beam centred at (l_k, m_k):

$$V_{pq}^{(k)}(u,v) = \iint A(\ell - l_k,\, m - m_k)\; I(\ell, m)\; e^{-2\pi i(u\ell + vm)}\; d\ell\, dm$$

The pointings are not independent – they are coupled through overlapping primary beam responses.

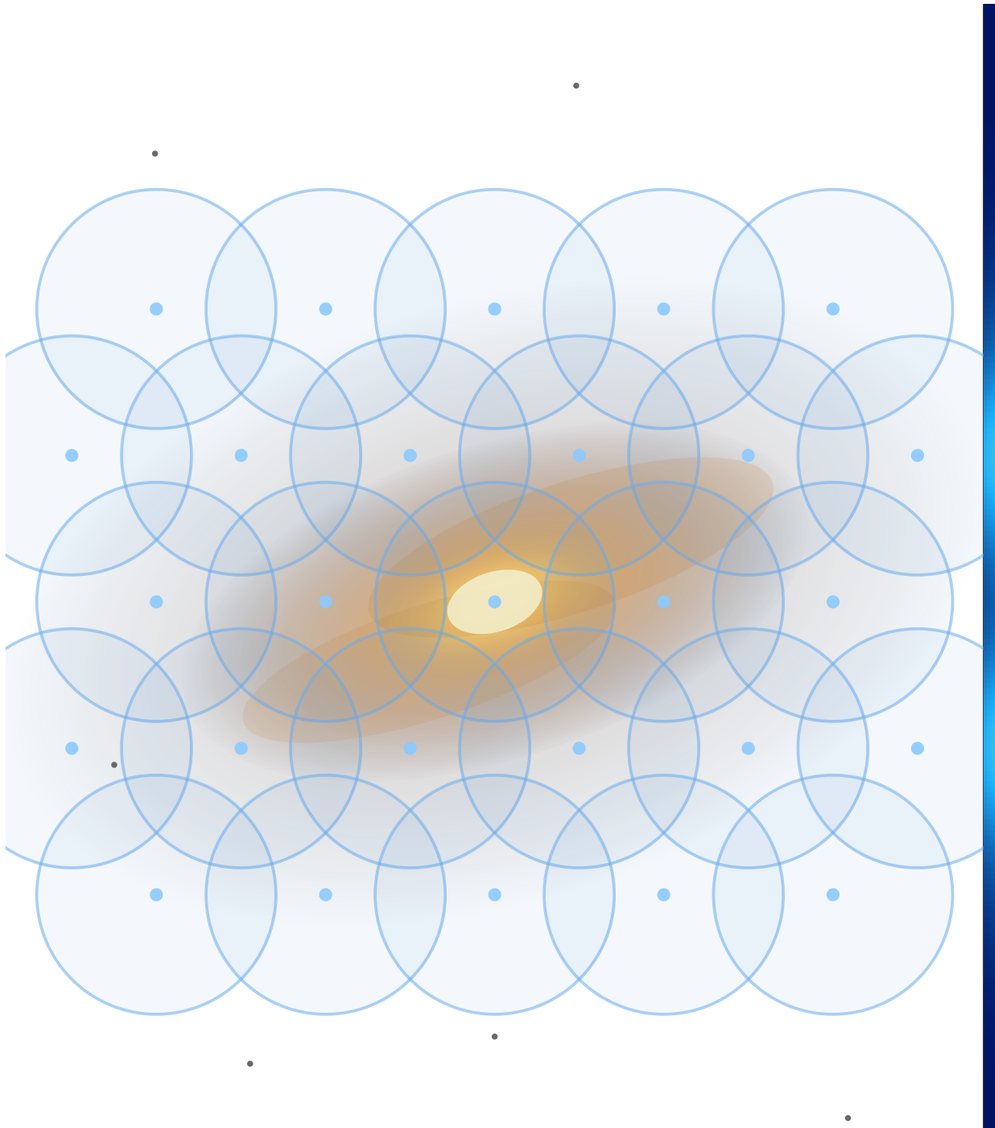
When one pointing is not enough

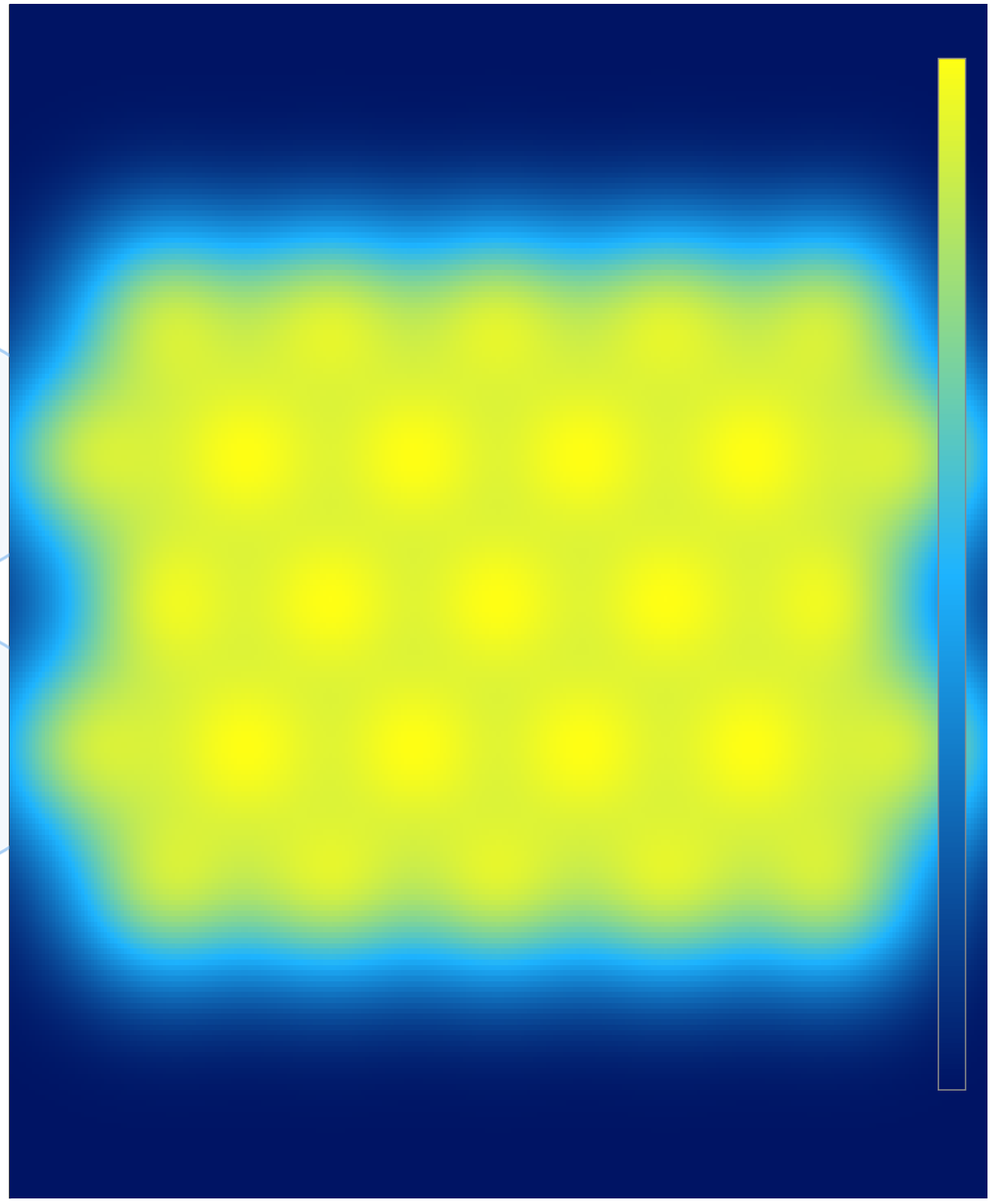




483 ALMA pointings covering M83. Koda et al. 2023.

Covering an extended source





Sensitivity across the mosaic

The sensitivity at any point on the sky is the sum of squared beam responses from all pointings that cover it:

$$Z(l,m) = \sum_k f^2(\theta_k(l,m)) \qquad f(\theta) = e^{-\theta^2/\sigma^2}$$

- k – pointing index; $\theta_k(l,m)$ – angular distance from pointing k to sky position (l,m)
- $f(\theta)$ – primary beam response (Gaussian approximation)
- $Z(l,m)$ – summed squared beam weight at (l,m)

The noise at that point is:

$$\sigma_F^2(l,m) = \frac{\sigma_D^2}{Z(l,m)}$$

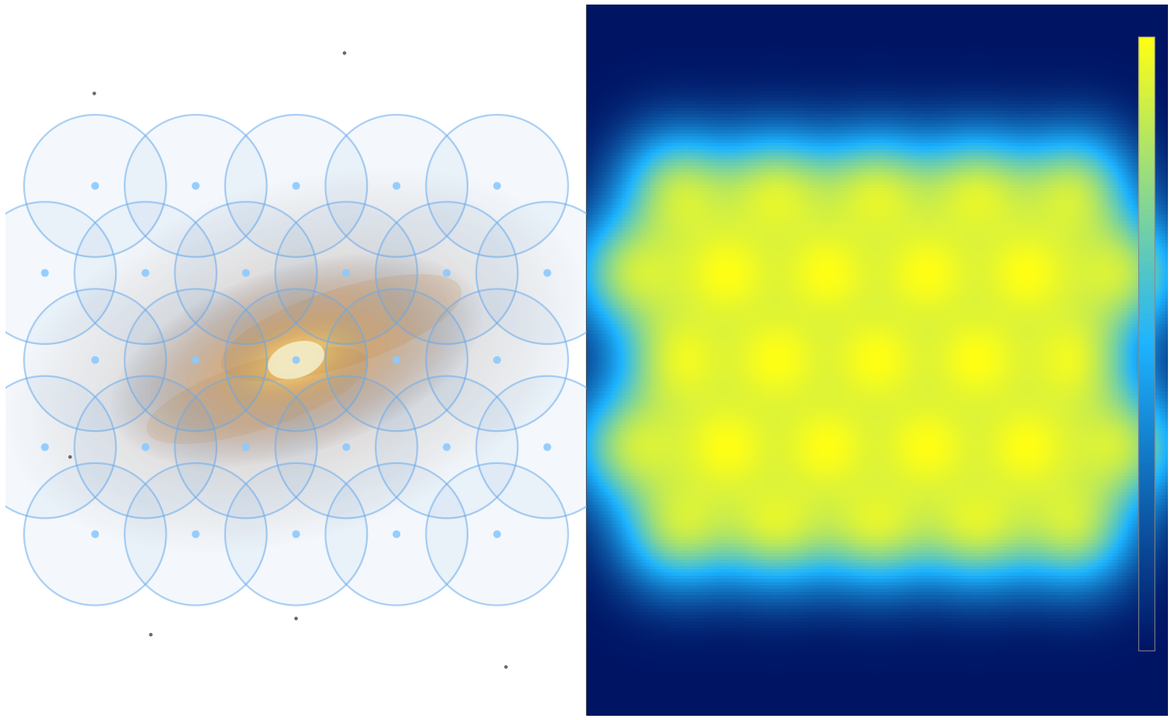
- σ_D^2 – per-visibility noise variance; noise falls where more pointings overlap

For a **hex pattern** with spacing $\theta_{\rm hex} = \theta_P/\sqrt{2}$ (where θ_P is the primary beam FWHM):

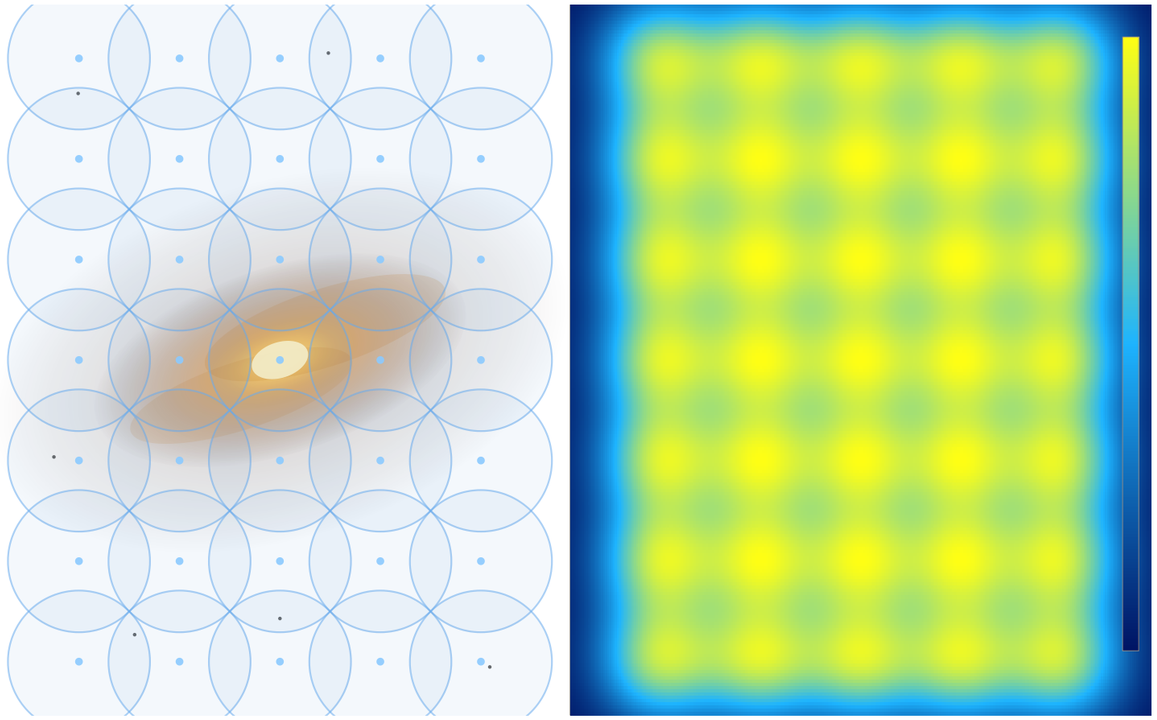
$$Z_{\max} = 1 + 6f^2(\theta_{\rm hex}) = 1.375 \qquad Z_{\min} = 3f^2\left(\theta_{\rm hex}/\sqrt{3}\right) = 1.191$$

Noise varies by only 15% across the covered field. Different patterns give different $Z(l,m)$ maps – and different noise uniformity.

Pointing patterns and sensitivity

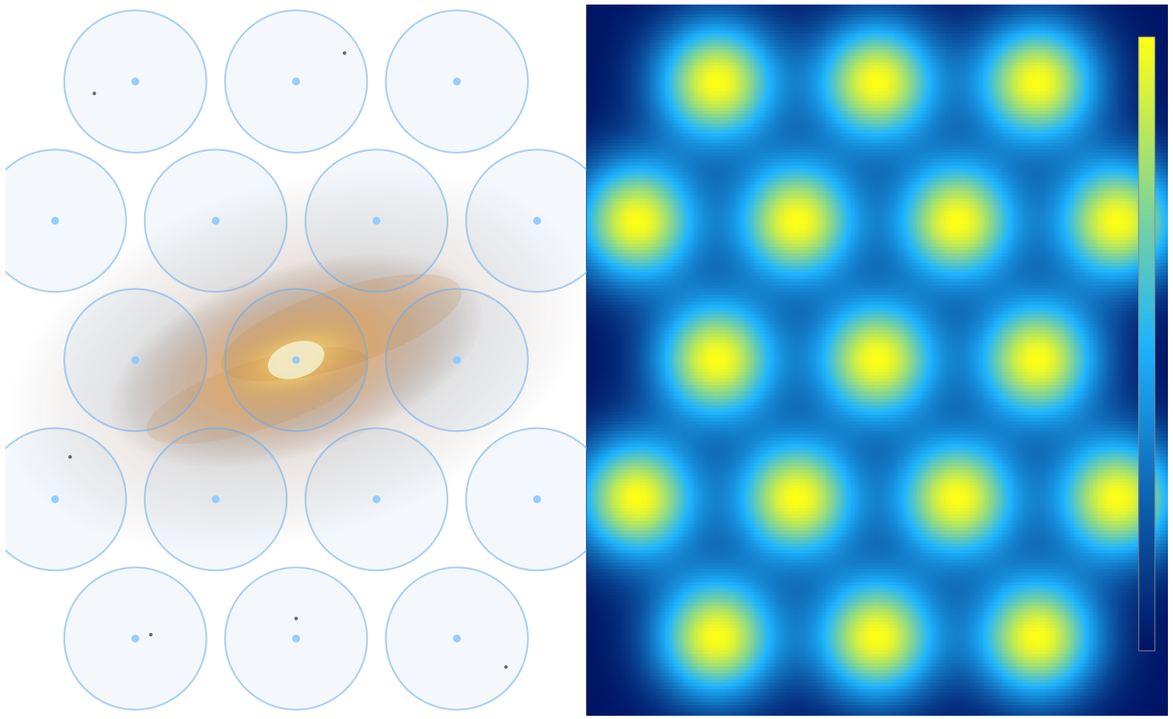


Hex – flat noise, optimal uniformity

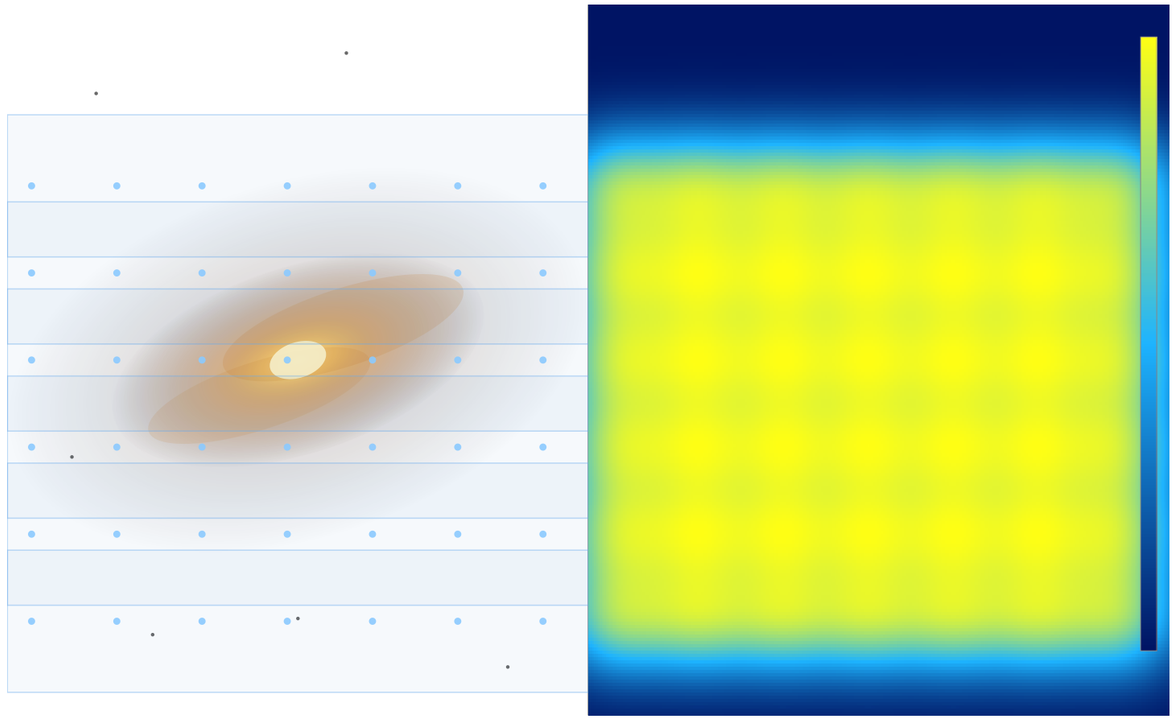


Square – sensitivity drops in corners

Pointing patterns and sensitivity



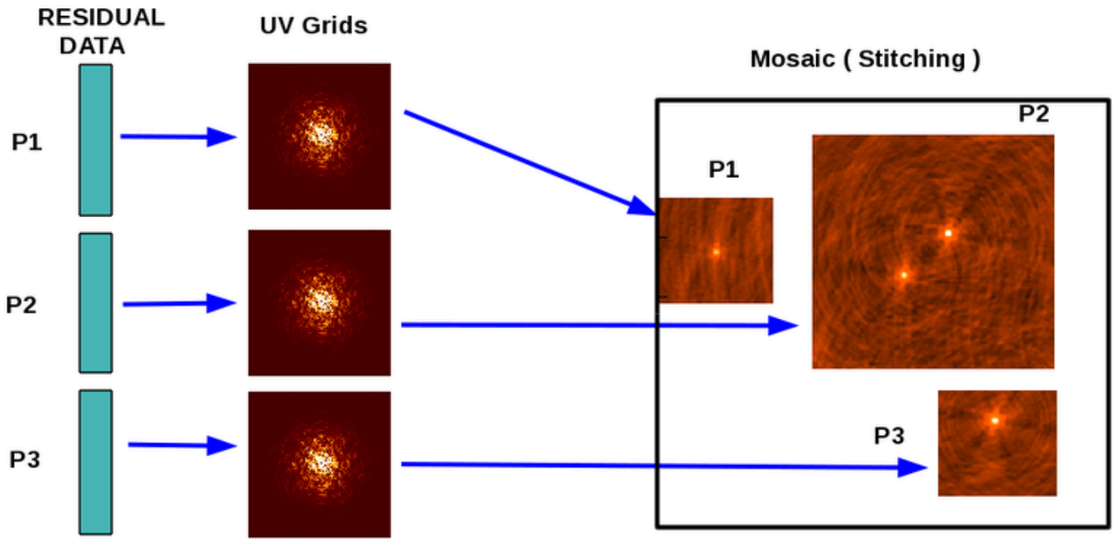
Triangular (sparse) – holes in coverage



On-the-fly strips – dots mark phase centres

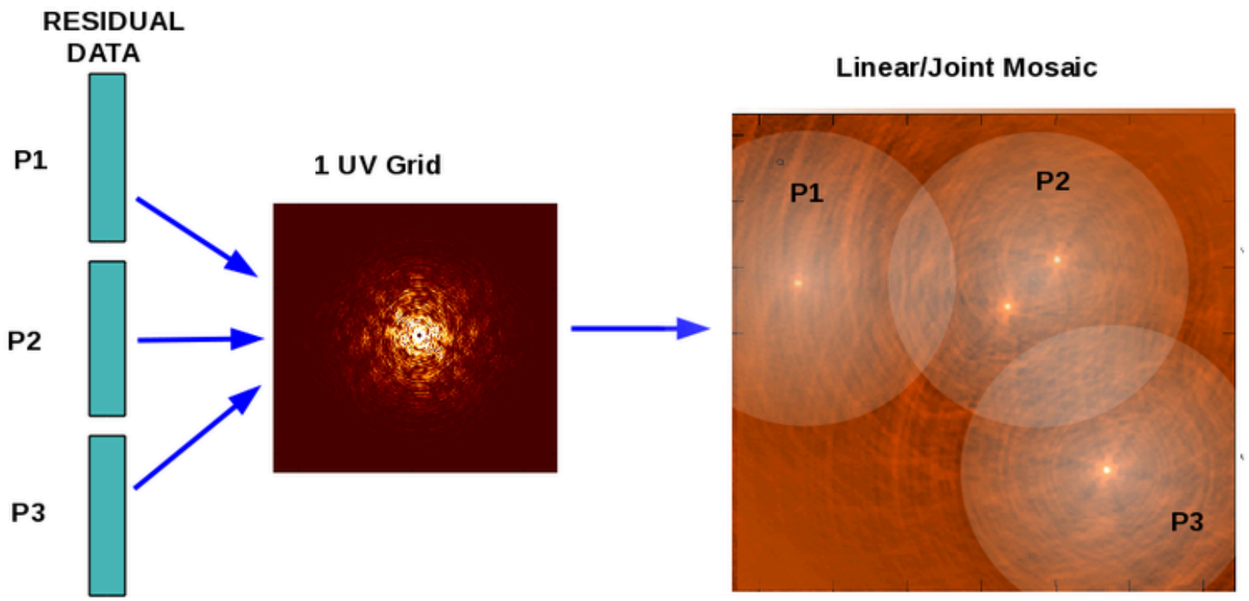
The pattern choice sets the noise map $Z(l, m)$; the combination strategy – image-plane or visibility-plane – sets how the beam responses across pointings are handled during deconvolution.

Image-plane mosaicking



Deconvolve each pointing independently, divide out its primary beam, co-add. Where PSFs differ significantly across pointings, sources near boundaries can produce ghosting in overlap regions.

Visibility-plane mosaicking



Adjacent pointings are gridded onto a single shared uv grid with a phase shift corresponding to the pointing separation. This produces a single PSF across the mosaic – provided the pointings have approximately similar uv coverage.

A sky shift $(\Delta l_k, \Delta m_k)$ is a phase ramp in the uv plane:

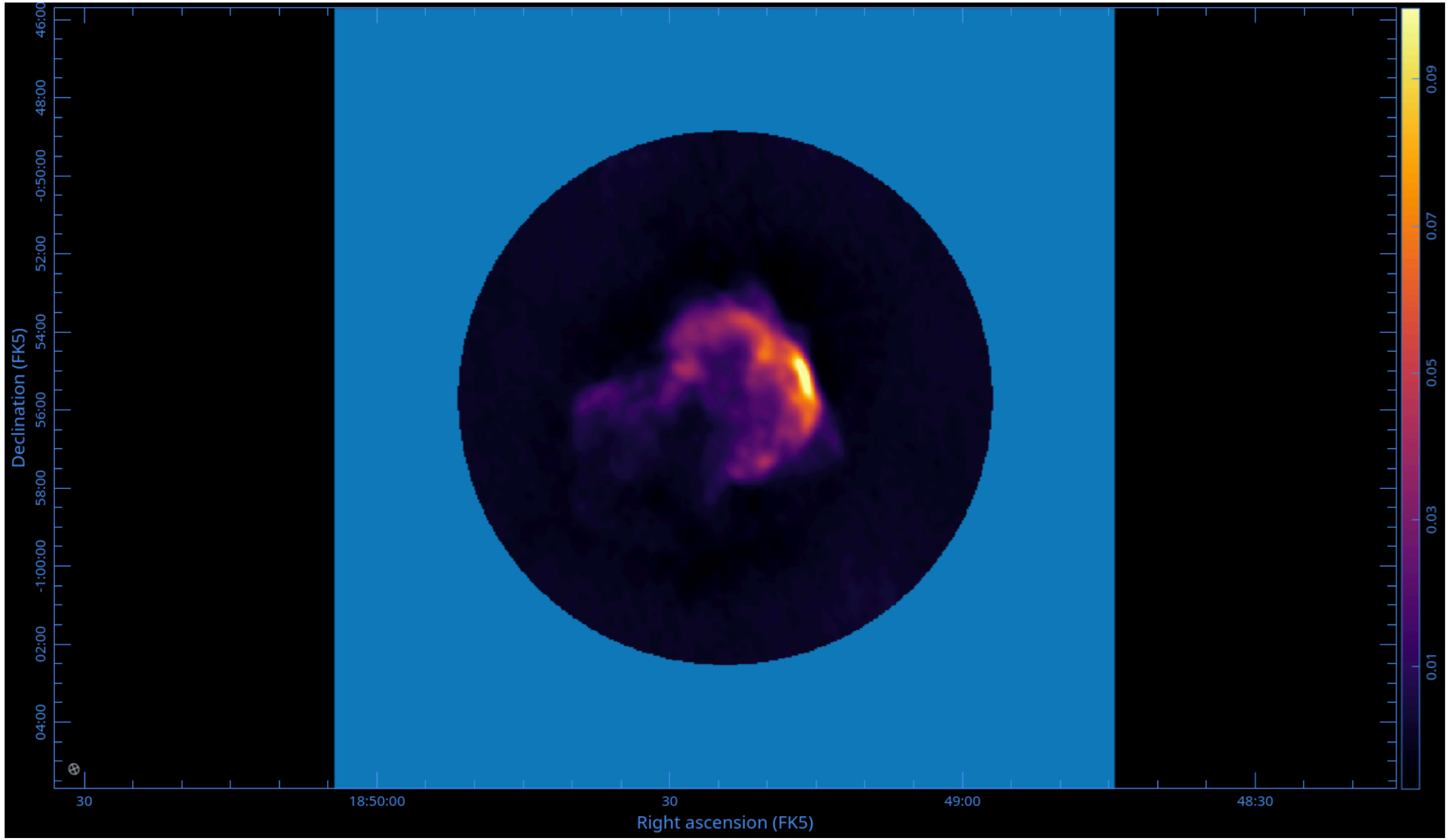
$$\tilde{V}(u,v) = \sum_k V_k(u,v)\; e^{2\pi i(u\,\Delta l_k + v\,\Delta m_k)}$$

So the gridding kernel for pointing k just picks up that phase:

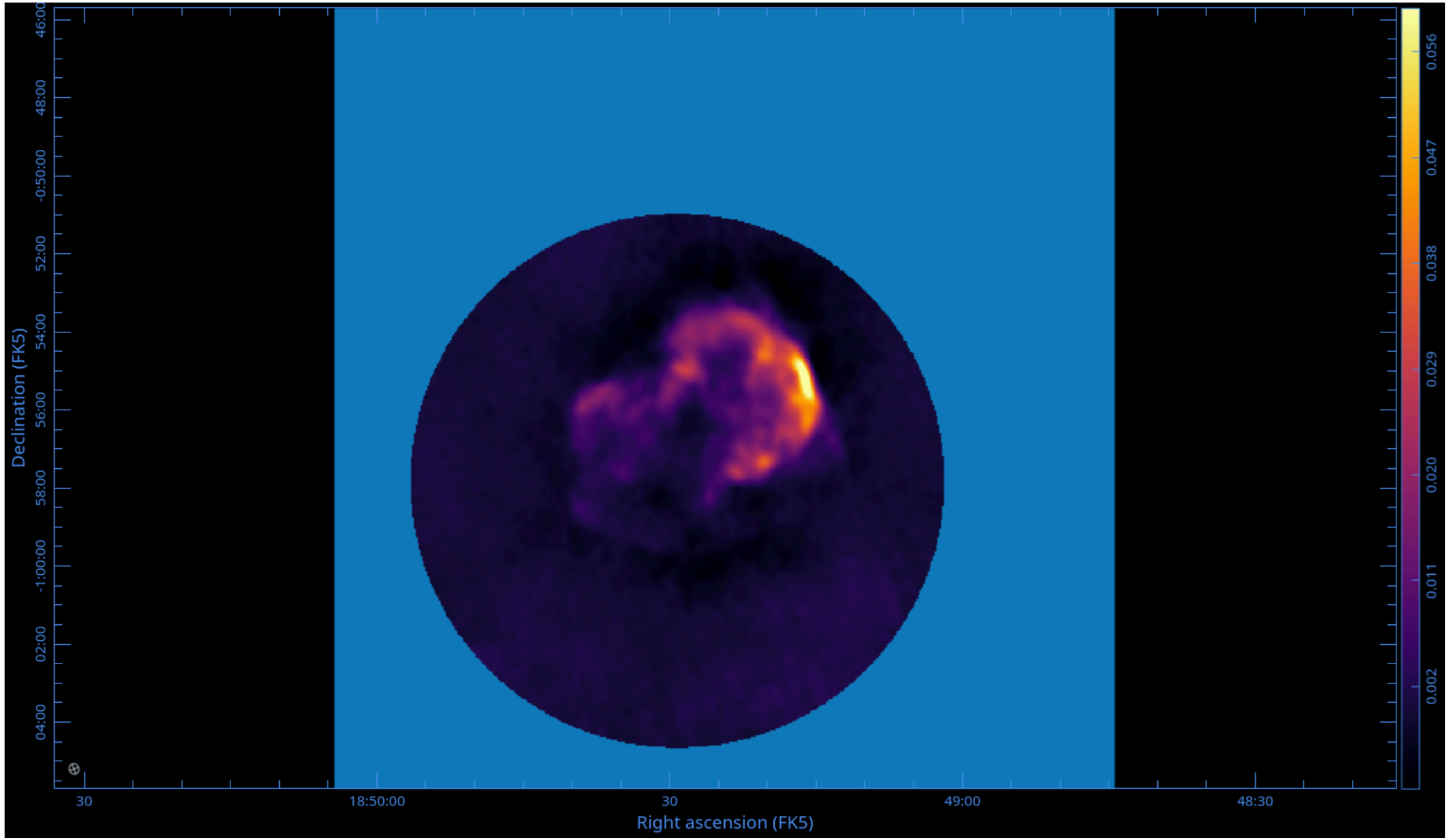
$$\mathcal{K}_k(u,v) = e^{i\phi_k}, \quad \phi_k = 2\pi(u\,\Delta l_k + v\,\Delta m_k)$$

For wideband mosaics with PB spectral index variation – see Urvashi's talk.

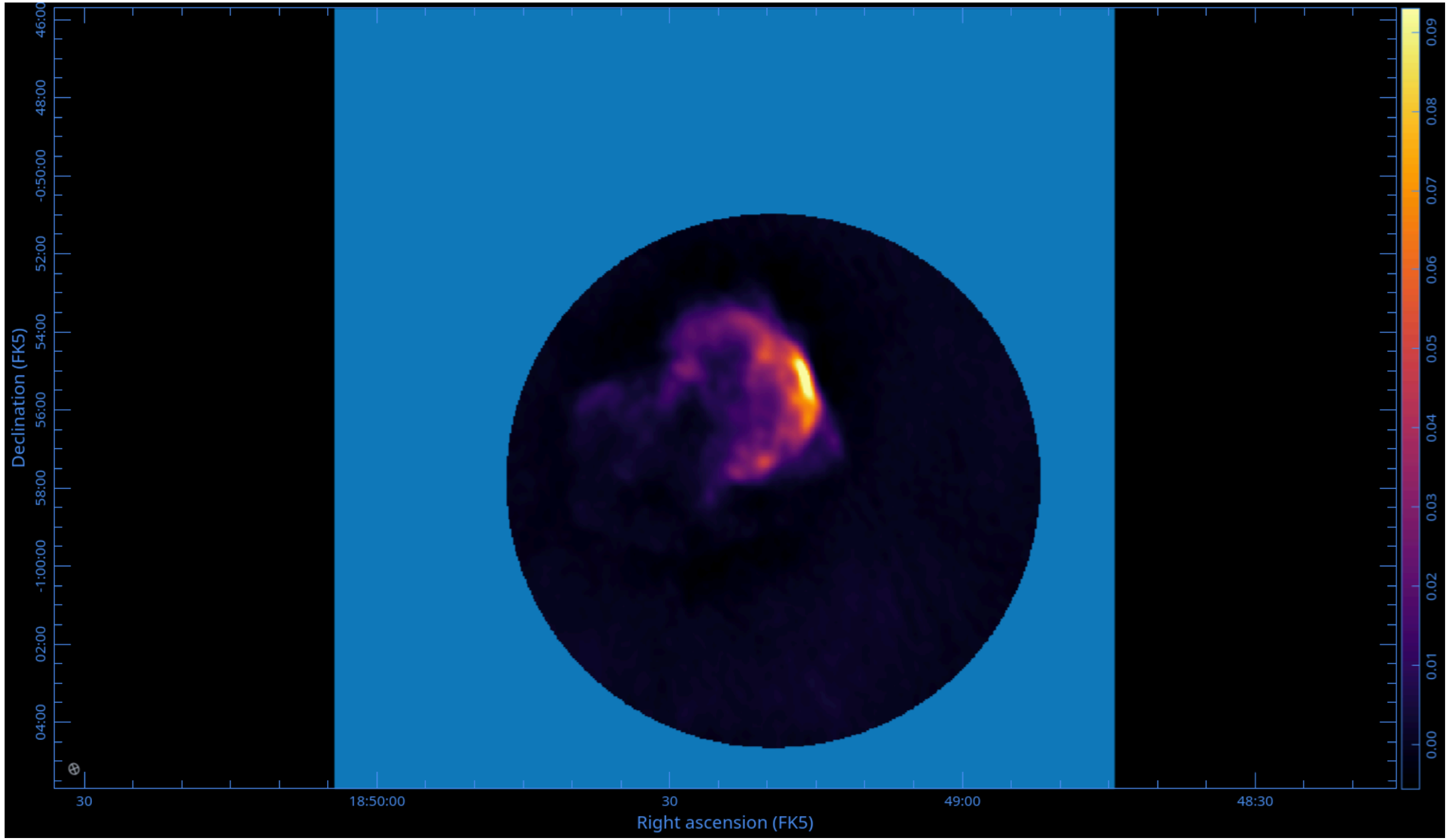
Linear mosaic: individual pointings



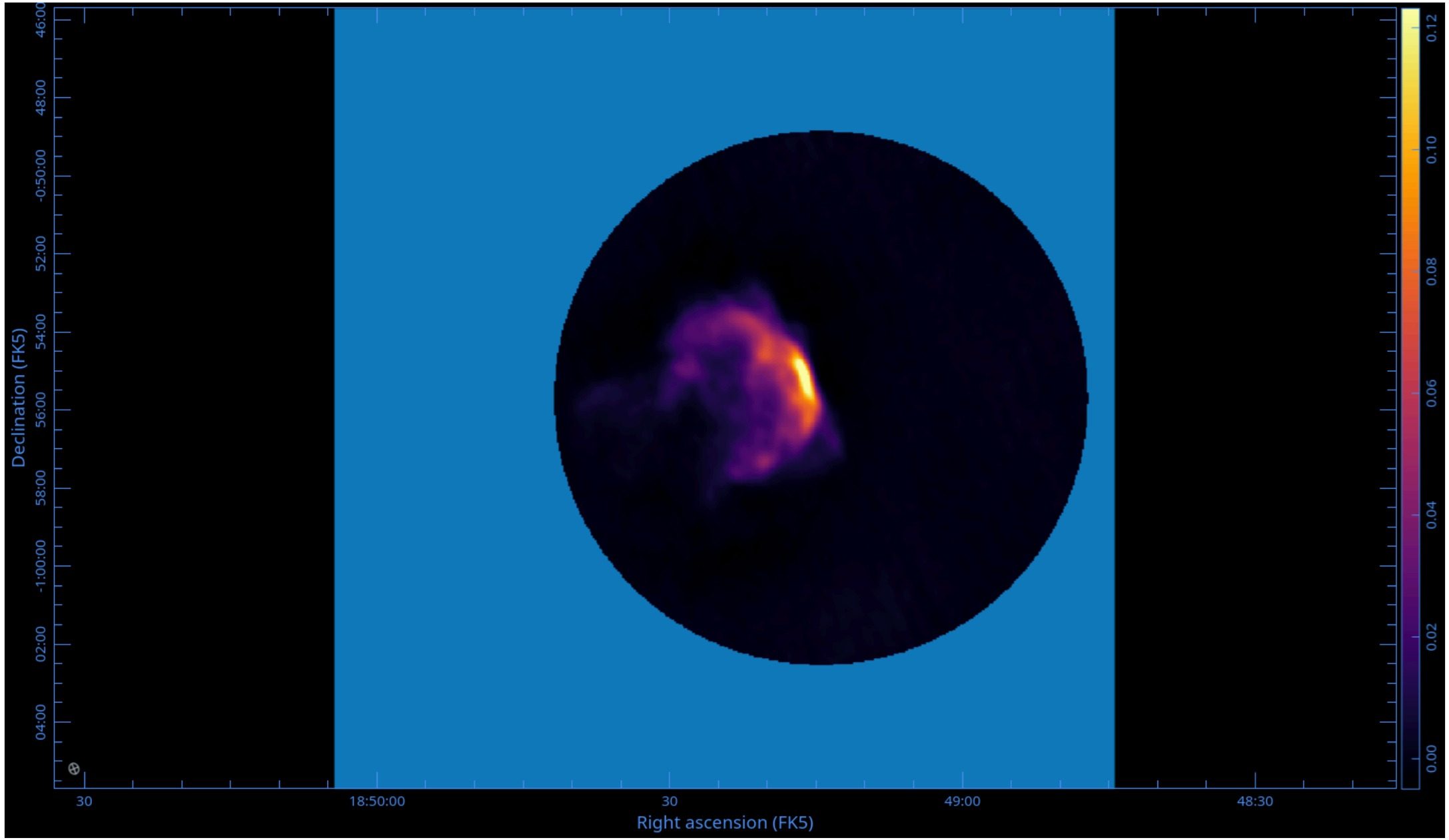
Linear mosaic: individual pointings



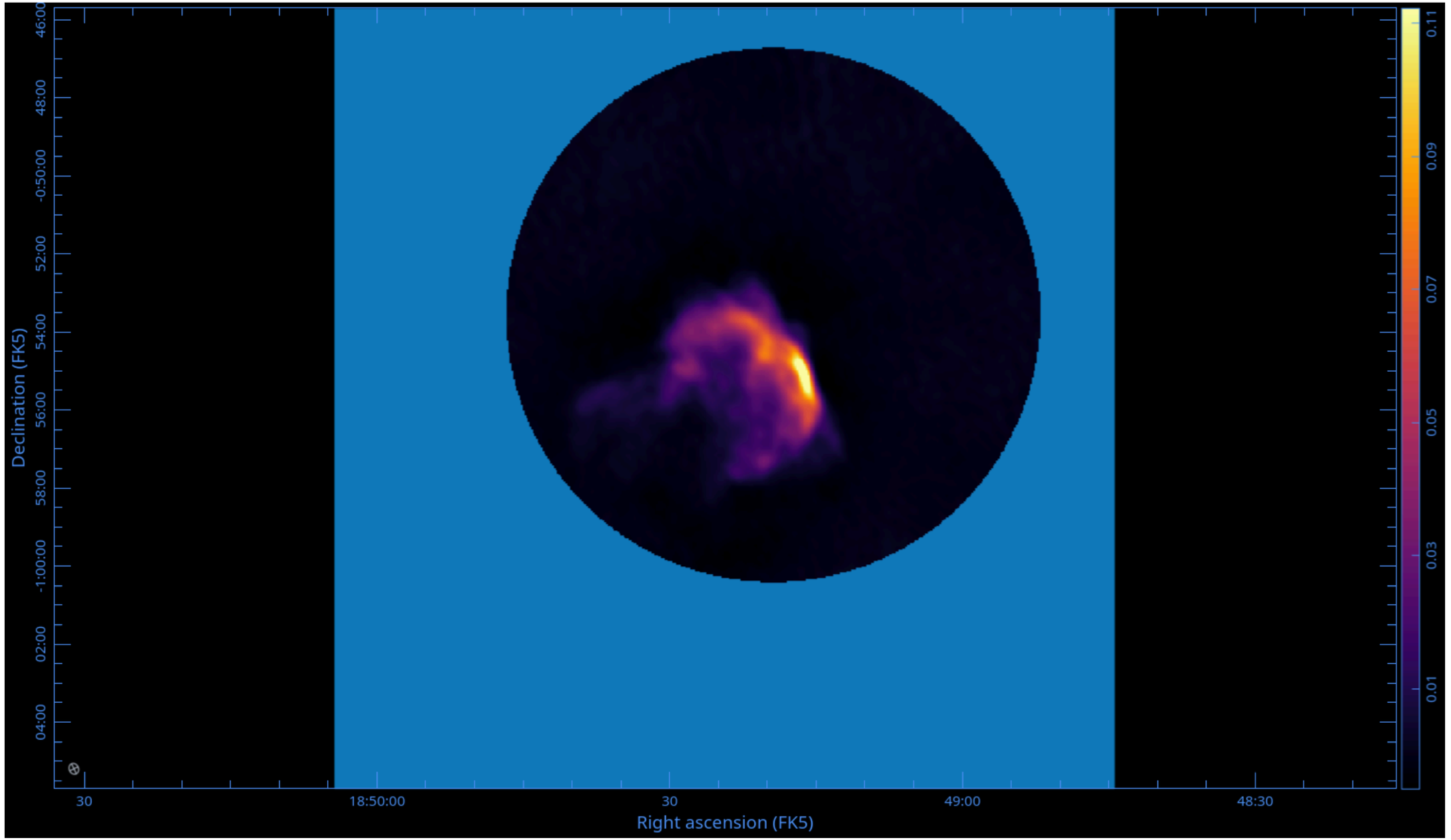
Linear mosaic: individual pointings



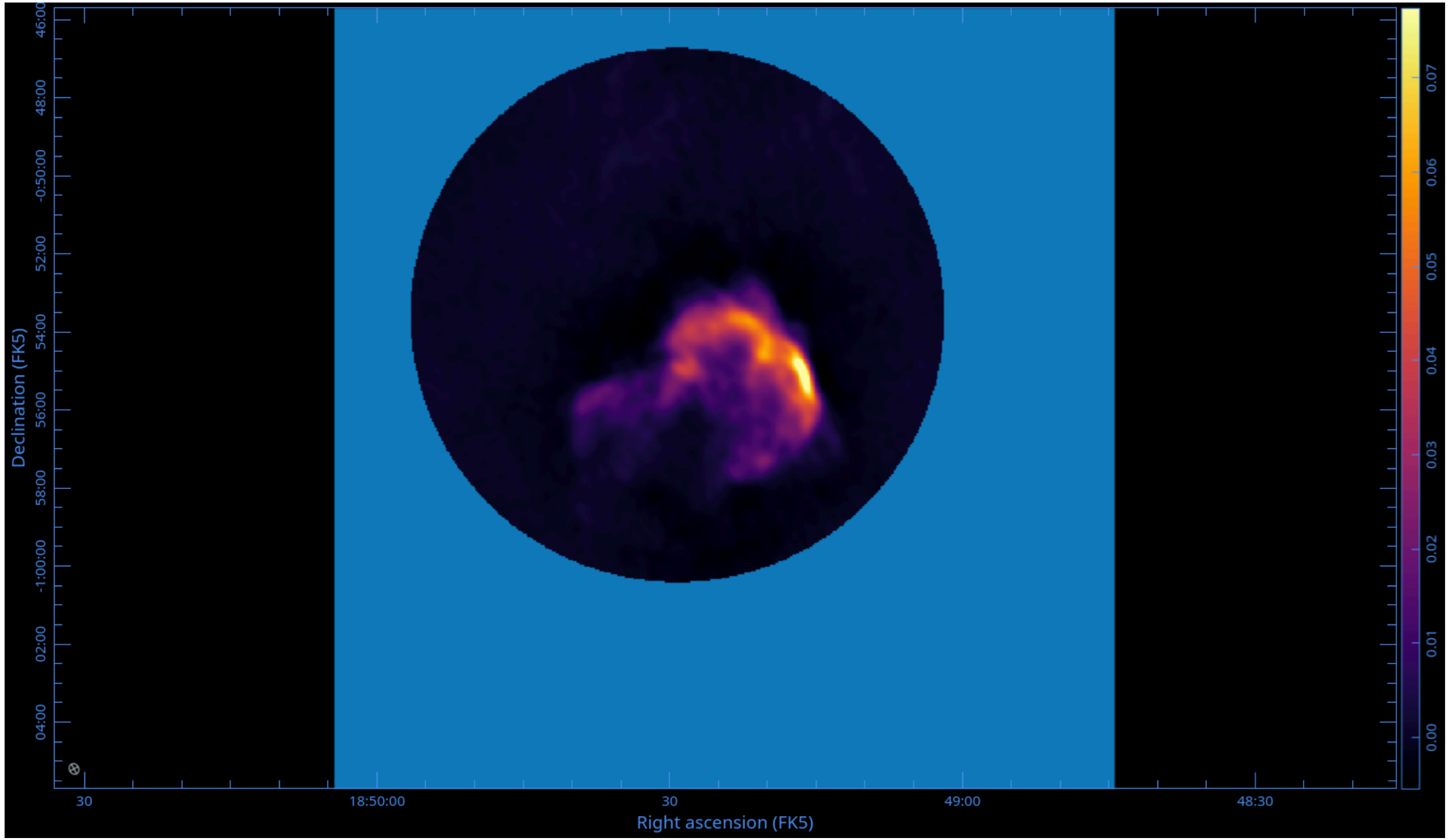
Linear mosaic: individual pointings



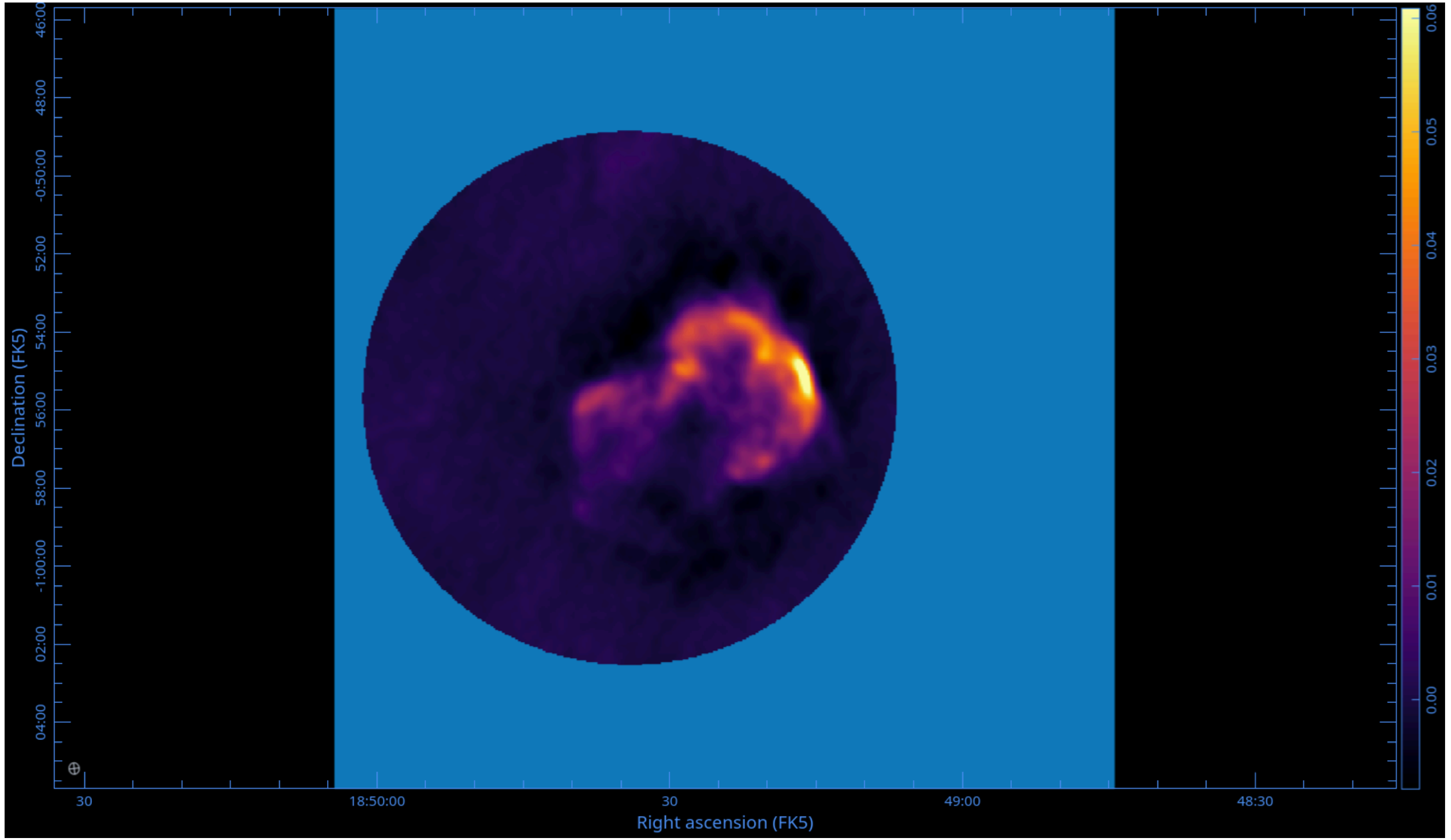
Linear mosaic: individual pointings



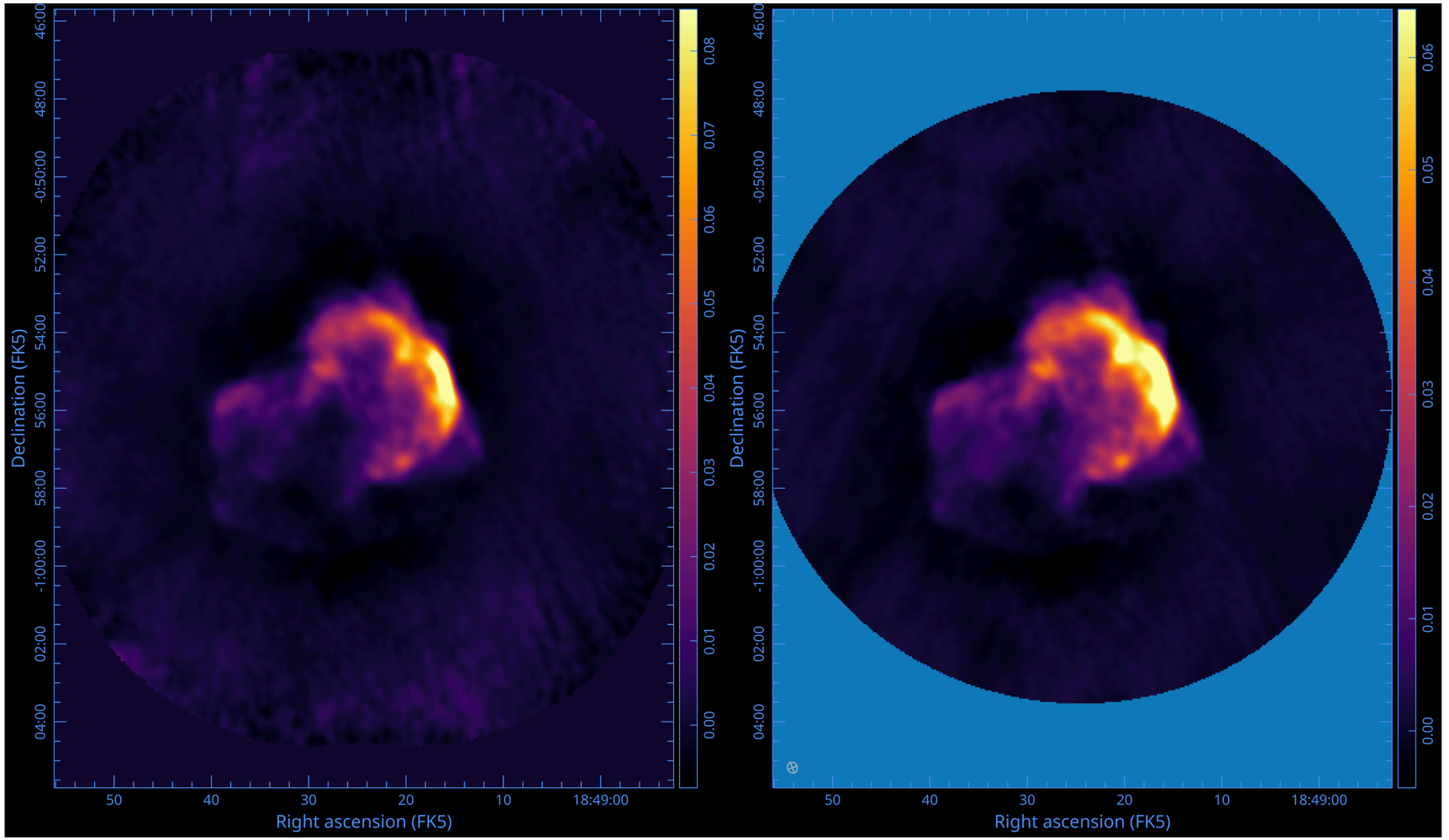
Linear mosaic: individual pointings



Linear mosaic: individual pointings



Linear mosaic vs joint UV mosaic



Left: linear mosaic – each pointing cleaned independently to a shallower depth, combined in the image plane. Right: joint UV mosaic – all pointings gridded onto a shared uv grid and deconvolved jointly to a deeper threshold. The joint mosaic recovers more extended emission and produces a uniform PSF across the field.

Summary

- The standard imaging equation assumes a flat sky and a static primary beam. Both assumptions fail for wide fields.
- The **w-term** is a per-baseline, per-direction phase. W-Projection and faceting both correct it – same result, different cost model.
- The **primary beam** changes with direction, time, and frequency. A-Projection corrects it during gridding by using a model aperture
- **Gridding cost** scales as $N_{\mathrm{vis}} \times N_{\mathrm{sup}}^2$. A 200-pixel w-kernel costs approximately 4000 times the prolate spheroidal. Only pay for what your science needs.
- **Mosaicking** couples pointings through overlapping primary beams. Image-plane co-addition after independent deconvolution is valid. Visibility-plane joint deconvolution is more complete.
- Residual pointing errors survive A-Projection and require a separate solve (see backup).

Thank you

Questions

Backup slides

Residual DDE after A-Projection

A-Projection corrects DDE known a priori: beam shape, rotation with parallactic angle.

What remains: **per-antenna mechanical pointing errors** – unknown, time-varying, up to a few arcminutes. At long integrations their effect is comparable in magnitude to the PA rotation term.

Antenna i with pointing offset l_i observes a shifted beam. The forward model A_{ij} is wrong – residual errors remain in the image even after A-Projection.

The ME in full matrix form

The observed visibility 4-vector in Stokes coordinates:

$$\mathbf{V}_{ij}^{\mathrm{obs}} = M_{ij}^{\mathrm{DI}} \left(M_{ij}^{S} \otimes (M_{ij}^{S})^{*} \right) \mathbf{V}_{ij}^{\mathrm{sky}}$$

The direction-dependent Mueller matrix:

$$M_{ij}^{S}(l, m; \nu, t) = J_i(l, m; \nu, t) \otimes J_j^{*}(l, m; \nu, t)$$

where the Jones chain for antenna i is:

$$J_i = G_i \cdot D_i \cdot P_i(\chi_i)$$

- G_i – complex gain (direction-independent, feeds calibration)
- D_i – polarization leakage (D-terms)
- $P_i(\chi_i)$ – parallactic angle rotation matrix

Full calibration matrix: direction-independent terms

The direction-independent calibration equation:

$$\mathbf{V}_{ij}^{\text{obs}} = (G_i \otimes G_j^*)\, \mathbf{V}_{ij}^{\text{model}} + \mathbf{n}_{ij}$$

In the 2×2 Jones/correlation matrix form:

$$\begin{pmatrix} V^{pp} \\ V^{pq} \\ V^{qp} \\ V^{qq} \end{pmatrix}_{ij}^{\text{obs}} = \underbrace{\begin{pmatrix} g_i^p g_j^{p*} & 0 & 0 & 0 \\ 0 & g_i^p g_j^{q*} & 0 & 0 \\ 0 & 0 & g_i^q g_j^{p*} & 0 \\ 0 & 0 & 0 & g_i^q g_j^{q*} \end{pmatrix}}_{G_i \otimes G_j^*} \begin{pmatrix} V^{pp} \\ V^{pq} \\ V^{qp} \\ V^{qq} \end{pmatrix}_{ij}^{\text{model}}$$

Diagonal structure: direction-independent gains do not couple feed hands.

Full calibration matrix: direction-dependent beam

Including the primary beam $E_i(l, m)$ inside the integral:

$$V_{ij}^{ab,\text{obs}}(\nu, t) = \iint \sum_{c,d} E_i^{ac}(l, m; \nu, t)\, I^{cd}(l, m)\, E_j^{bd*}(l, m; \nu, t)\, e^{-2\pi i(ul+vm+w(n-1))}\, \frac{dl\, dm}{n}$$

In Mueller form (Stokes basis):

$$\mathbf{V}_{ij}^{\text{obs}} = \iint M_{ij}^{S}(l, m; \nu, t)\, \mathbf{I}(l, m)\, e^{-2\pi i(ul+vm+w(n-1))}\, \frac{dl\, dm}{n}$$

where $M_{ij}^{S} = \mathcal{T}^{-1}(J_i \otimes J_j^*)\mathcal{T}$ and $\mathcal{T}$ is the feed-to-Stokes transformation.

Full 4×4 matrix M_{ij}^{S} has non-zero off-diagonal terms when the beam is not circularly symmetric or is frequency-dependent.

Pointing SelfCal – the idea

(Bhatnagar and Cornwell 2017)

A pointing error is a **phase gradient across the aperture** – just 2 parameters per antenna (azimuth and elevation offset l_i).

Parametrize A_{ij} as a function of l_i and minimize χ^2:

$$\frac{\partial \chi^2}{\partial l_i^*} = -2 \sum_{j \neq i} \mathrm{Re}\left(R_{ij}^{\dagger} \cdot W_{ij} \cdot \nabla_i V_{ij}^{M} \right)$$

The gradient $\nabla_i A_{ij} = \partial A_{ij} / \partial l_i$ is analytic.

where E is the antenna far-field electric field pattern. The solver signal is $\left| \partial E / \partial l \otimes E^* \right|$ – a double-hump, zero at the PB centre, **peaking at the half-power point**. Flux spread across the FoV drives the solution. SNR scales as $\sqrt{\Delta\nu \, \tau_{\mathrm{sol}} \, (N_{\mathrm{ant}} - 1)}$.

Algorithm: image with A-Projection $\rightarrow$ solve for $l_i \rightarrow$ re-image. Iterate.