EVLA Data Reduction Workshop
Oct. 2017, Socorro

Wide-band Wide-field Imaging

S. Bhatnagar

U. Rau, K. Golap, J. Robnett, P. Jagannathan

Why new algorithms?

- Instantaneous wide-band capability of the EVLA is the single dominant parameter that enables new scientific capabilities

$$Noise \propto \frac{T_{sys}}{A_{eff}\sqrt{\Delta\nu \, \Delta T}}$$

  - More instantaneous information about the emission
    - Spectral Index, RM,...

- $$V_{ij}(\nu) = G_{ij}^{DI} \; W_{ij} \int \underbrace{P_{ij}(\mathbf{s},\nu,t) \; I(\mathbf{s},\nu) \; e^{\iota \mathbf{s}.\mathbf{b}_{ij}}}_{\text{Direction Dependent (DD) terms}} \; d\mathbf{s}$$

  - Terms inside the integral cannot be accounted-for before imaging
    - Conventional imaging ignores DD terms
    - Also ignores time, frequency and polarization dependence
  - Solutions: Project-out the effects during imaging + model frequency dependence of the sky during deconvolution
  - Or resort to spectral cube imaging + image-plane corrections/averaging

Why new algorithms?

- Instantaneous wide-band capability of the EVLA is the single dominant parameter that enables new scientific capabilities

$$Noise \propto \frac{T_{sys}}{A_{eff}\sqrt{\Delta\nu\,\Delta T}}$$

  - More instantaneous information about the emission
    - Spectral Index, RM,...

- $$V_{ij}(\nu) = G_{ij}^{DI}\; W_{ij} \int P_{ij}(\mathbf{s},\nu,t)\; I(\mathbf{s},\nu)\; e^{\iota\, \mathbf{s}.\mathbf{b}_{ij}}\; d\,\mathbf{s}$$

PB Effects (A-Projection)

Direction Dependent (DD) terms

  - Terms inside the integral cannot be accounted-for before imaging
    - Conventional imaging ignores DD terms
    - Also ignores time, frequency and polarization dependence
  - Solutions: Project-out the effects during imaging + model frequency dependence of the sky during deconvolution
  - Or resort to spectral cube imaging + image-plane corrections/averaging

Why new algorithms?

- Instantaneous wide-band capability of the EVLA is the single dominant parameter that enables new scientific capabilities

$$Noise \propto \frac{T_{sys}}{A_{eff}\sqrt{\Delta\nu\,\Delta T}}$$

  - More instantaneous information about the emission
    - Spectral Index, RM,...

Spectral Index effects (MT-MFS)

- $$V_{ij}(\nu) = G_{ij}^{DI}\; W_{ij} \int \underbrace{P_{ij}(\mathbf{s},\nu,t)\; I(\mathbf{s},\nu)\; e^{\iota\, \mathbf{s}.\mathbf{b}_{ij}}}_{\text{Direction Dependent (DD) terms}}\; d\,\mathbf{s}$$

  - Terms inside the integral cannot be accounted-for before imaging
    - Conventional imaging ignores DD terms
    - Also ignores time, frequency and polarization dependence
  - Solutions: Project-out the effects during imaging + model frequency dependence of the sky during deconvolution
  - Or resort to spectral cube imaging + image-plane corrections/averaging

Why new algorithms?

- Instantaneous wide-band capability of the EVLA is the single dominant parameter that enables new scientific capabilities

$$Noise \propto \frac{T_{sys}}{A_{eff}\sqrt{\Delta\nu\,\Delta T}}$$

  - More instantaneous information about the emission
    - Spectral Index, RM,...

Non co-planar baselines (W-term)
(W-Projection)

- $$V_{ij}(\nu) = G_{ij}^{DI}\; W_{ij} \int \underbrace{P_{ij}(s,\nu,t)\; I(s,\nu)\; e^{\iota\, s.b_{ij}}}_{\text{Direction Dependent (DD) terms}}\; ds$$

Direction Dependent (DD) terms

  - Terms inside the integral cannot be accounted-for before imaging
    - Conventional imaging ignores DD terms
    - Also ignores time, frequency and polarization dependence
  - Solutions: Project-out the effects during imaging + model frequency dependence of the sky during deconvolution
  - Or resort to spectral cube imaging + image-plane corrections/averaging

Plan

- Wide-band Imaging
  - Account for frequency dependent sky brightness distribution
  - Algorithm: Multi-term Multi-Frequency Synthesis (MT-MFS, MS-MFS)

[Rau & Cornwell, A&A, 2011]

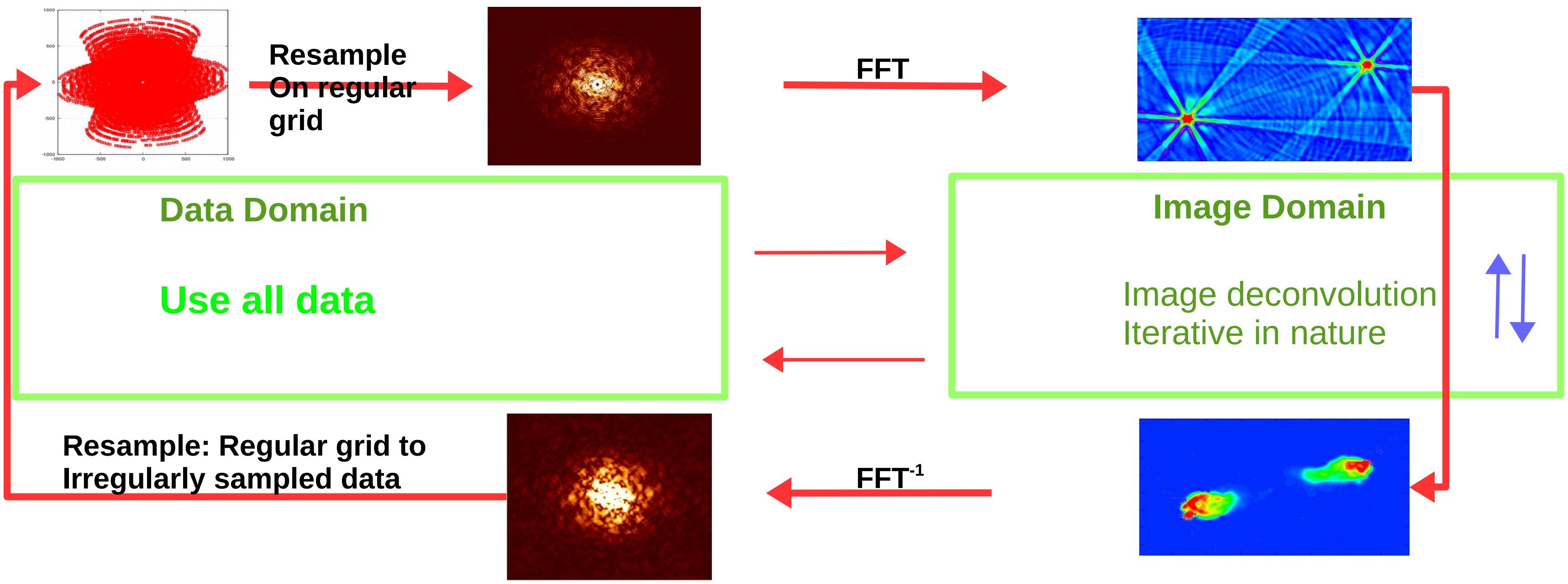
- Wide-field Imaging: Includes any effect that increases with R
  - Non co-planar baseline effect (W-term) **[Cornwell, Golap, Bhatnagar, Proc. IEEE, 2009]**
  - Effect of antenna PB: Time- and Poln.-dependence **[Bhatnagar et al., A&A, 2008]**
  - Algorithm: W-Projection, (WB) A-Projection

[General review: Rau et al., Proc. IEEE, V. 97 (8) 2009]

- Wide-band Wide-field Imaging
  - All of the above + PB frequency dependence
  - Algorithm: MT-MFS + (WB) AW-Projection (+ Mosaic) **[Bhatnagar et al., A&A, 2012]** **[WB Mosaic: Rau, Bhatnagar In prep.**
- Full-polarization imaging, Computing load and solutions

Imaging & Deconvolution: A recap

- Compute residuals using the original data
  - Needs Gridding and de-Gridding during major-cycle iterations



W-Projection

A-Projection

WB A-Projection

Standard gridding

CS-Clean

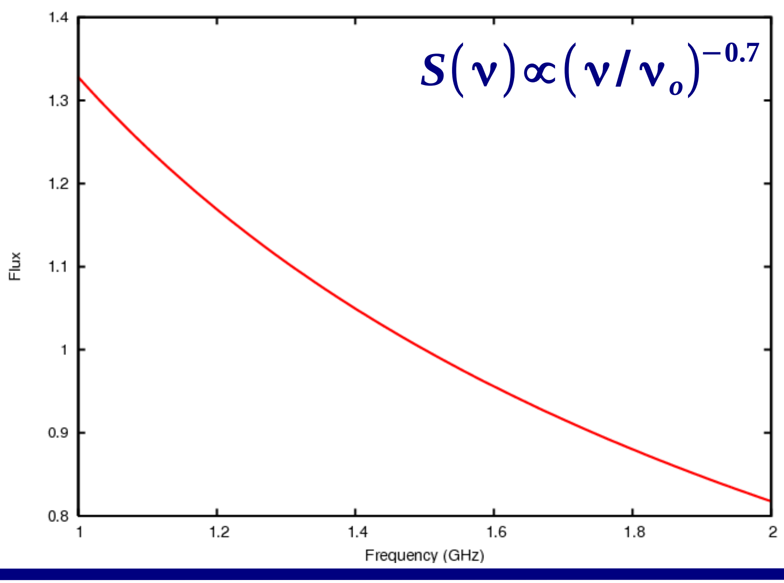
MS-Clean

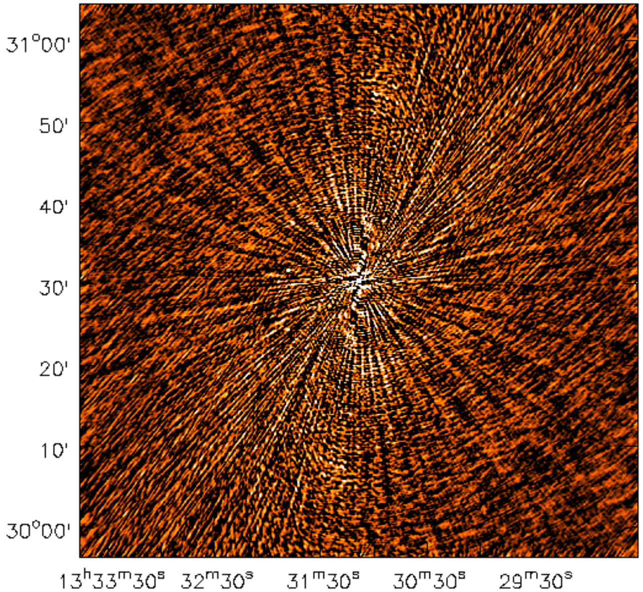
MT-MFS/MS-MFS

Hogbom Clean

What do we call wide-band?

- When fractional signal bandwidth used for imaging > ~20%
  - Plus source spectral index >= -1.0
  - Plus target dynamic range > 1000
- Spectral effects for higher source spectral index will become significant at lower bandwidth ratios
  - Empirical Dynamic range : $\frac{I\alpha}{100}$
  - Spectral line imaging, by definition, does not require wide-band imaging algorithms





Wide-band Imaging Sensitivity

Frequency Range :	(1 – 2 GHz)	(4 – 8 GHz)	(8 – 12 GHz)
Bandwidth : $\nu_{max} - \nu_{min}$	1 GHz	4 GHz	4 GHz
Bandwidth Ratio : $\nu_{max} : \nu_{min}$	2 : 1	2 : 1	1.5 : 1
Fractional Bandwidth : $(\nu_{max} - \nu_{min}) / \nu_{mid}$	66%	66%	40%

Broad-band receivers increase the 'instantaneous' imaging sensitivity of an instrument

Continuum sensitivity : (at field-center) $\sigma_{cont} \propto \frac{T_{sys}}{\sqrt{N_{ant}(N_{ant}-1)\ \delta\tau\,\delta\nu}}$

50 MHz → 2 GHz   Theoretical sensitivity improvement : $\sqrt{\frac{2\,GHz}{50\,MHz}} \approx 6$ times.

In practice, effective broadband sensitivity for imaging depends on bandpass shape, data weights, and regions of the spectrum flagged due to RFI (radio-frequency interference).

Use narrow-band channels – avoid bandwidth smearing

In the early days of continuum-observing, only one visibility was computed across the entire bandwidth of the receiver, and attributed to the reference (or middle) frequency ν_0. Delay-tracking was also done only at ν_0.

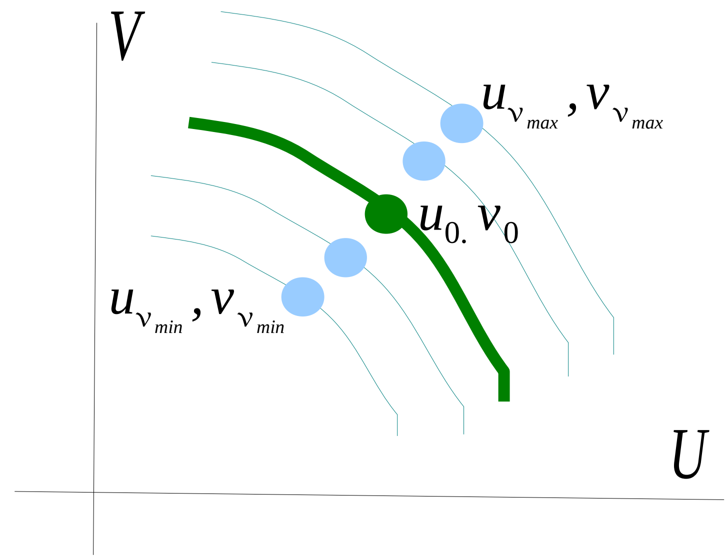
The visibility $V(u_\nu)$ is mistakenly mapped to $u_0 = \frac{b\nu_0}{c} = \frac{\nu_0}{\nu} u_\nu$

Similarity theorem of Fourier-transforms :

=> A radial shift in the source position, with frequency.

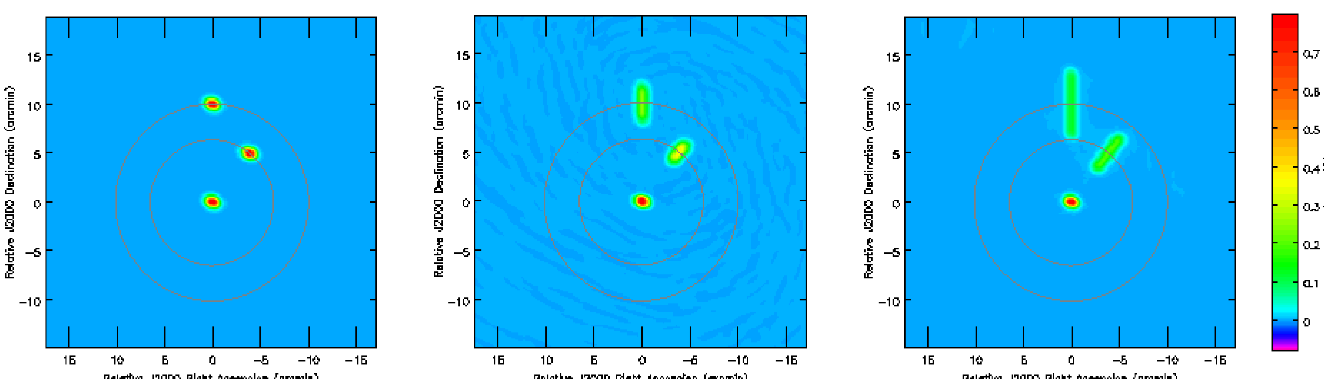
=> Radial smearing of the brightness-distribution

Note : Excessive channel-averaging has a similar effect.



An (exaggerated) example of bandwidth-smearing with a 1-2 GHz signal......

2 MHz 200 MHz 1.0 GHz



Bandwidth Smearing Limits at 1.4 GHz

33 MHz (VLA D-config),
10 MHz (VLA C-config),
3 MHz (VLA B-config),
1 MHz (VLA A-config)

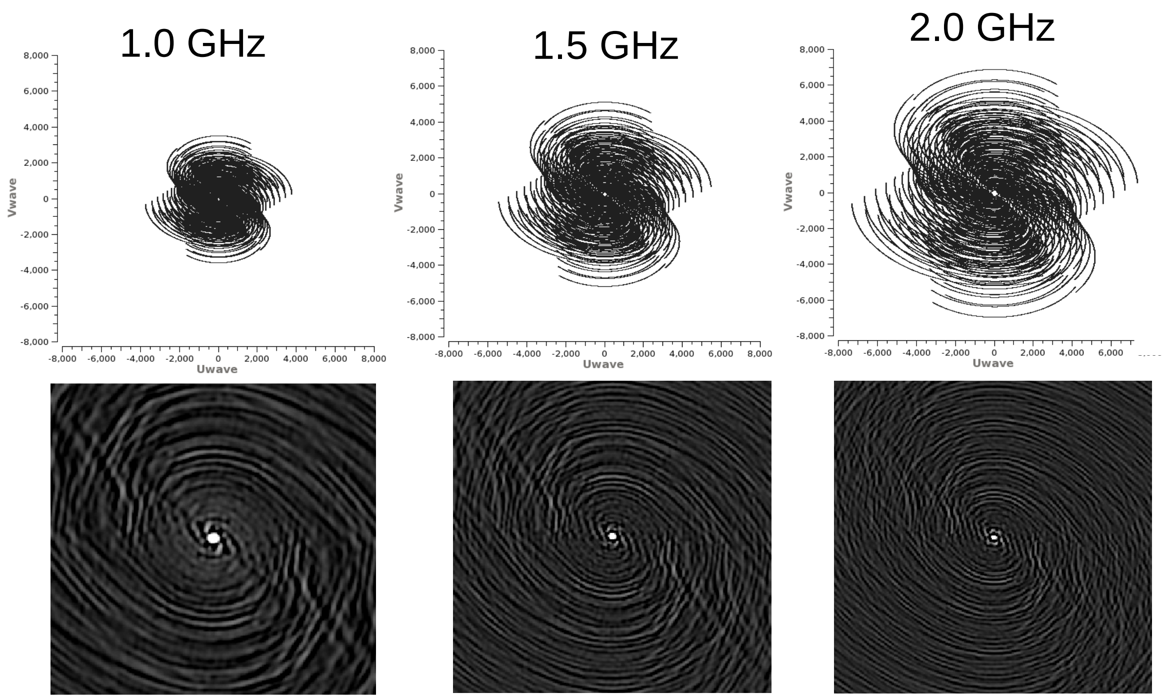
Contours represent 5 and 10 arcmin distances from the phase-center.

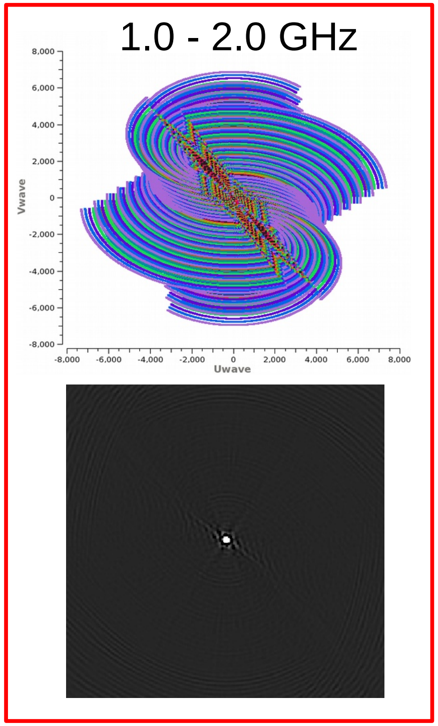
Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency :

- Angular-resolution increases at higher frequencies
- Sensitivity to large scales decreases at higher frequencies
- Wideband UV-coverage has fewer gaps => lower Psf sidelobe levels

$$S(u,v)_{\nu} = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\,\nu}{c}$$





But, when the source intensity varies with frequency, different channels measure the visibility function of different sky-brightness distributions

the $V(u_{\nu}, v_{\nu}) = \iint I(l,m,\nu)\, e^{2\pi i (u_{\nu} l + v_{\nu} m)}\, dl\, dm$

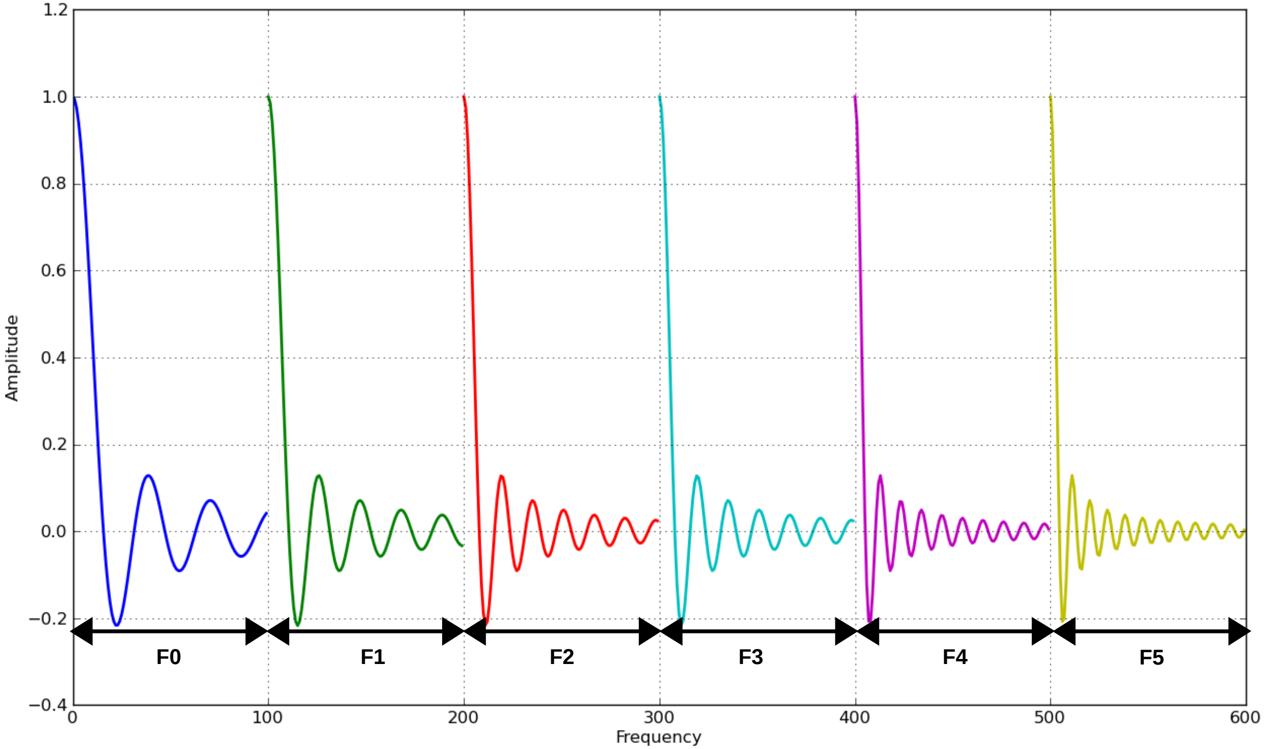
=> Need to model the spectrum as part of image reconstruction

Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency:

- PSF structure scales with frequency

$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\,\nu}{c}$$



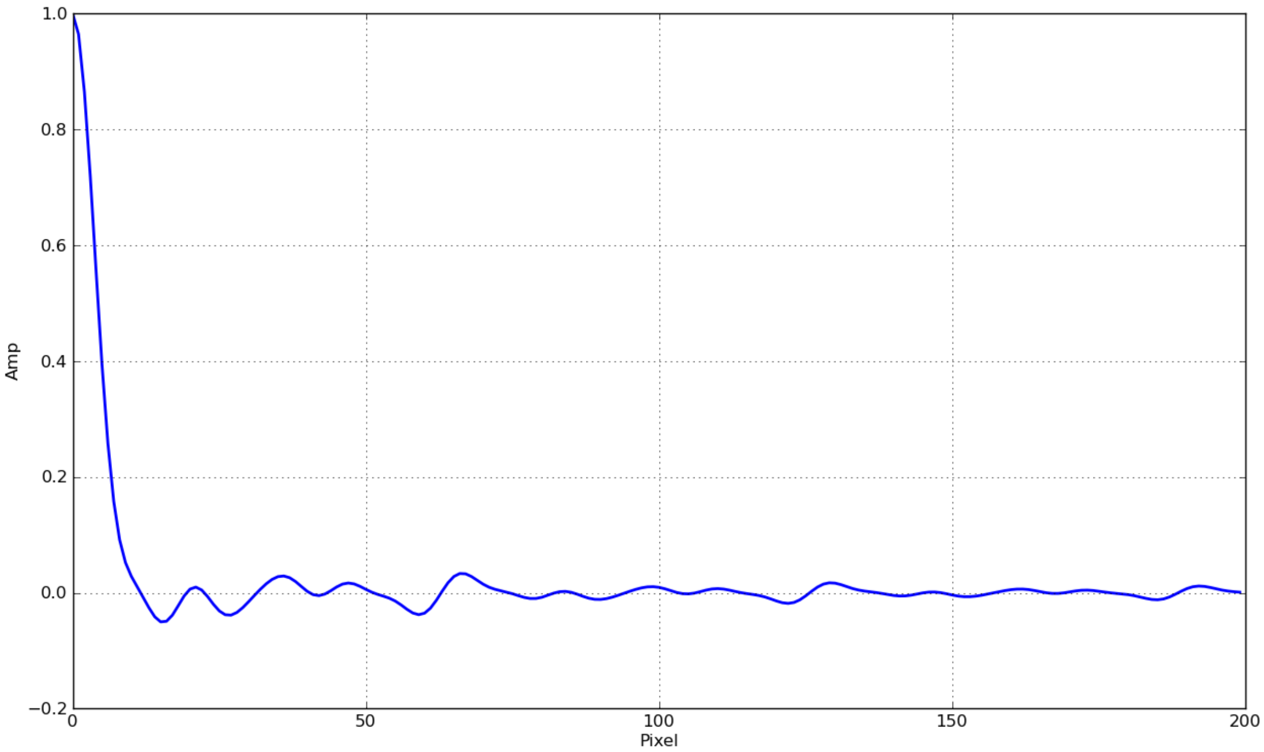
Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency:

- PSF structure scales with frequency

$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\,\nu}{c}$$

$$PSF_{Continuum} = \sum_\nu PSF(\nu)$$

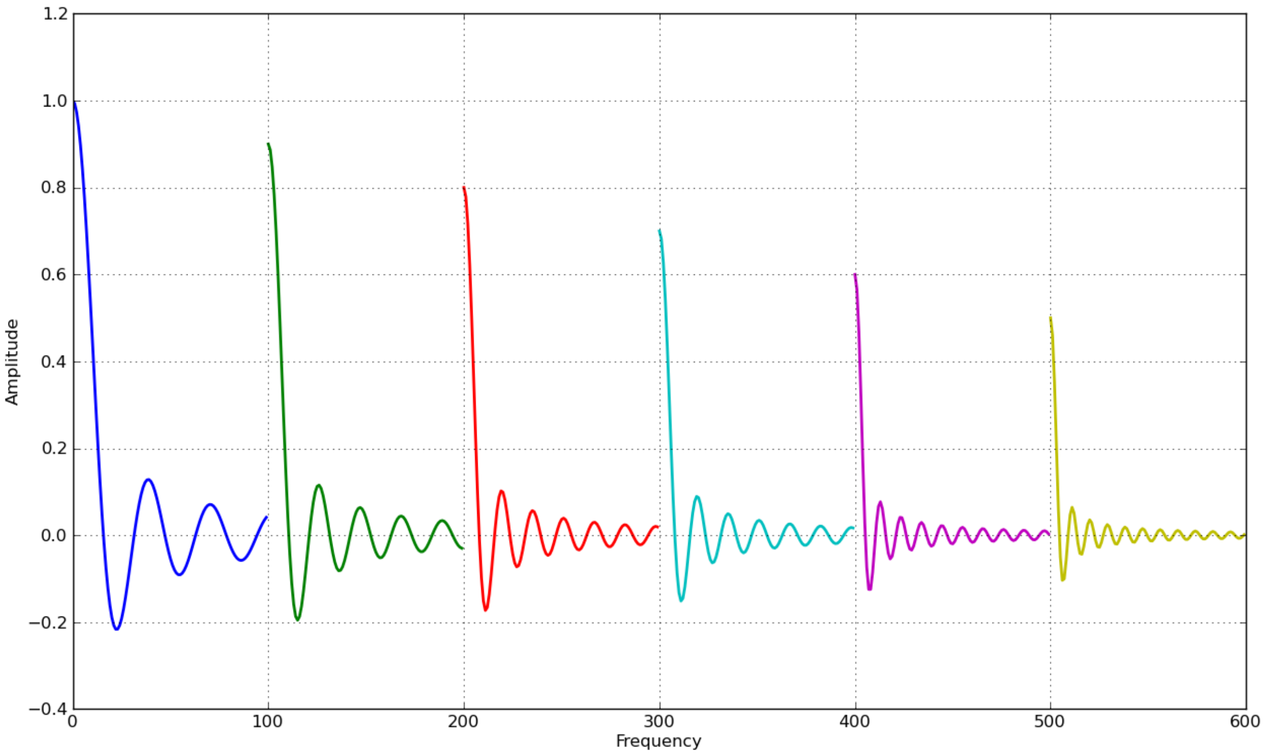


Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency:

$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\nu}{c}$$

- PSF structure scales with frequency
- Due to source Spectral Index, PSF amplitude also changes with frequency



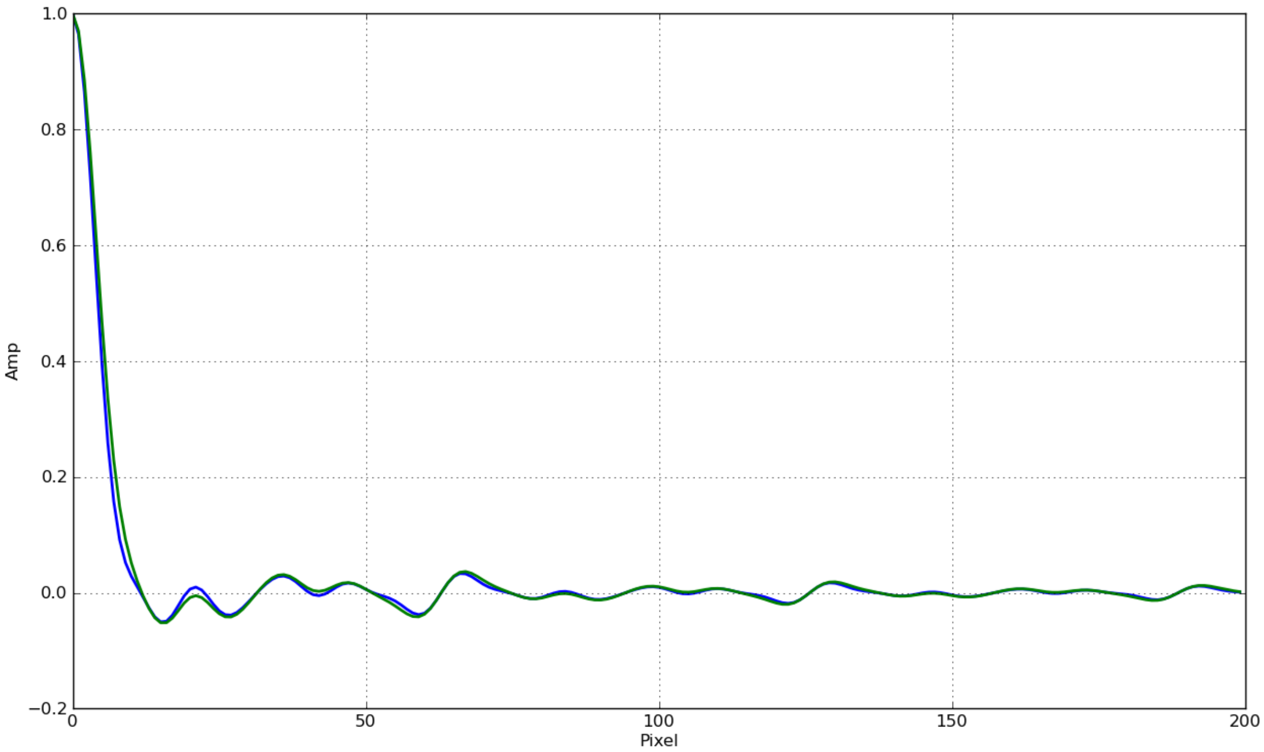
Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency:

$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\,\nu}{c}$$

- PSF structure scales with frequency
- Due to source Spectral Index, PSF amplitude also changes with frequency

$$PSF(x_o)_{Continuum} = \sum_\nu I(x_o,\nu)\, PSF(x - x_o, \nu)$$



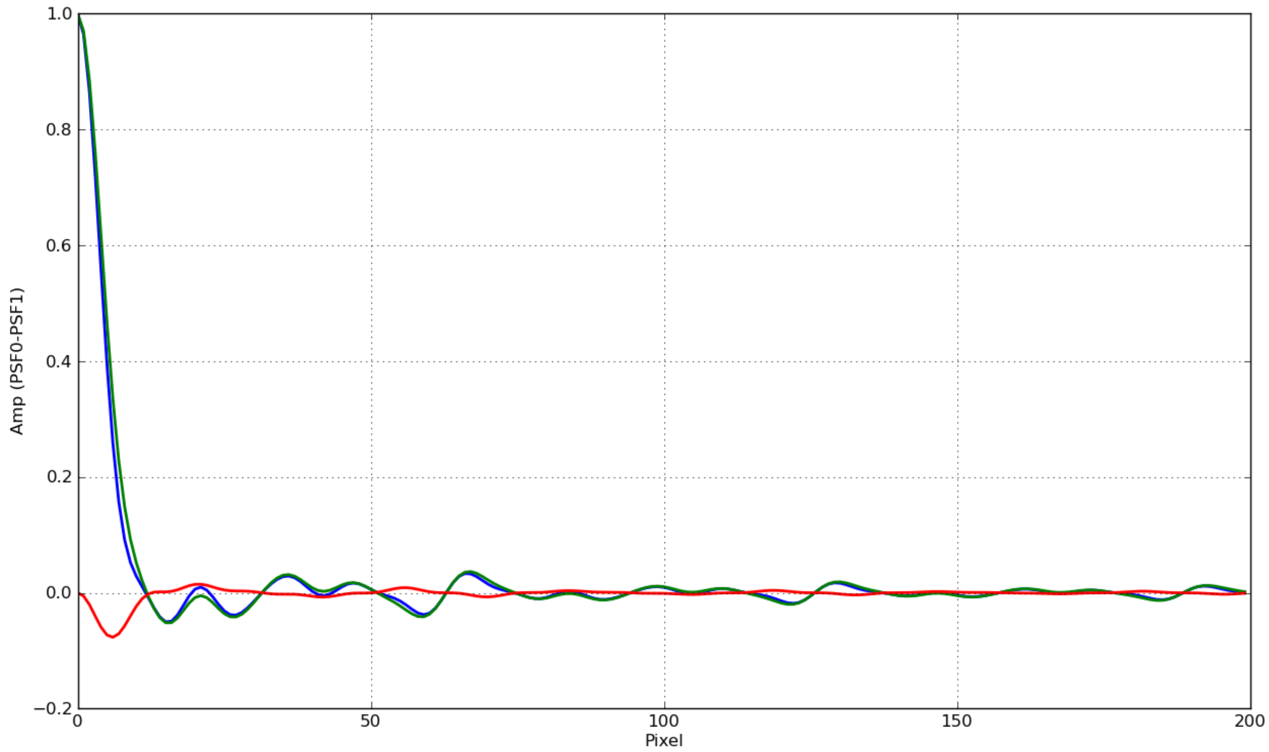
Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency:

$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\nu}{c}$$

- PSF structure scales with frequency
- Due to source Spectral Index, PSF amplitude also changes with frequency

$$Res(x_o)_{Continuum} = \sum_\nu PSF(x - x_o, \nu) - \sum_\nu I(x_o, \nu) PSF(x - x_o \nu)$$



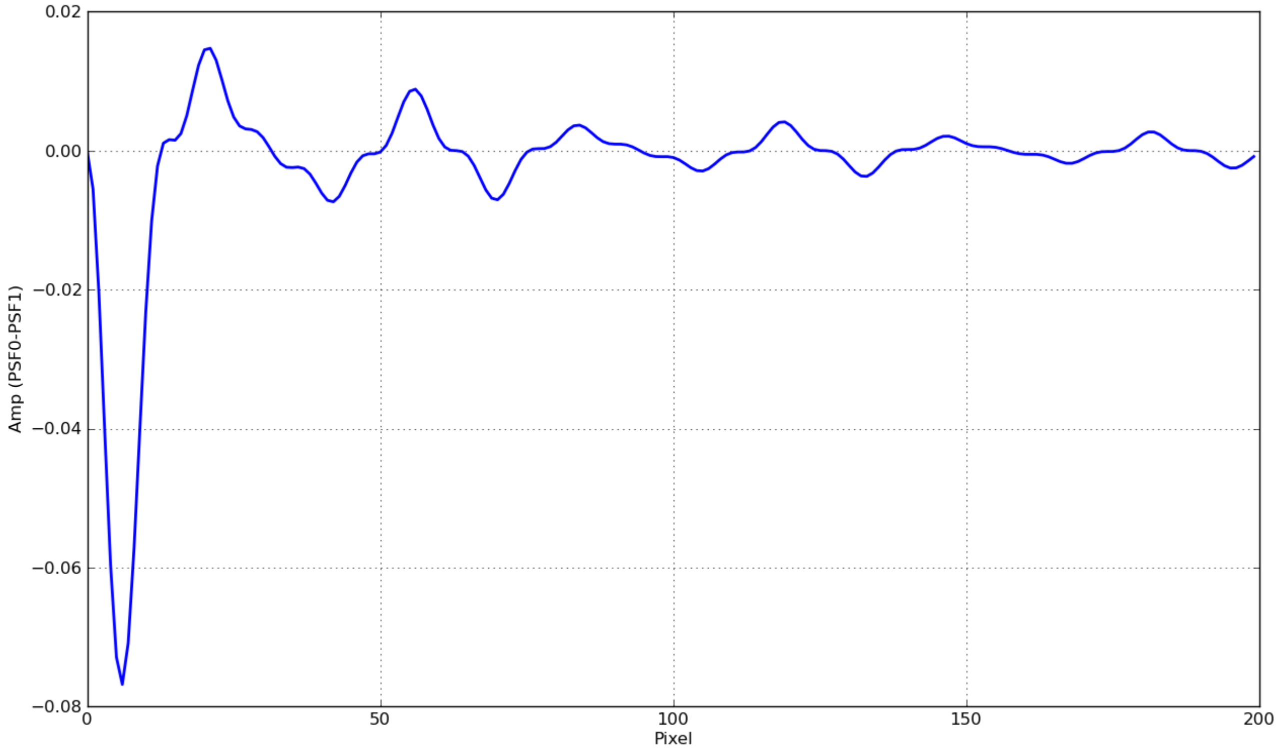
Frequency-dependent UV-coverages and PSFs

Spatial-frequency coverage and imaging properties change with frequency:

$$S(u,v)_\nu = \frac{\vec{b}}{\lambda} = \frac{\vec{b}\,\nu}{c}$$

- PSF structure scales with frequency
- Due to source Spectral Index, PSF amplitude also changes with frequency

$$Res(x_o)_{Continuum} = \sum_\nu PSF(x - x_o, \nu) - \sum_\nu I(x_o, \nu) PSF(\nu)$$



Wideband Imaging Options

(1) Make images for each channel / SPW separately.

- Signal-to-noise ratio : one SPW
- Angular resolution varies with SPW (smooth to lowest)
- Imaging fidelity may change across SPWs
- Primary beam correction can be done per SPW

Cube imaging will suffice for sources with simple spatial structures, and where the added uv-coverage, sensitivity and angular resolution is not required for the target science.

(2) Combine all frequencies during imaging (MFS : multi-frequency synthesis)

- Signal-to-noise ratio : all SPWs
- Angular resolution is given by the highest frequency
- Imaging fidelity is given by the combined uv-coverage
- Wideband PB correction is required (average gain and spectrum)

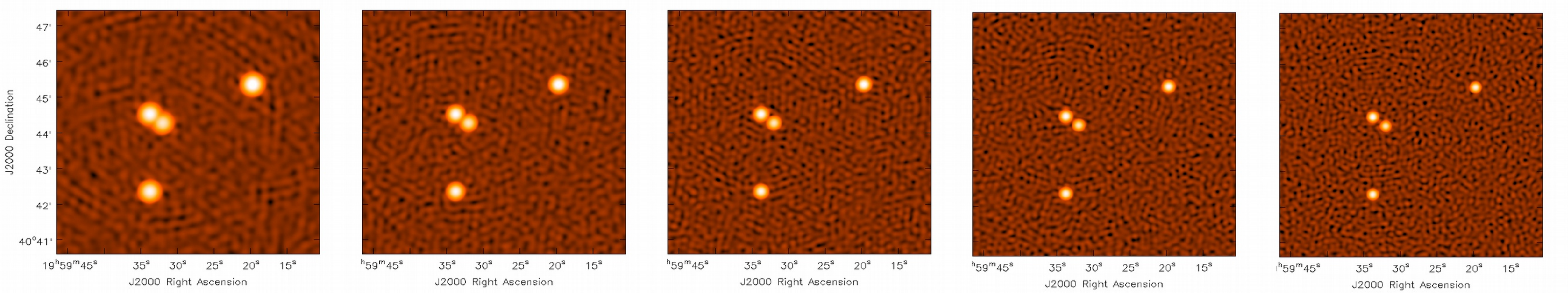
Multi-frequency-synthesis is needed to fully utilize the wideband uv-coverage and sensitivity during image reconstruction.

The frequency dependence of the sky and instrument must be taken into account

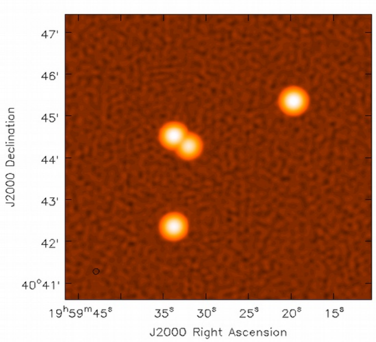
Single-channel vs MFS imaging – Angular Resolution

Simulated Example : 3 flat-spectrum sources + 1 steep-spectrum source (1-2 GHz)

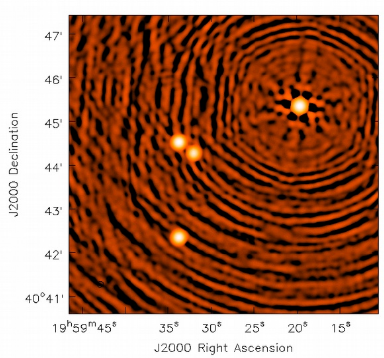
Images made separately at different frequencies between 1 and 2 GHz



Combine all single-frequency images (after smoothing)

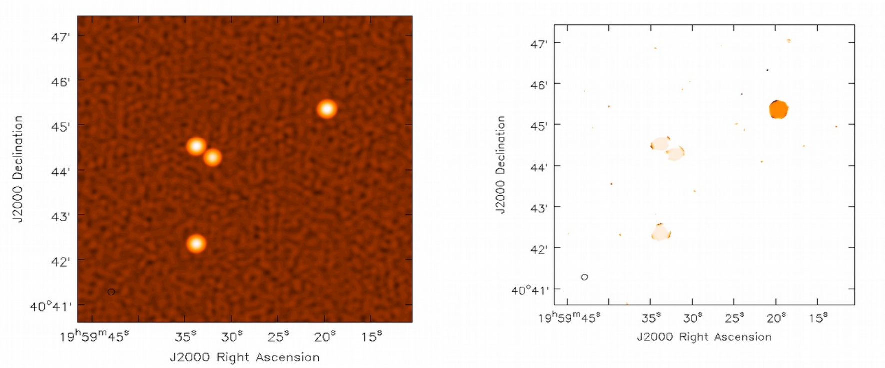


Use all UV-coverage together, but ignore spectra



Use all UV-coverage together
+ Model and fit for spectra too

Output : Intensity and Spectral-Index



=> Imaging with a spectrum model : higher angular resolution + continuum sensitivity.

Continuum Imaging : (multi-scale) multi-frequency-synthesis

Sky Model : Collection of multi-scale flux components whose amplitudes follow a polynomial in frequency

$$I_\nu^{sky} = \sum_t I_t \left(\frac{\nu - \nu_0}{\nu_0} \right)^t \qquad \text{where } I_t = \sum_s [I_s^{shp} * I_{s,t}]$$

Algorithm : Linear least squares + deconvolution

Parameters : mode='mfs', nterms=2, reffreq='1.5GHz', multiscale=[0,6,10]

Data Products : Taylor-Coefficient images $I_{0,}^m I_{1,}^m I_{2,}^m \ldots$ that represent the observed spectrum

Interpretation :

- As a power-law (spectral index and curvature) $I_\nu = I_{\nu_0} \left(\frac{\nu}{\nu_0} \right)^{\alpha + \beta \log(\nu/\nu_0)}$

$$I_0^m = I_{\nu_0} \qquad I_1^m = I_{\nu_0} \alpha \qquad I_2^m = I_{\nu_0} \left(\frac{\alpha(\alpha-1)}{2} + \beta \right)$$

- PB-correction : Model the average PB-spectrum with a Tayor-polynomial, and do a post-deconvolution Polynomial-Division

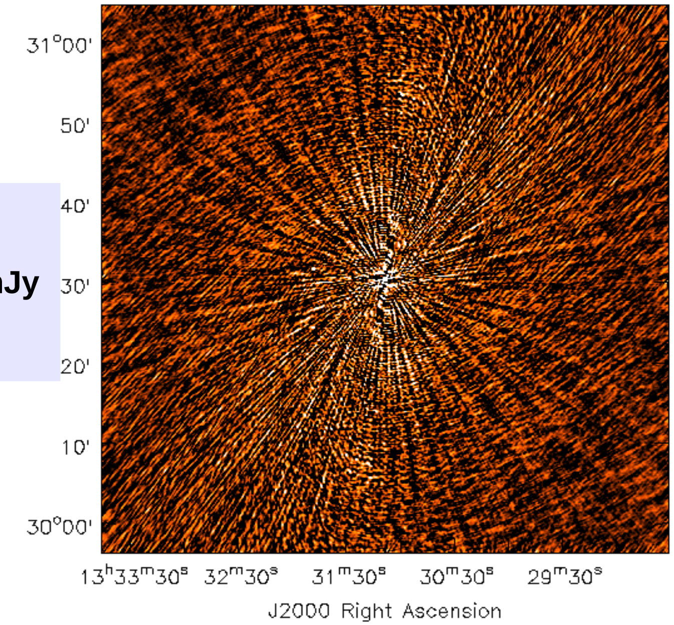
$$\frac{(I_{0,}^m I_{1,}^m I_{2,}^m \ldots)}{(P_{0,} P_{1,} P_{2,} \ldots)} = (I_{0,}^{sky} I_{1,}^{sky} I_2^{sky} \ldots)$$

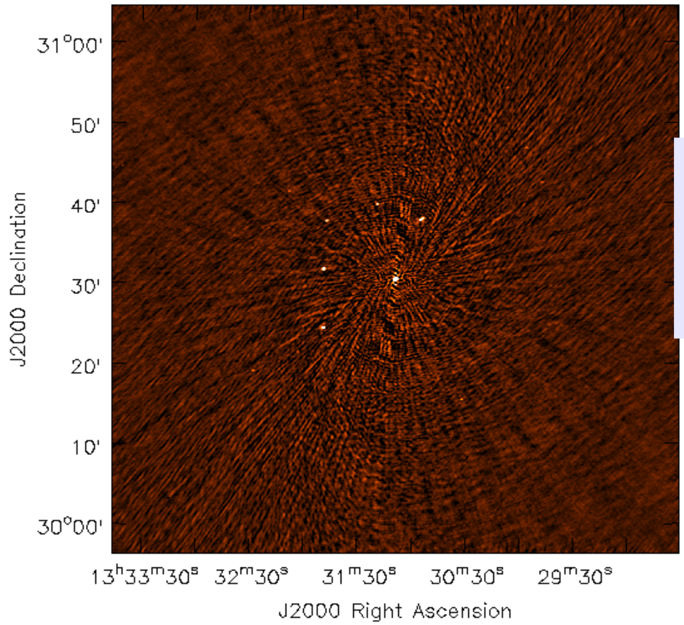
Dynamic-range with MS-MFS : 3C286 example : Nt=1,2,3,4

NTERMS = 1

Rms : 9 mJy -- 1 mJy

DR : 1600 -- 13000





NTERMS = 2

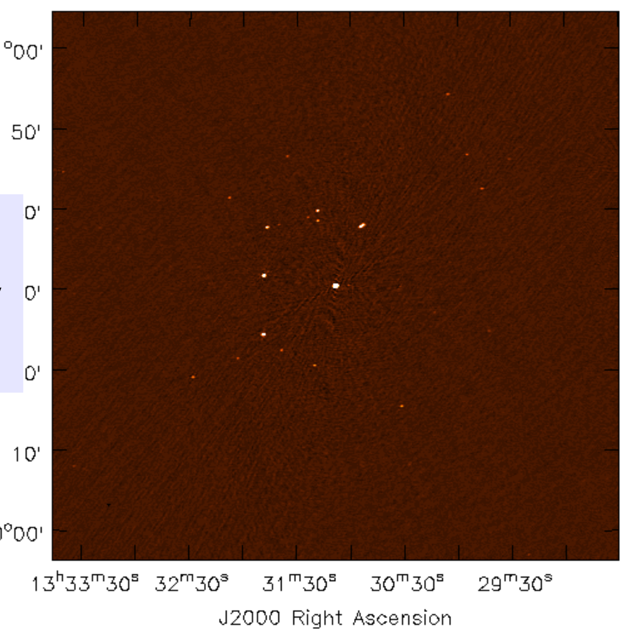
Rms :1 mJy -- 0.2 mJy

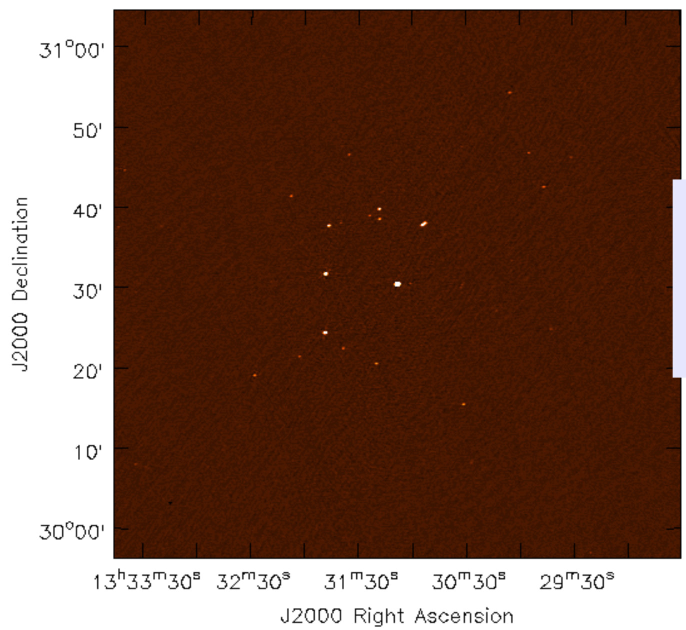
DR :10,000 -- 17,000

NTERMS = 3

Rms : 0.2 mJy -- 85 uJy

DR : 65,000 -- 170,000





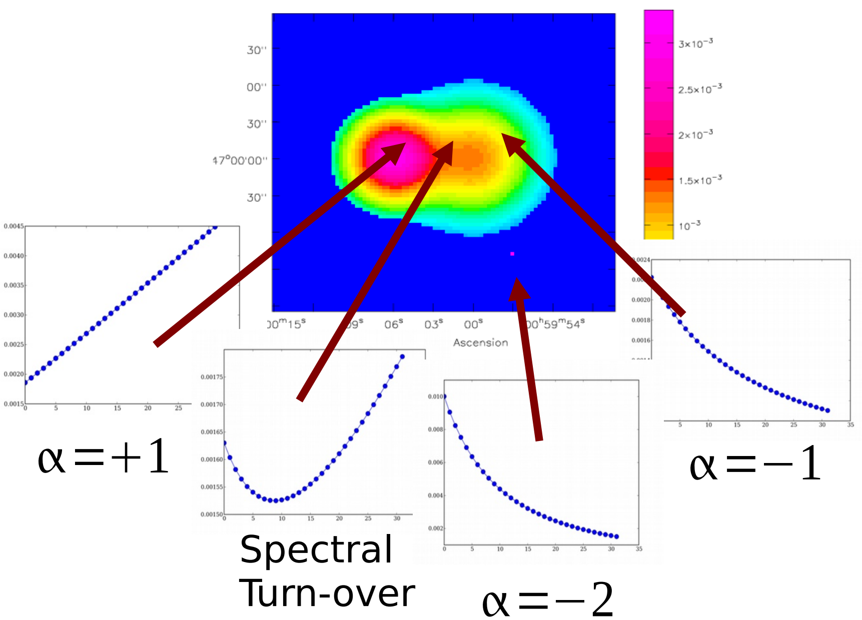
NTERMS = 4

Rms 0.14 mJy -- 80 uJy

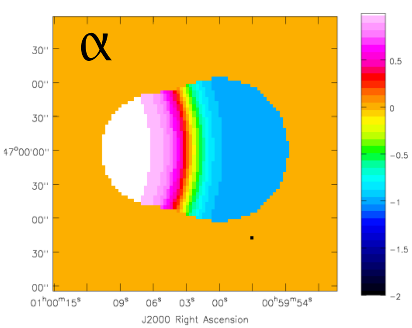
DR :>110,000 -- 180,000

Example of wideband-imaging on extended-emission

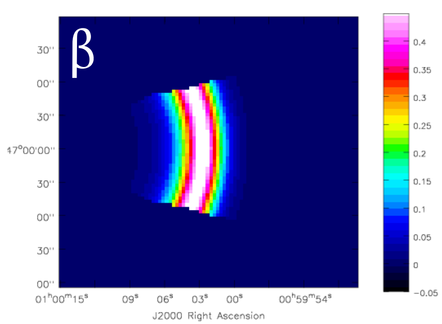
Intensity Image

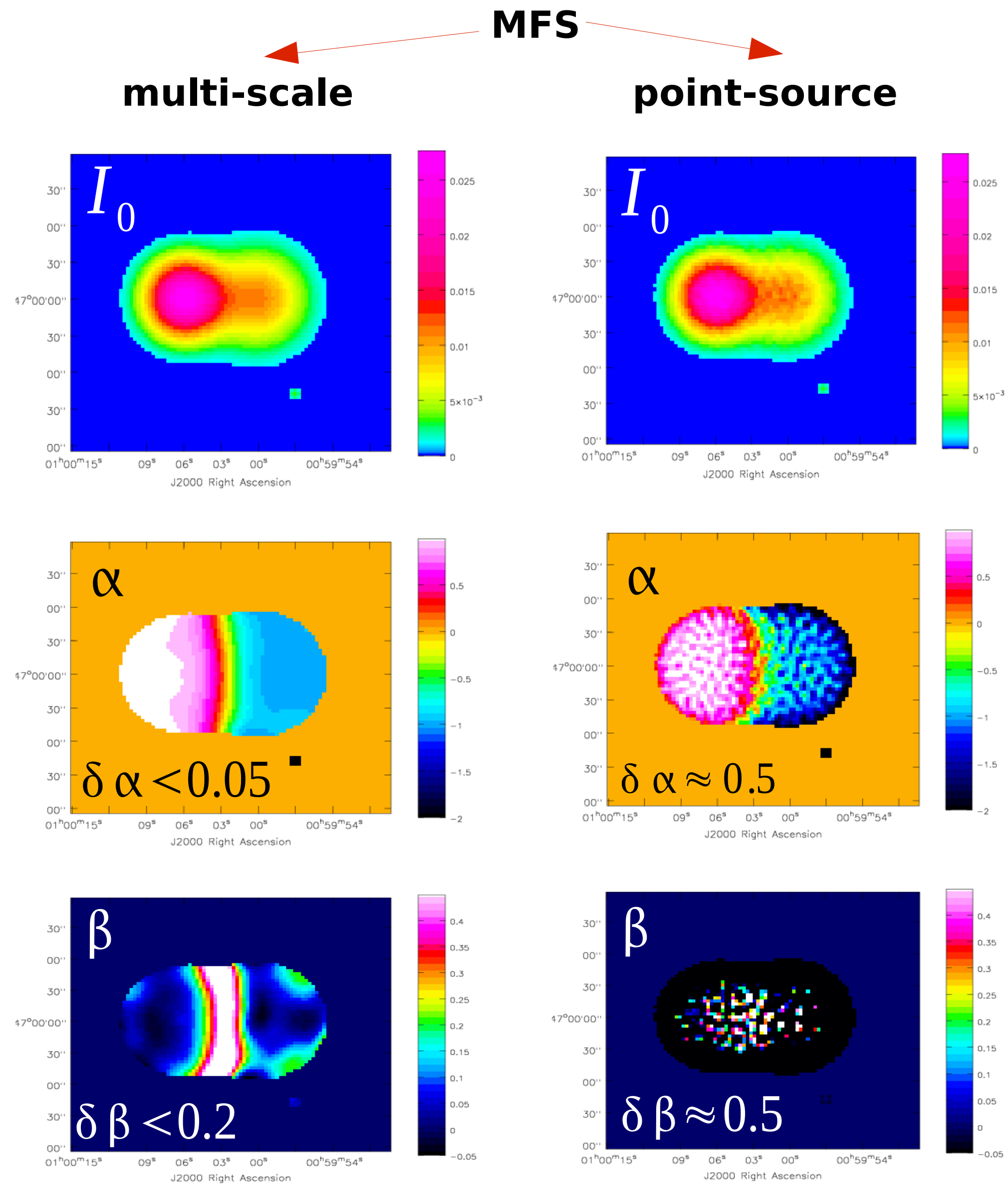


Average Spectral Index



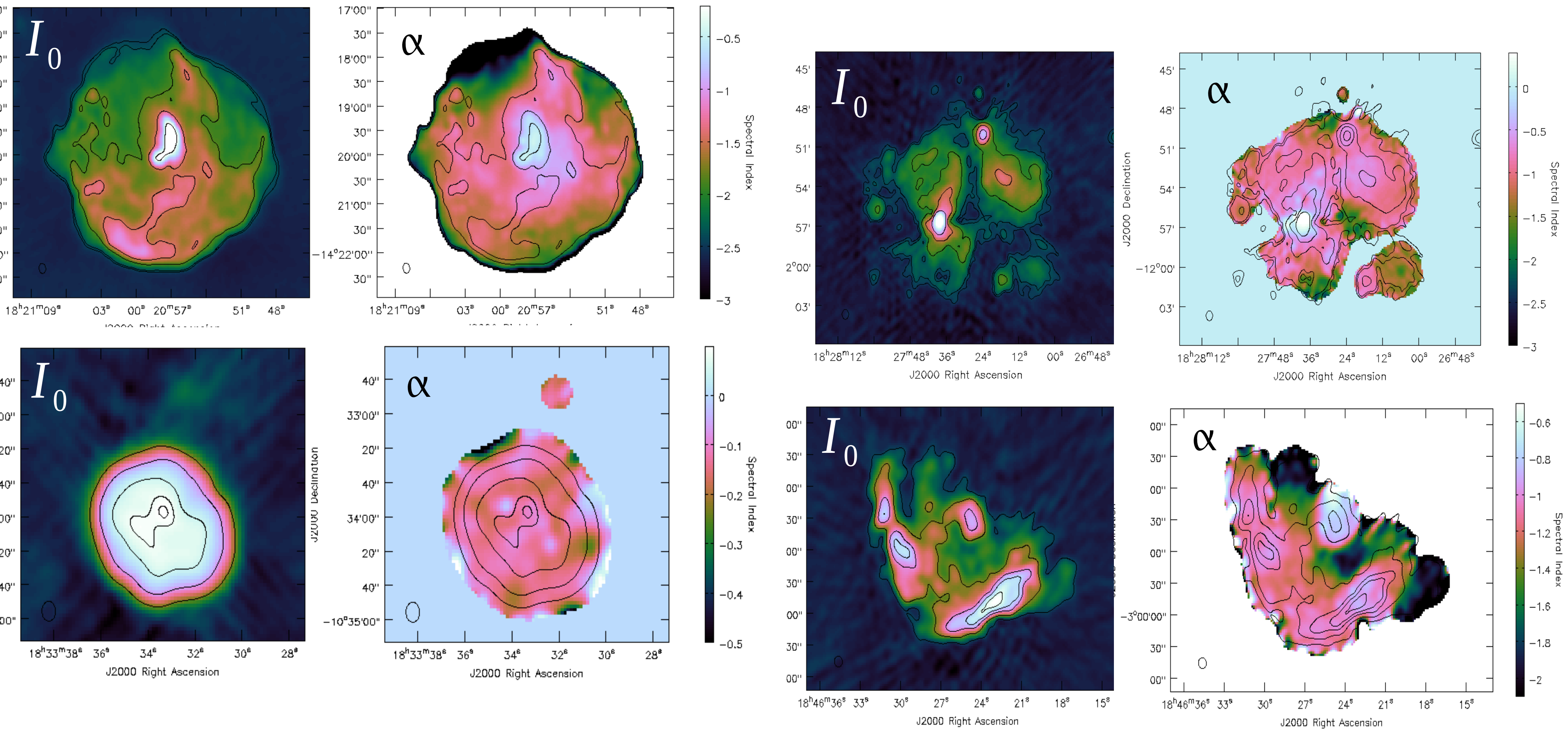
Gradient in Spectral Index





=> For extended emission - spectral-index error is dominated by 'division between noisy images'
– a multi-scale model gives better spectral index and curvature maps

Extended emission – SNR example (a realistic expectation)



These examples used nterms=2, and about 5 scales.

=> Within 1-2 Ghz and 4-8 GHz, can tell-apart regions by their spectral-index (+/- 0.2) if SNR>100.
(this accuracy will increase with wider bandwidths – 1-3 GHz CABB)

=> These images have a dynamic-range limit of few x 1000   ---> residuals are artifact-dominated

Errors in polynomial fitting + Imaging (empirical)

For a 1 Jy point source with spectral index of -1.0 ...

– If spectra are ignored during MFS imaging => Errors increase with bandwidth.

Dynamic-range limits for VLA uv-coverage (natural)

1-2 GHz => ~ 1000
1-3 GHz => few 100

– If spectra are modeled + High signal-to-noise => Need higher-order polynomials to fit a power-law

1 term (flat spectrum) => peak intensity error of 0.1 (on 1 Jy)
2 terms (linear spectrum) => peak intensity error of 0.02, spectral index error of 0.1
3 terms (quadratic spectrum) => intensity error of 0.0001, spectral index error of 0.05

– If spectra are modeled + Low signal-to-noise => Higher-order polynomials give more errors

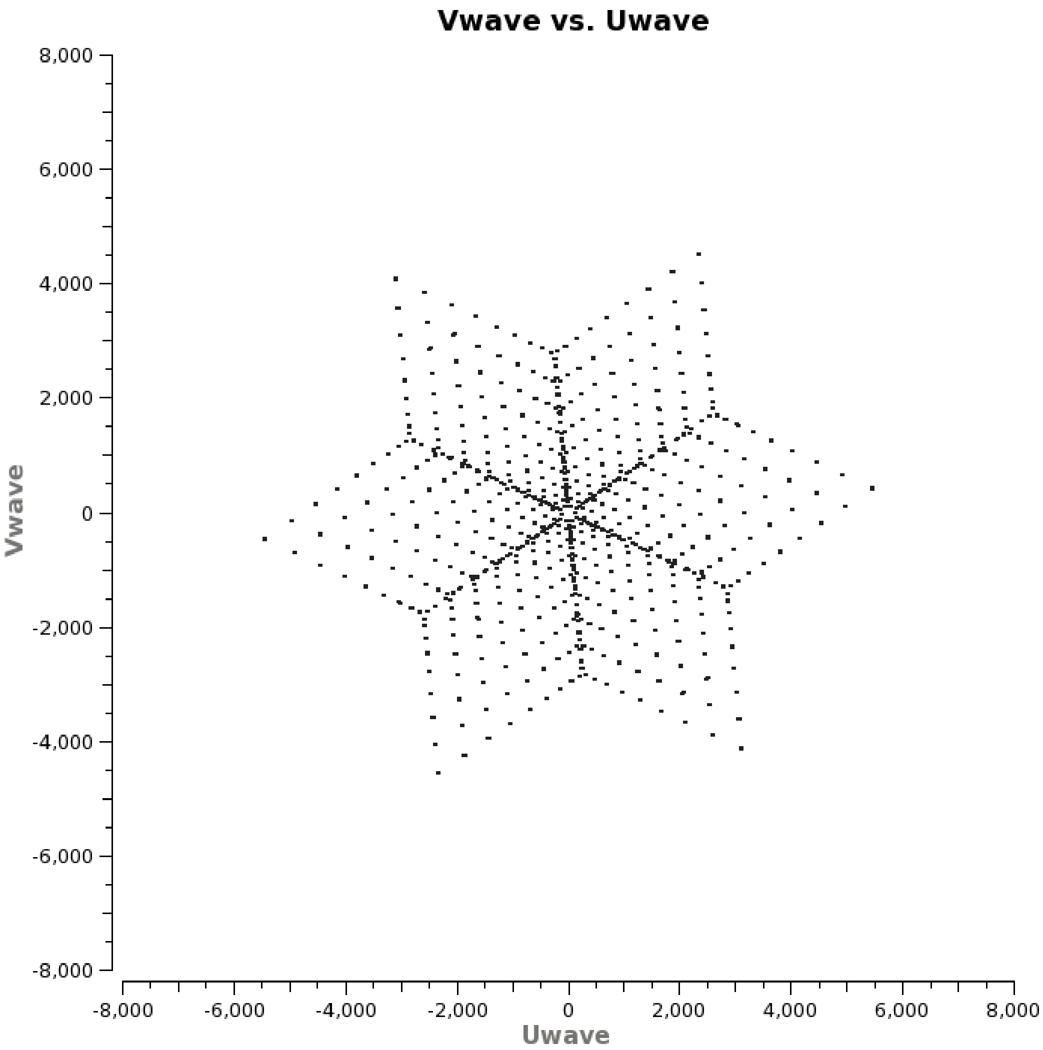
The following situations give similar error on spectral index (~ 0.1) for a point source....

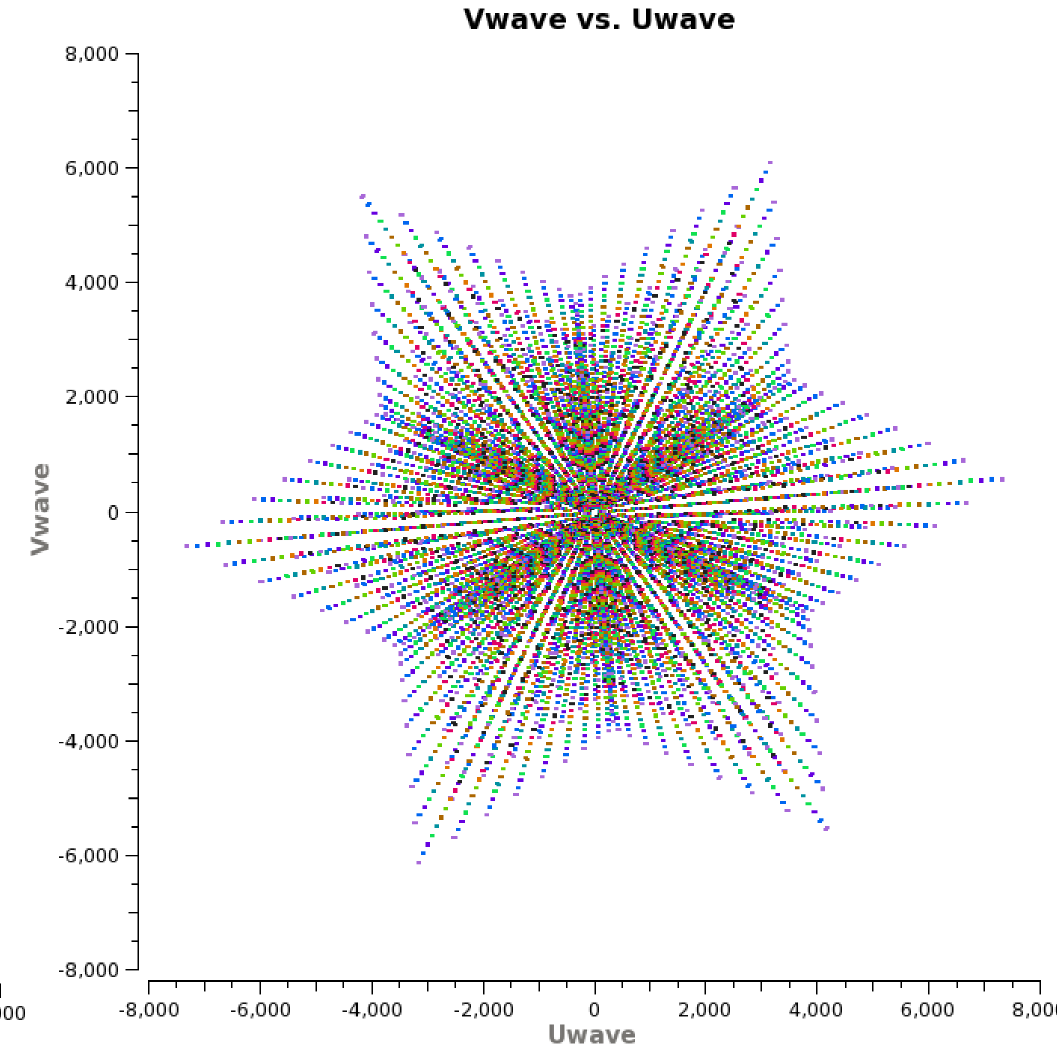
L-Band + C-Band : 1-8 GHz : Sources with signal-to-noise ratio of 10~20
L-Band only (1-2 GHz) or C-Band only (4-8 GHz) : Sources with SNR ~ 40

For extended emission, spectral index errors <= 0.2 only for SNR > 100......

Multi-Frequency-Synthesis : Snapshot

Observing tip.....

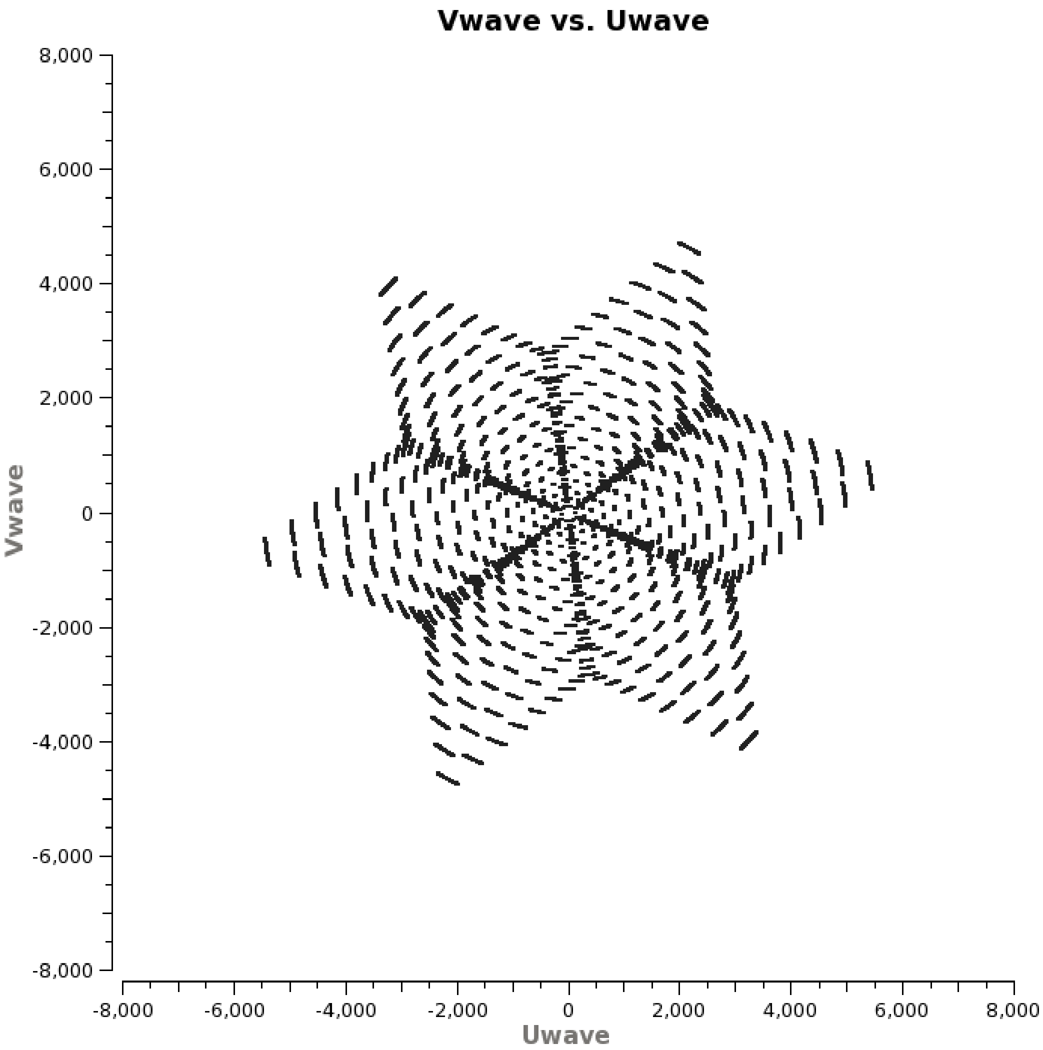


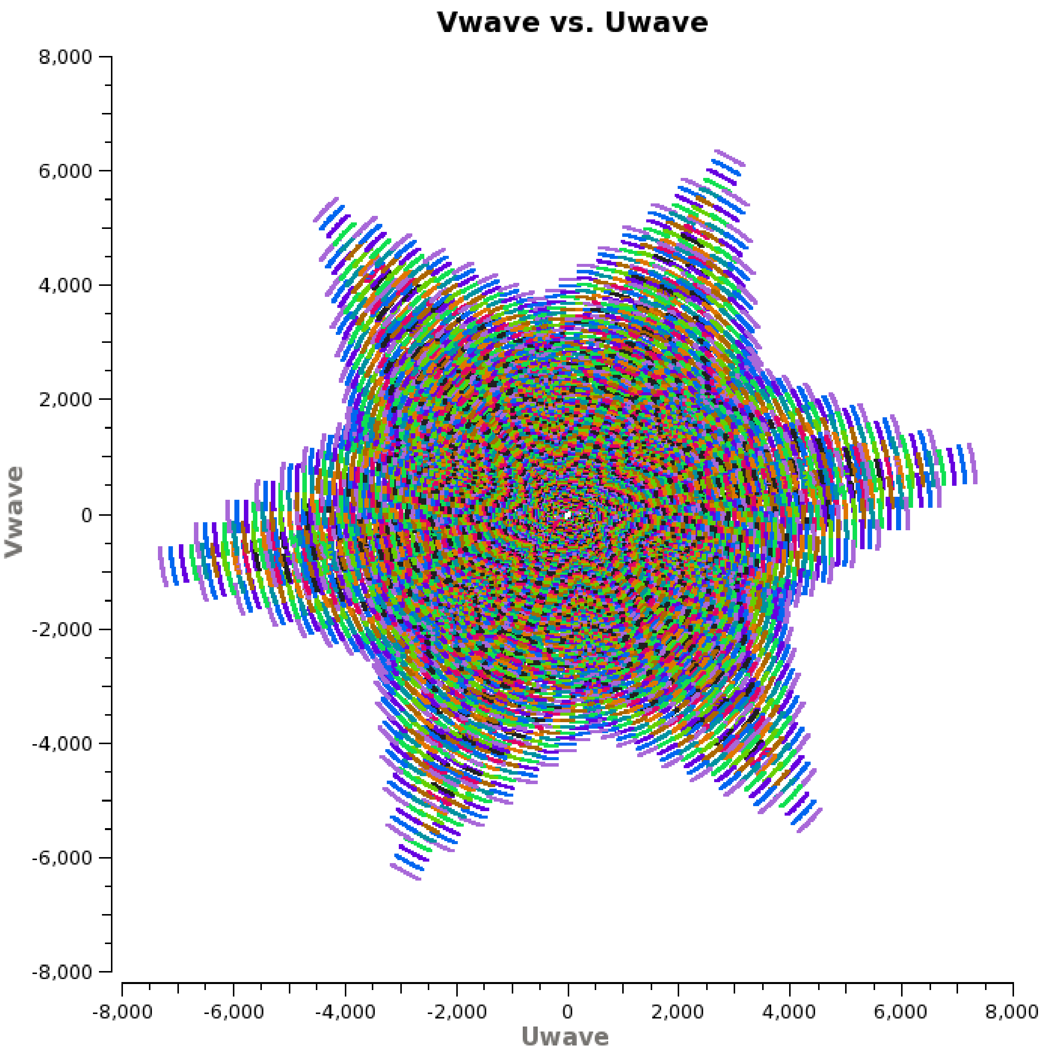


Wideband UV-coverage fills the UV-plane radially.....

Multi-Frequency-Synthesis : 30 min

Observing tip.....



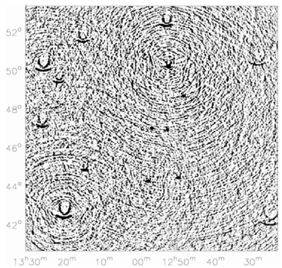


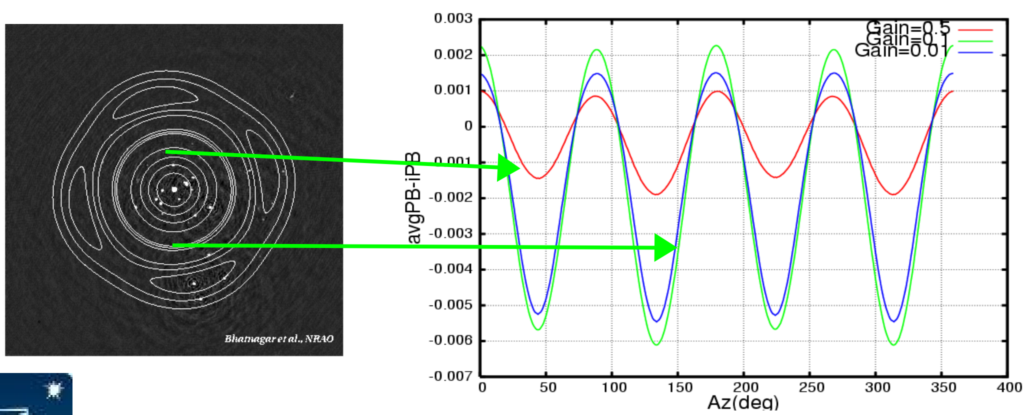
Small time-increments generate good uv-filling => Plan wideband observations in small time-chunks, spread out in time to cover more spatial-frequencies at-least once.

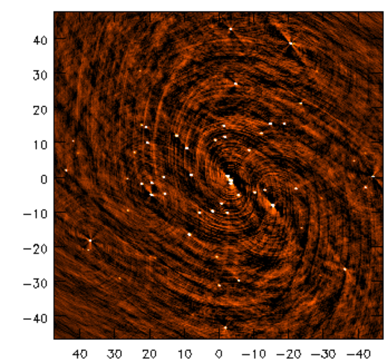
What do we call wide-field?

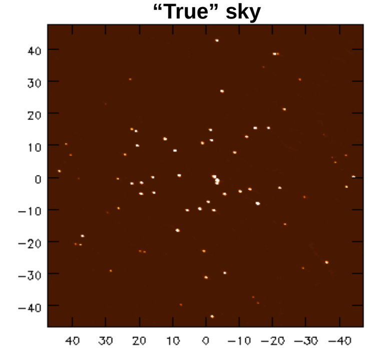
- Imaging that requires invoking any of the following:
  - Corrections for non co-planar baseline effects
  - Corrections for the rotational asymmetry of the PB
    - Imaging beyond 50% point, mosaicking
  - Corrections for the frequency or polarization dependent effects
  - Noise limited imaging at 4-,P-,L-, S- (and probably C-Band)
    - Because of the radio brightness distribution
- Noise limited imaging of structure comparable to the PB beam-width
- Mosaicking: imaging on scales larger than the PB beam-width

$$\frac{\lambda}{B_{max}} \leq \theta_f^2$$



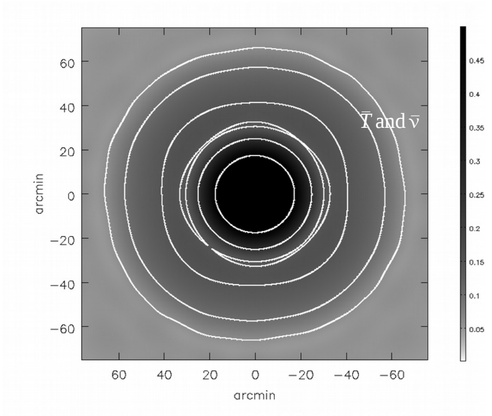


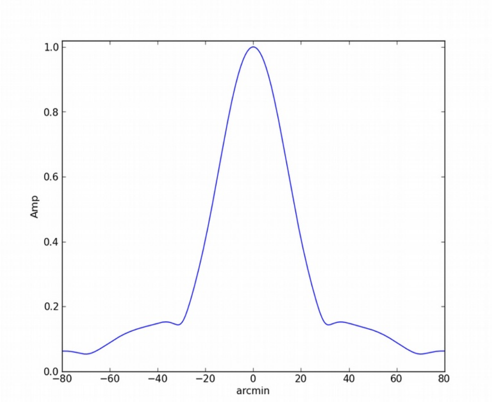




Why wide-field?

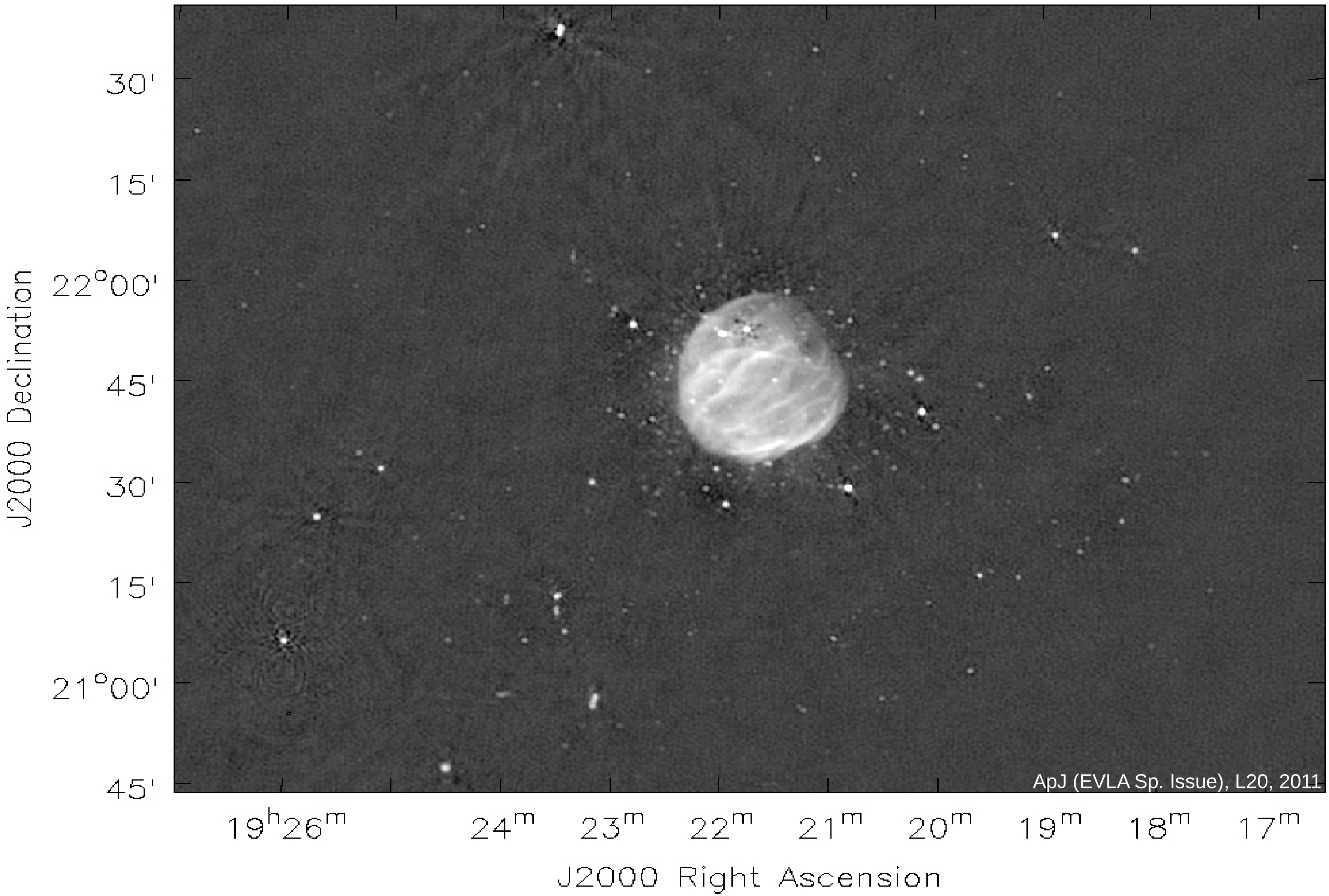
- Primarily due to improved continuum sensitivity
- @L-Band, PB gain ~1 deg. away can be up to 10%
  - In the EVLA sensitivity pattern, VLA sensitivity is achieved at the location of the VLA-null!
  - No null in the EVLA sensitivity pattern





- E.g. a 1% PSF side lobe due to a source away from the center is now significantly above continuum thermal noise limit
  - This is a largely independent of the total integration time

Wide-field Issues



Wide-field sensitivity because of wide-bandwidths

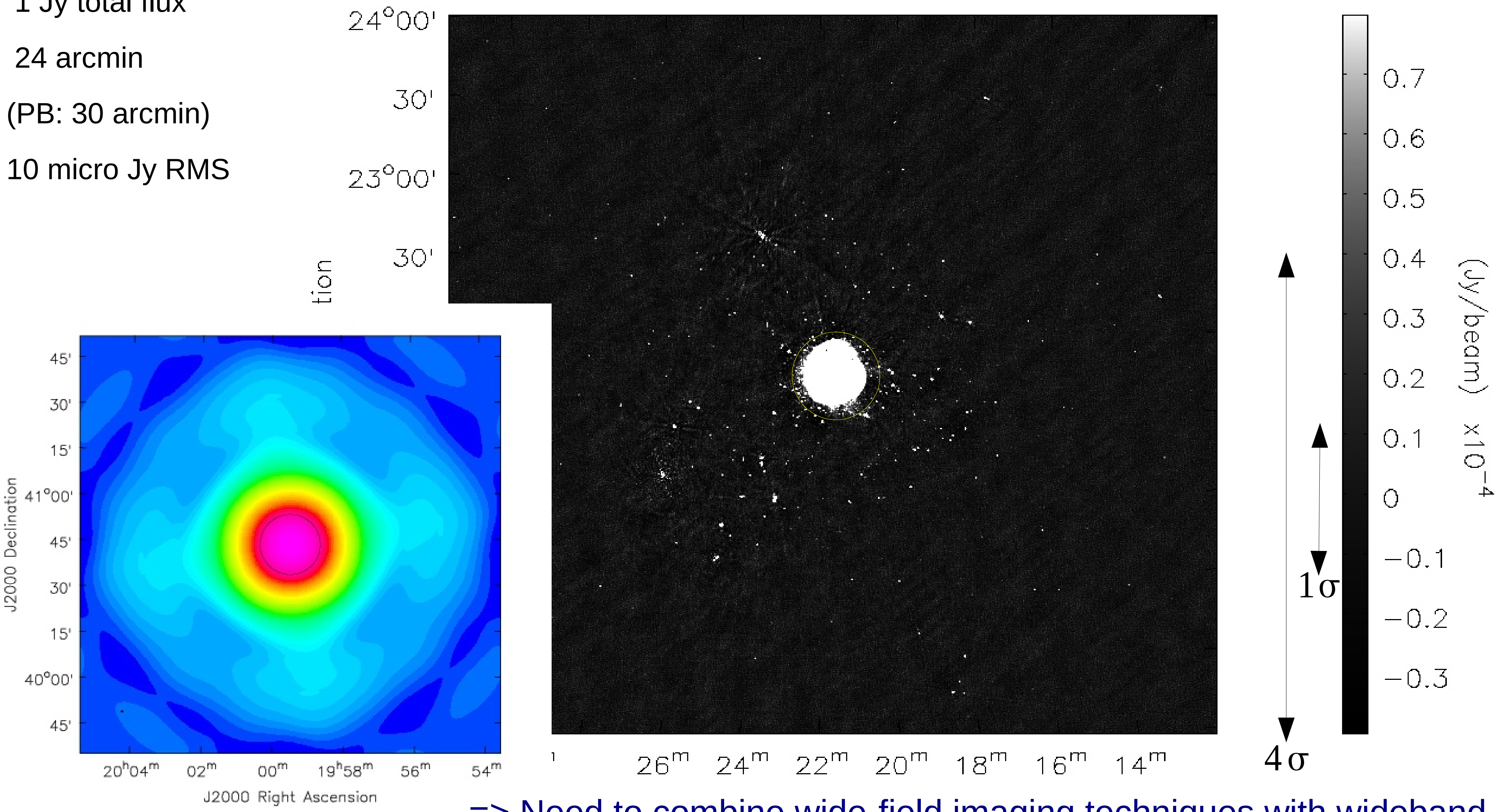
G55.7+3.4 : Galactic supernova remnant : 4 x 4 degree field-of-view from one EVLA pointing

1 Jy total flux

24 arcmin

(PB: 30 arcmin)

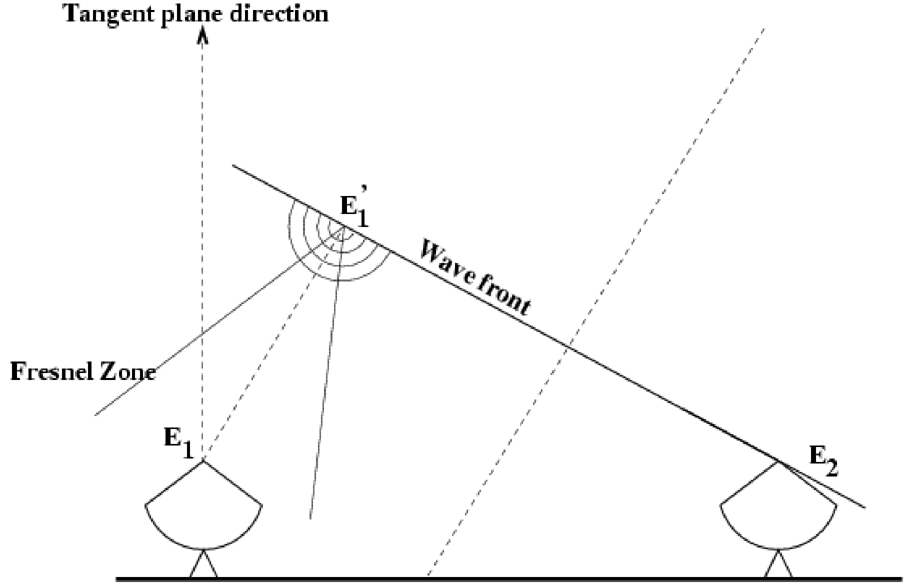
10 micro Jy RMS

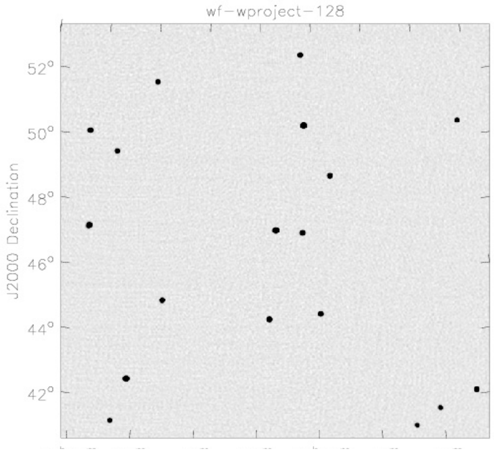


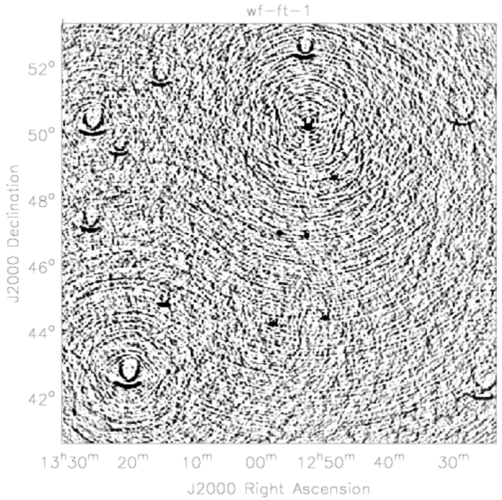
=> Need to combine wide-field imaging techniques with wideband..

Non co-planar baseline: The W-term

- 2D FT approximation of the Measurement Equation breaks down



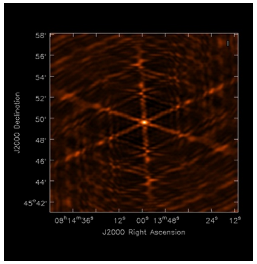




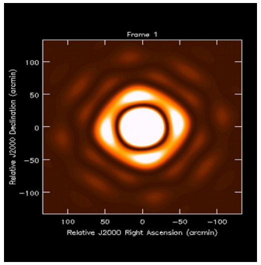
- We measure: $V_{12} = \langle \boldsymbol{E_1}(u, v, w=0)\, \boldsymbol{E_2}^*(0,0,0) \rangle$
- We interpret it as: $V^o_{12} = \langle \boldsymbol{E}'_1(u, v, w \neq 0)\, \boldsymbol{E_2}^*(0,0,0) \rangle$
- We should interpret E_1 as [E_1' x Fresnel Propagator]

PB Effects: Rotation asymmetry

- Only average quantities are available in the image domain
- Asymmetric PB rotation leads to time and direction dependent gains



$$\Delta I^R = \sum_{\psi} \left[PSF(\psi) - avgPSF \right] * \left[\left(PB(\psi) - avgPB \right) I^o \right]$$

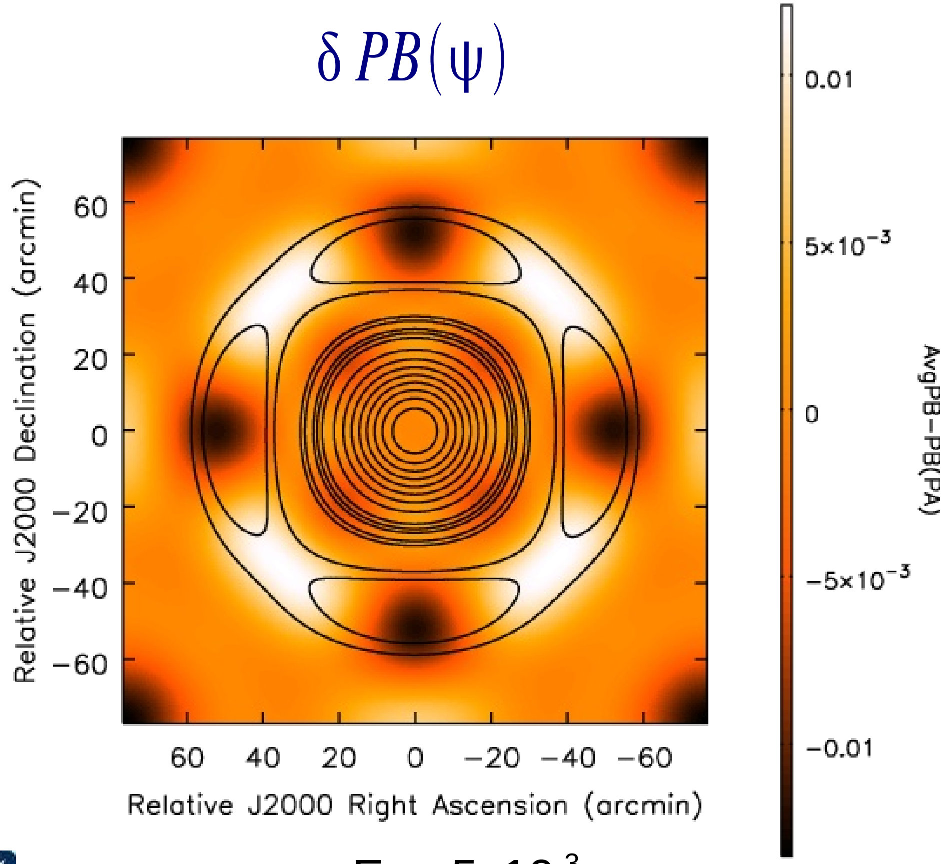


- Time-variability due to rotational asymmetry is stronger below ~10% point and in the side-lobes.
- Time-variability due to pointing errors is stronger at ~50% point.

PB Effects: Error Propagation

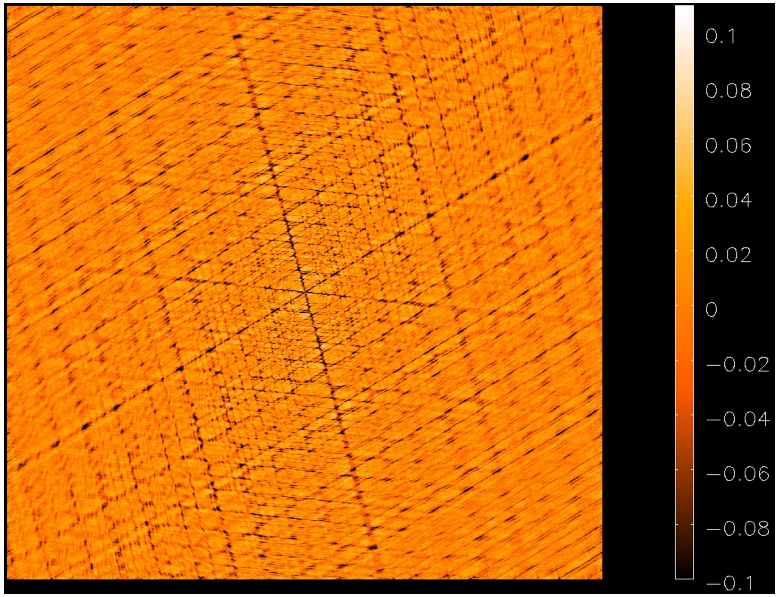
$$\Delta I^R = \sum_\psi \delta PSF(\psi) * [\delta PB(\psi) I^o]$$

$\delta PB(\psi)$



E.g. 5×10^{-3}

$\delta PSF(\psi)$



E.g. 2×10^{-2}

Projection algorithms

- Direction-dependent effects in the image domain are convolutional terms in the data domain

- Projection algorithms for DD corrections:
  - Project-out various DD effects as part of the gridding operator
    - ME:

$$V_{ij}^{Obs} = A_{ij} * V^{o} + N_{ij}$$

    - Construct D, such that

$$D_{ij}^{T} * A_{ij} \approx \text{Time/Freq./Pol.} \quad \text{indep.}$$

    - Imaging:

$$I = F^{-1} \sum_{ij} D_{ij}^{T} * V_{ij}^{Obs} = F^{-1} \frac{\sum_{ij} D_{ij}^{T} * A_{ij} * V_{ij}^{o} + D_{ij}^{T} * N_{ij}}{Normalization}$$

DI Corrections: Standard Calibration

- DI ME entirely in the visibility domain:

$$V_{ij}^{Obs} = \left[J_i \otimes J_j^*\right].\left[V_{ij}^o\right] = \left[M_{ij}\right].\left[V_{ij}^o\right]$$

$$\begin{bmatrix} V_{pp}^{Obs} \\ V_{pq}^{Obs} \\ V_{qp}^{Obs} \\ V_{qq}^{Obs} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix} . \begin{bmatrix} V_{pp}^o \\ V_{pq}^o \\ V_{qp}^o \\ V_{qq}^0 \end{bmatrix}$$

- Diagonal: “pure” poln. products
- Off-diagonal: Include poln. leakage

$$M_{ij} = g_i g_j^* = G_{ij}$$

- DI Correction

$$V_{ij}^{Corr} = G_{ij}^{-1} V_{ij}^{obs}$$

$$G_{ij}^{-1} = \frac{G^*}{\left|G_{ij}\right|^2}$$

gaincal, bandpass, gencal, applycal, polcal,...

DI Corrections: Standard Calibration

- DI ME entirely in the visibility domain:

$$V_{ij}^{Obs} = \left[J_i \otimes J_j^*\right] . \left[V_{ij}^o\right] = \left[M_{ij}\right] . \left[V_{ij}^o\right]$$

$$\begin{bmatrix} V_{pp}^{Obs} \\ V_{pq}^{Obs} \\ V_{qp}^{Obs} \\ V_{qq}^{Obs} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix} . \begin{bmatrix} V_{pp}^o \\ V_{pq}^o \\ V_{qp}^o \\ V_{qq}^0 \end{bmatrix}$$

- Diagonal: “pure” poln. products
- Off-diagonal: Include poln. leakage

$$M_{ij} = g_i g_j^* = G_{ij}$$

- Full-pol. DI Correction

$$V_{ij}^{Corr} = \left[M_{ij}^{M^{-1}}\right] . \left[V_{ij}^{Obs}\right] = \frac{adj(M_{ij}^M)}{det(M_{ij}^M)} . \left[V_{ij}^{Obs}\right]$$

Equivalent Complex math.: $G_i^{-1} = \frac{G^*}{|G|^2}$

No pol. leakage case: $= \frac{G_{q,ij}^{M^*}}{G_{p,ij}^M G_{q,ij}^{M^*}}$

gaincal, bandpass, gencal, applycal, polcal,...

DD Corrections: Projection algorithms

- DD ME entirely in the visibility domain:

$$V_{ij}^{Obs} = \left[J_i \otimes J_j^*\right] * \left[V_{ij}^o\right] = \left[M_{ij}\right] * \left[V_{ij}^o\right]$$

J : Each elements is a function

$\otimes$ = Element-by-element convolution

$$\begin{bmatrix} V_{pp}^{Obs} \\ V_{pq}^{Obs} \\ V_{qp}^{Obs} \\ V_{qq}^{Obs} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix} * \begin{bmatrix} V_{pp}^o \\ V_{pq}^o \\ V_{qp}^o \\ V_{qq}^o \end{bmatrix}$$

- Diagonal: "pure" poln. PBs
- Off-diagonal: In-beam poln. leakage

$$M_{pq} = J_{p,i} * J_{q,j}^*$$

- Full-pol. DD corrections

$$V_{ij}^{Corr} = \left[M_{ij}^{M^{-1}}\right] * \left[V_{ij}^{Obs}\right] = \frac{adj(M_{ij}^M)}{det(M_{ij}^M)} * \left[V_{ij}^{Obs}\right]$$

$$I_{ij}^{Corr} = \frac{F\left[adj(M_{ij}^{M^T})\right] * \left[V_{ij}^{Obs}\right]}{F\, det(M_{ij}^M)}$$

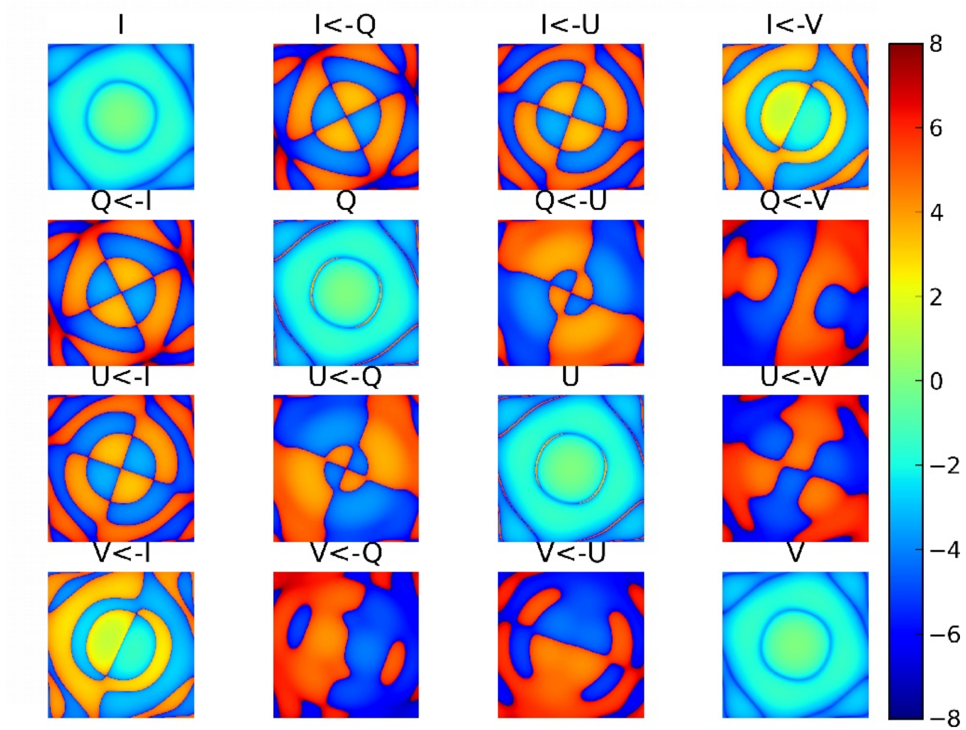
During gridding

Image plane normalization

DD Effects in Full-pol. Imaging

- DD "Mueller" matrix:

$$\begin{bmatrix} I_I^{Obs} \\ I_Q^{Obs} \\ I_U^{Obs} \\ I_V^{Obs} \end{bmatrix} =$$



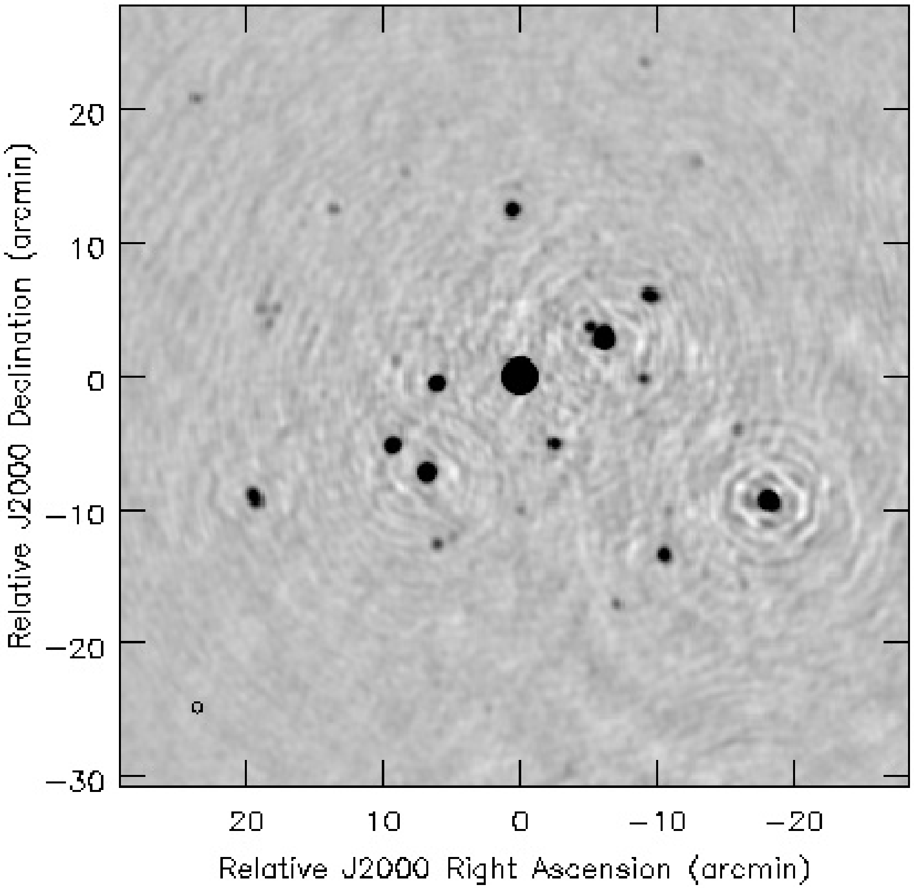
Stokes Basis

$$\cdot \begin{bmatrix} I_I^{o} \\ I_Q^{o} \\ I_U^{o} \\ I_V^{0} \end{bmatrix}$$

- Affects DR at the $10^{3\text{-}4}$ level
- PB Stokes-Q, -U is few% of Stokes-I

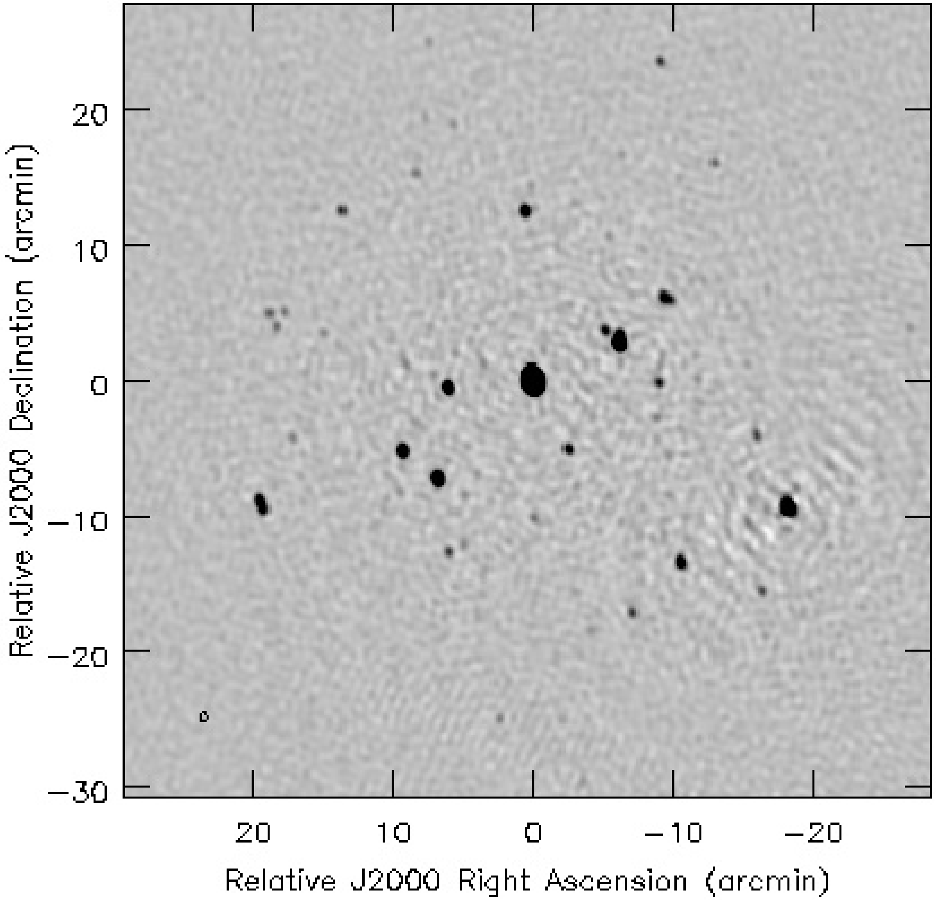
[Jagannathan, et al., AJ, 2017]

A-Projection: Stokes-I Before



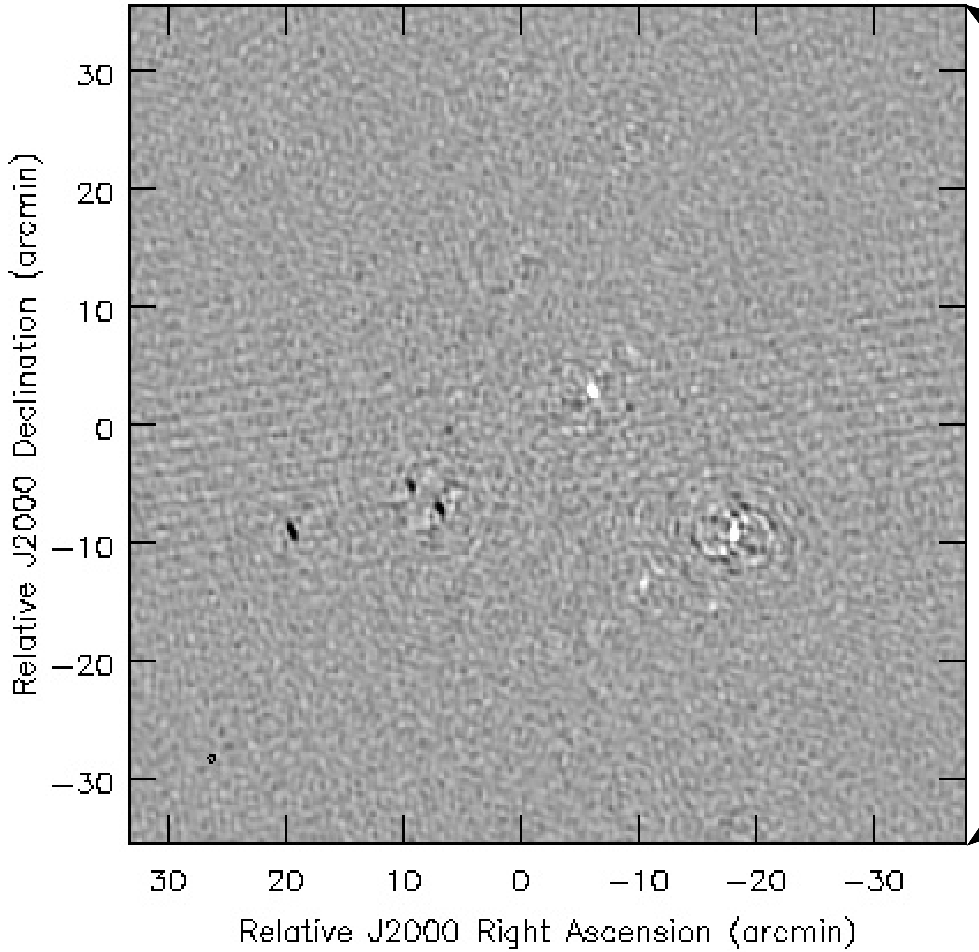
Effective PB is time-variant

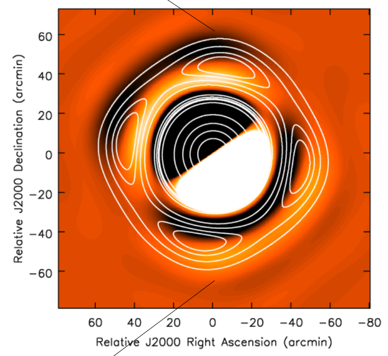
A-Projection: Stokes-I After



Effective PB is time-invariant

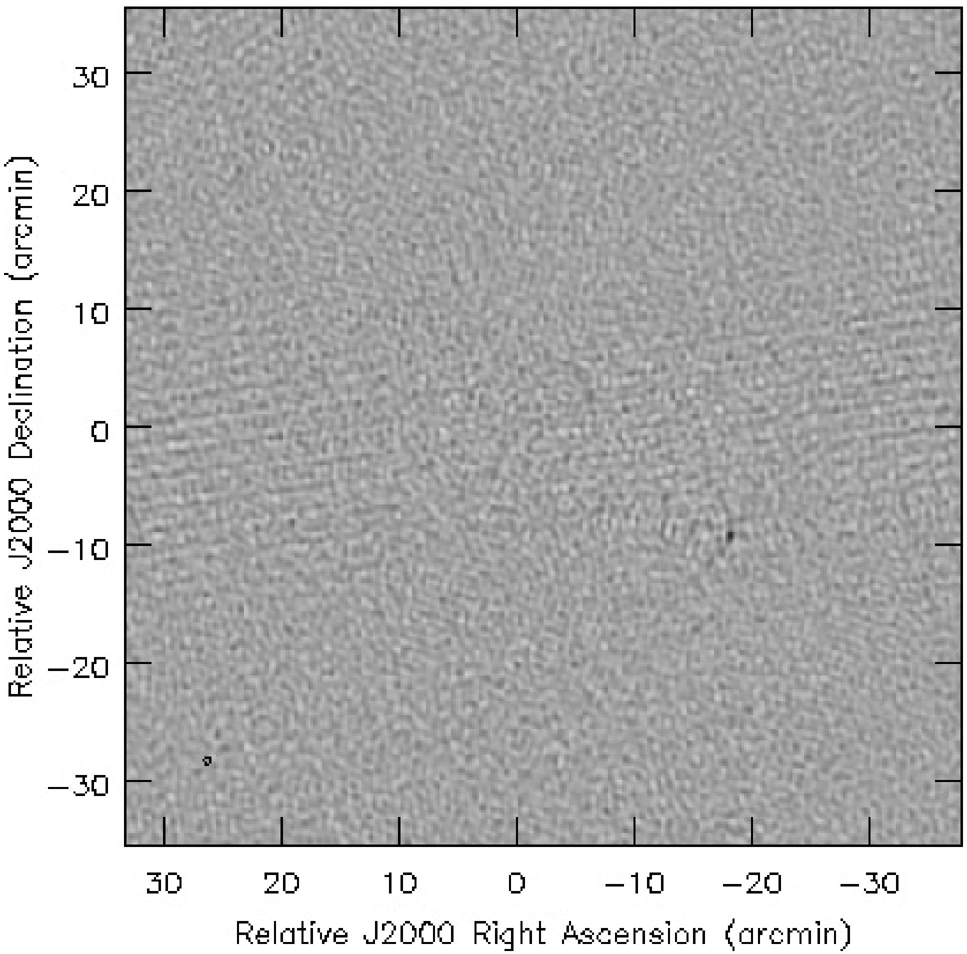
A-Projection: Stokes-V Before





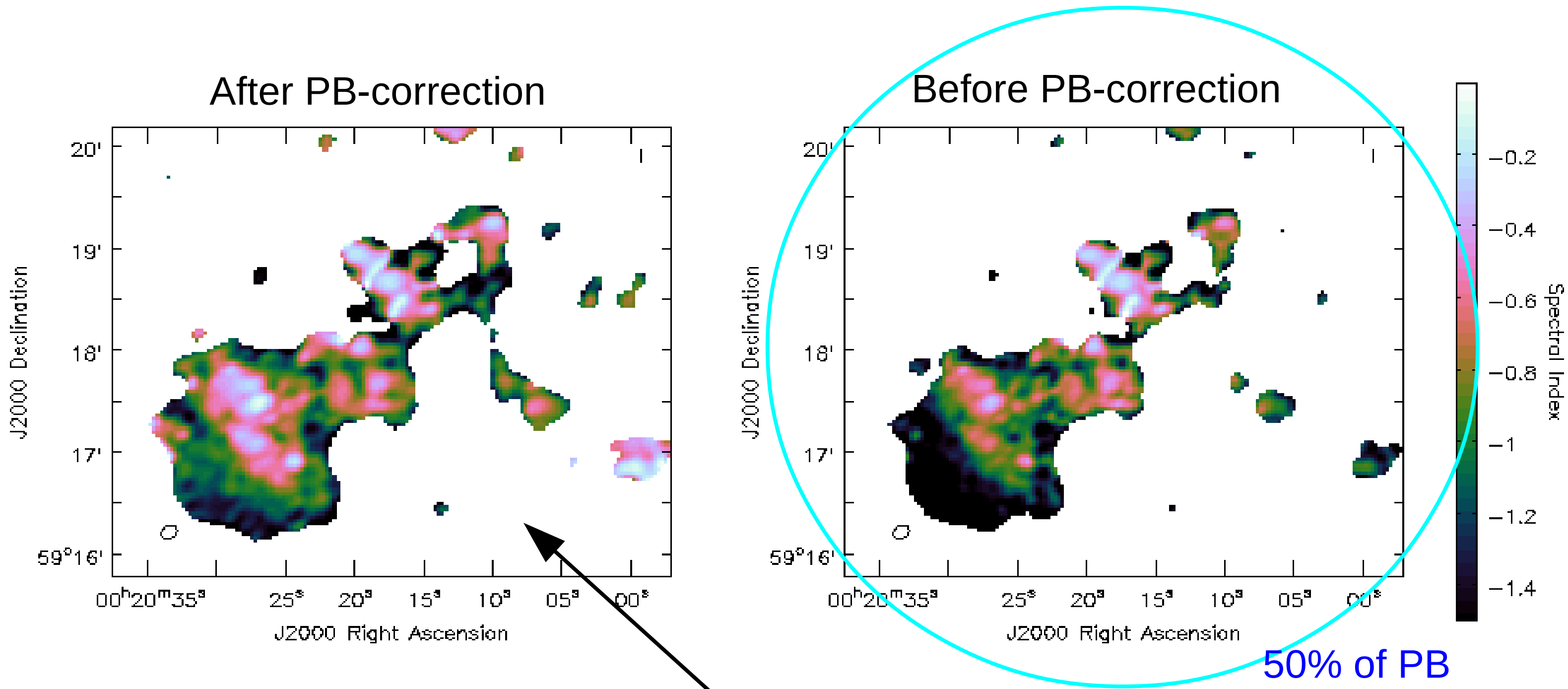
Effective PB is polarization-variant

A-Projection: Stokes-V After



Effective PB is polarization-invariant

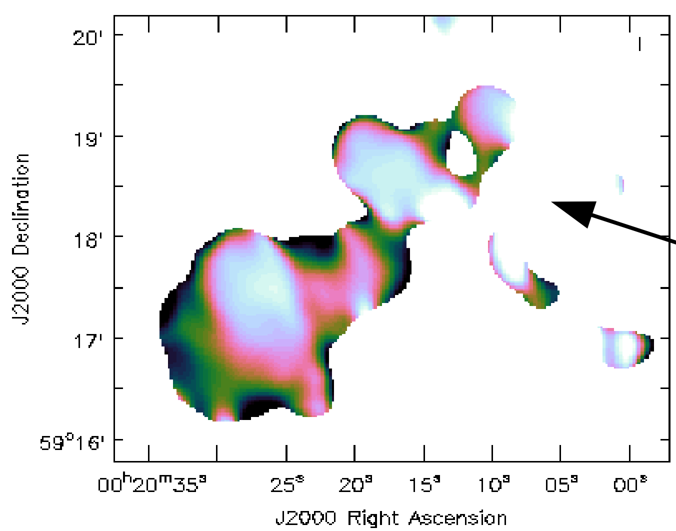
Continuum (MS-MFS) vs Cube Imaging (with PB-correction)



IC10 Dwarf Galaxy :

Spectral Index across C-Band.

Dynamic-range ~ 2000 (~ noise-limited image obtained)

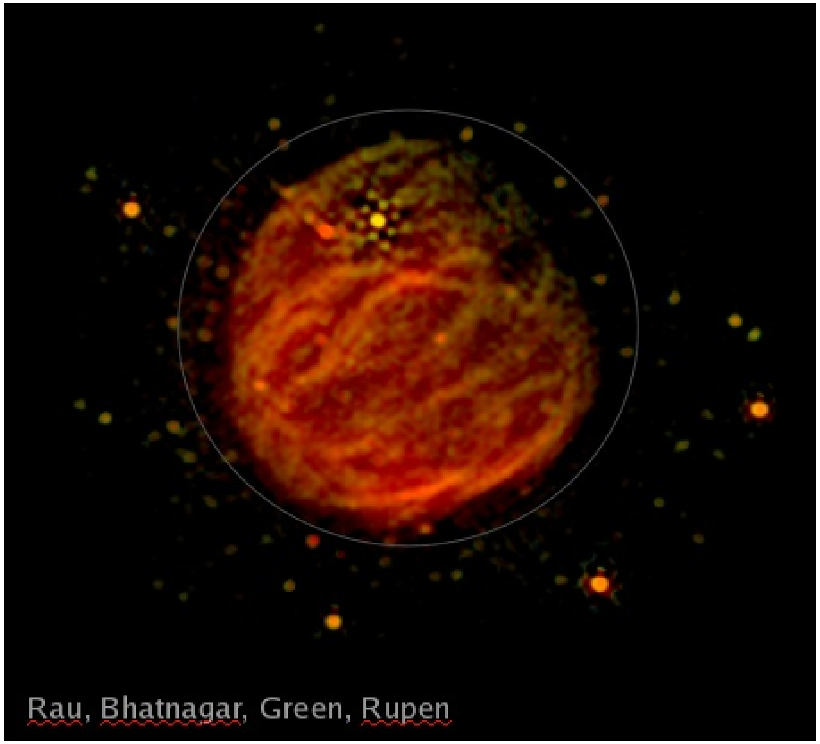


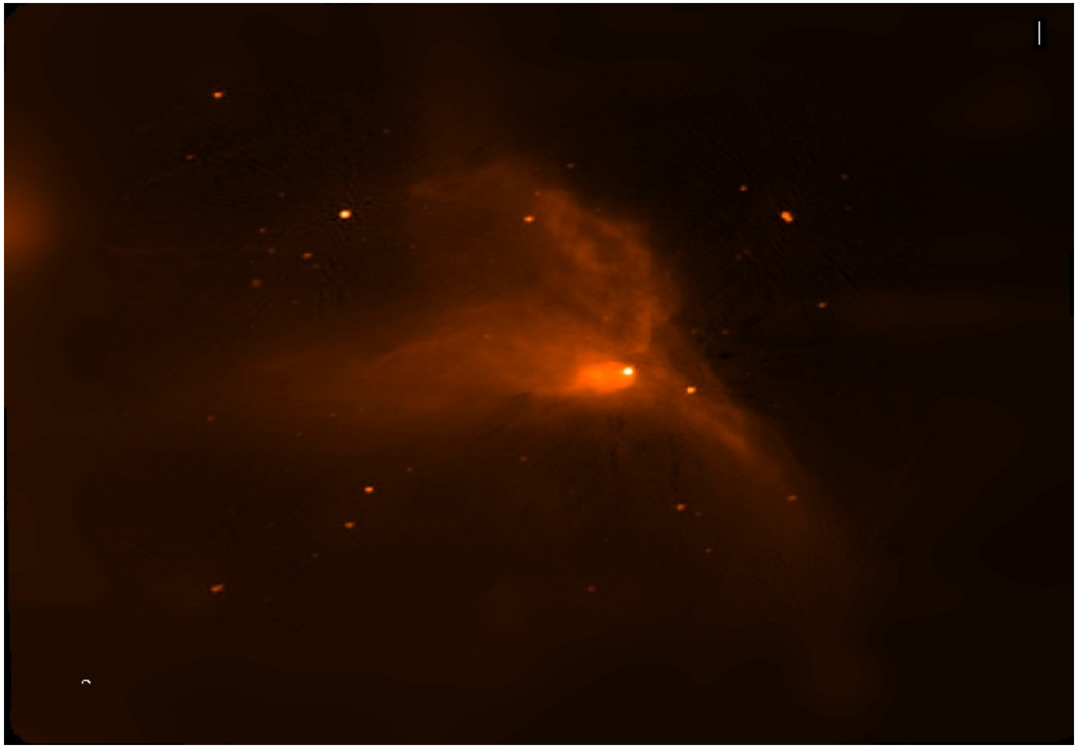
MS-MFS :
Result of wide-band PB-correction after MT-MS-MFS.

Cube :
Spectral-index map made by PB-correcting single-SPW images smoothed to the lowest resolution.

Scales for Multi-Scale Deconvolution

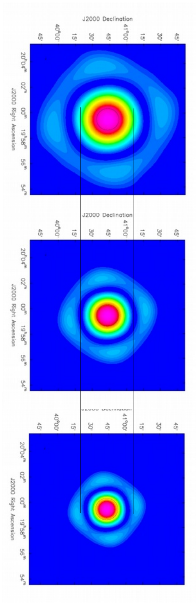
- Thumb rule for selecting largest scale
  - Smallest dimension of the largest scale in the image

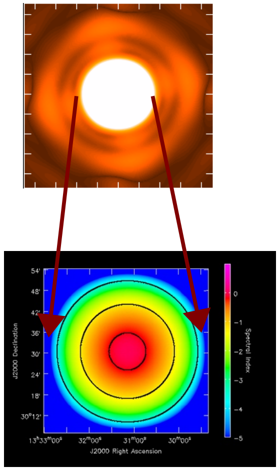




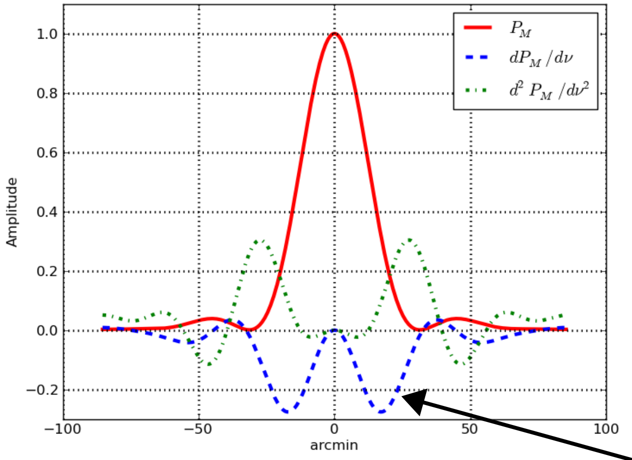
Wide-band Wide-field Imaging

- Wide band data to image beyond the ~50% point of the PB at a reference frequency
  - Bandwidth ratio > ~20%
  - FoV > ~HPBW @ reference frequency
  - Variable PB:
    - Long integration (rotation), Mosaicking (pointings at different PA), in-beam polarization is large (AA)



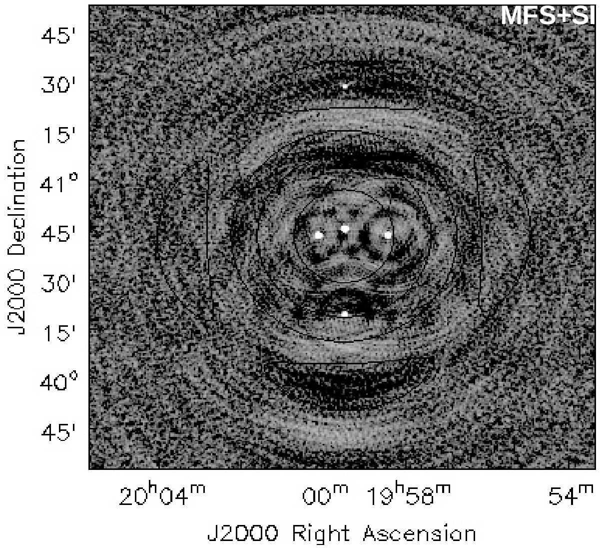


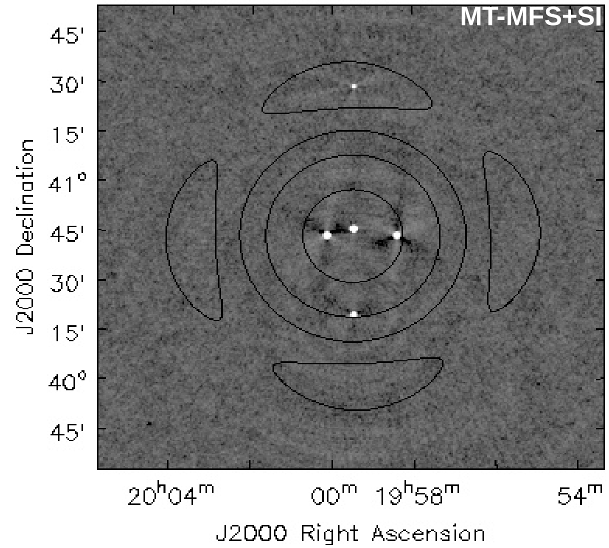
PB "Spectral Index"

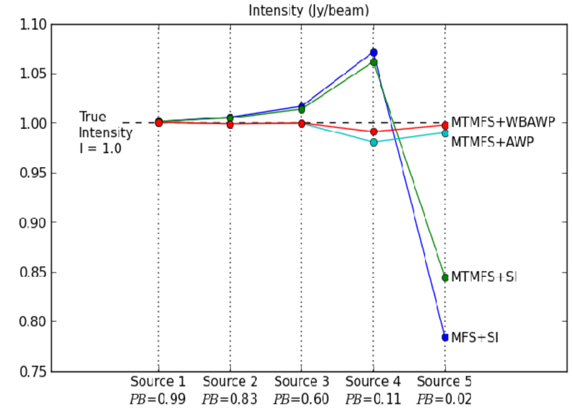


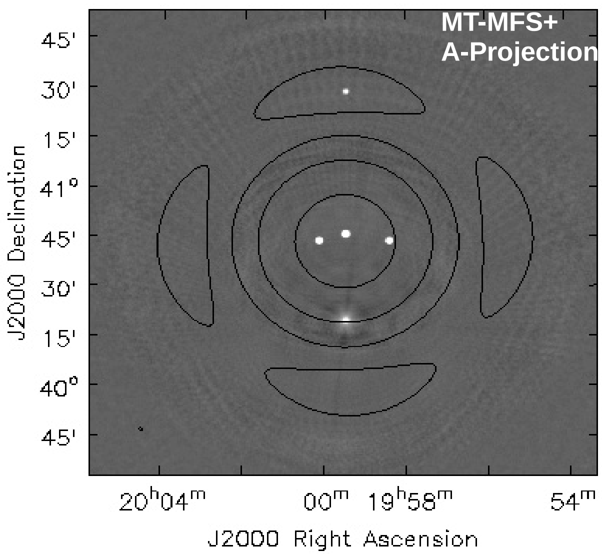
Wide-band Wide-field Imaging

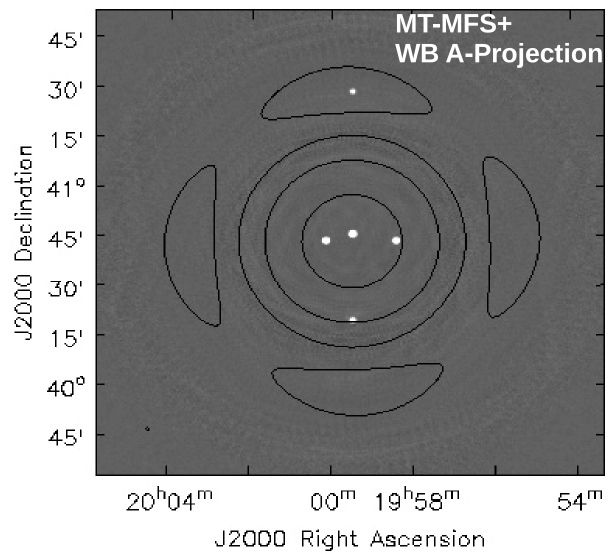
- Characterization of the (WB) A-Projection + MT-MFS

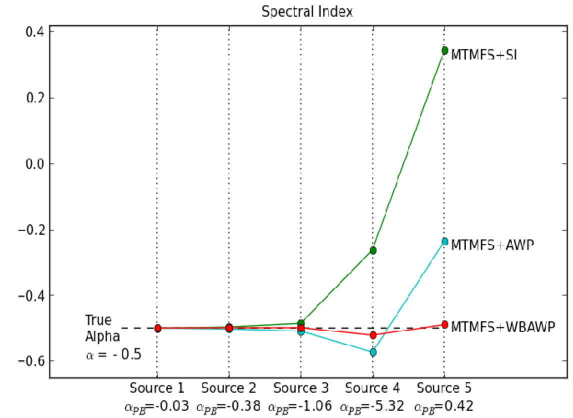






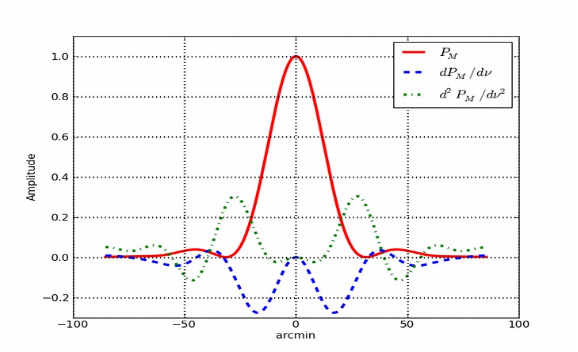


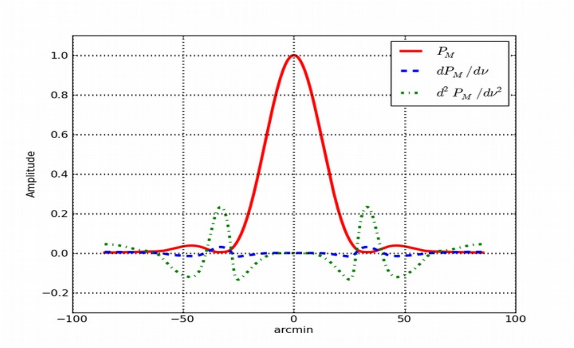




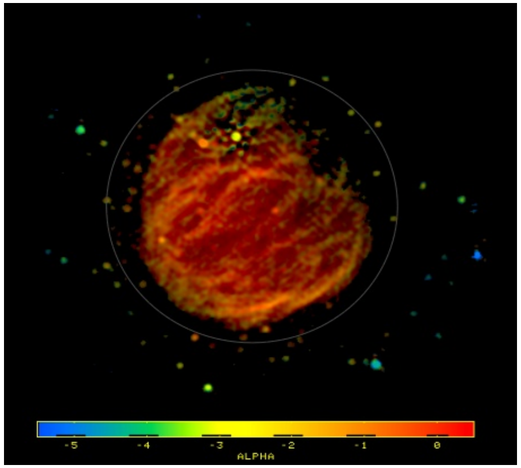
Wide-band Wide-field Imaging

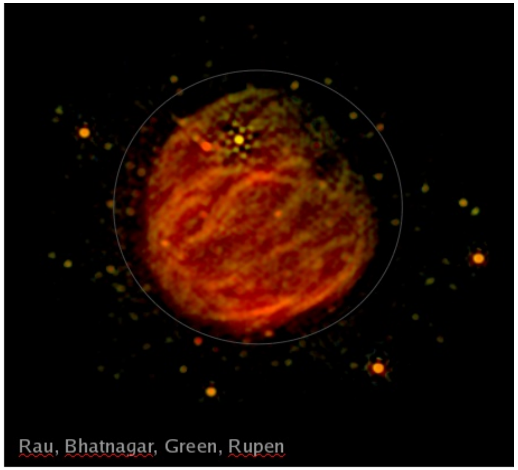
- WB A-Projection + MT-MFS
  - WB A-Projection for PB





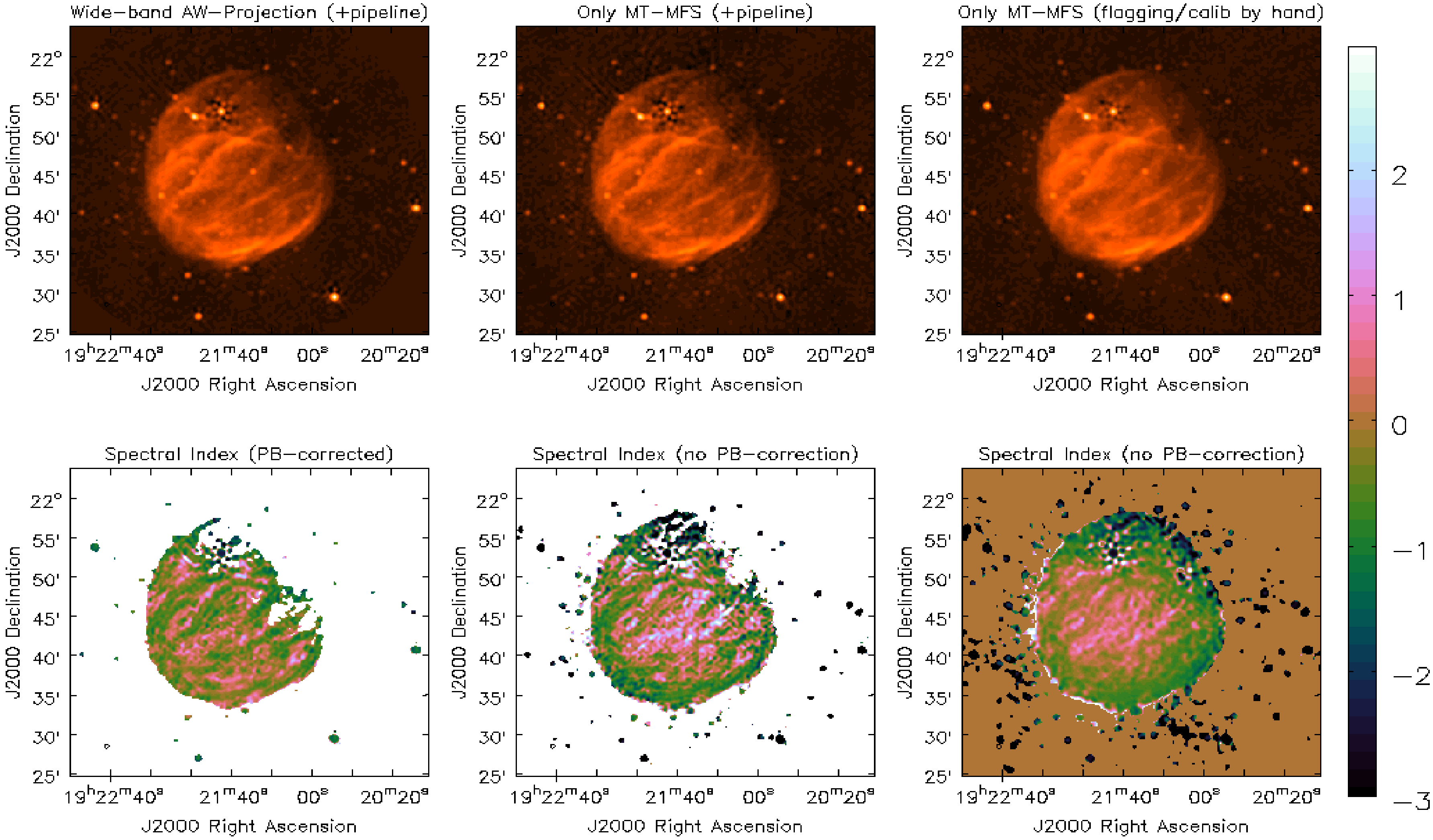
  - MT-MFS for sky





Wide-band Wide-field Imaging

- WB A-Projection + MT-MFS

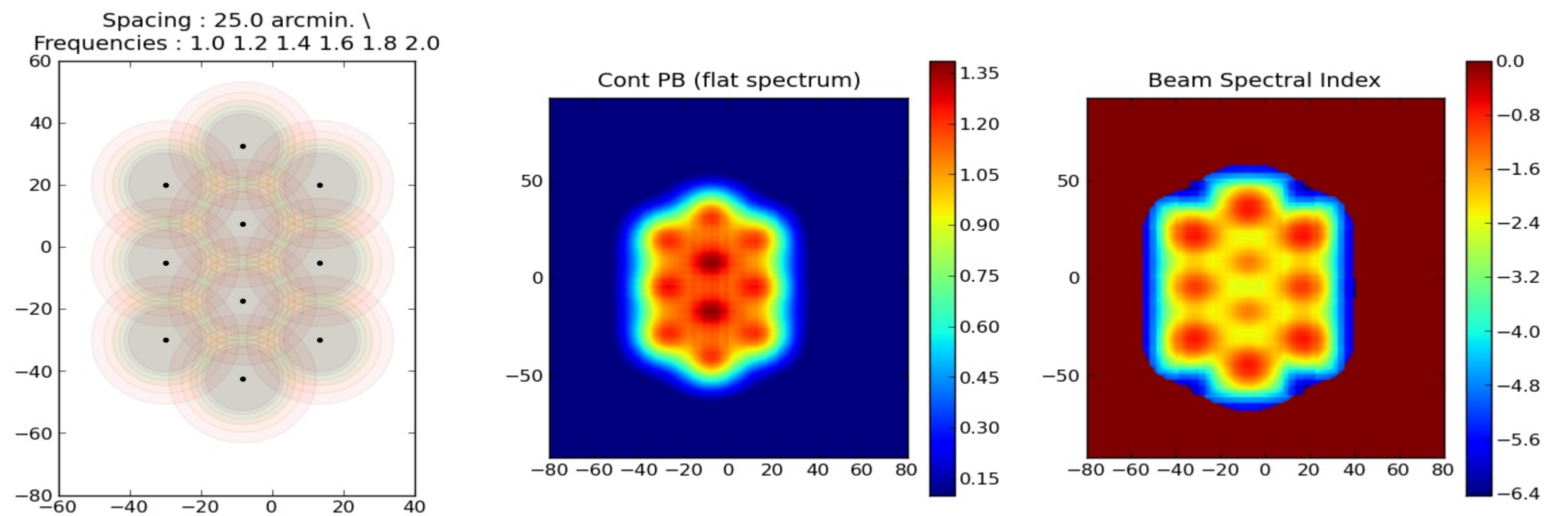


Wideband Primary Beams – Mosaic

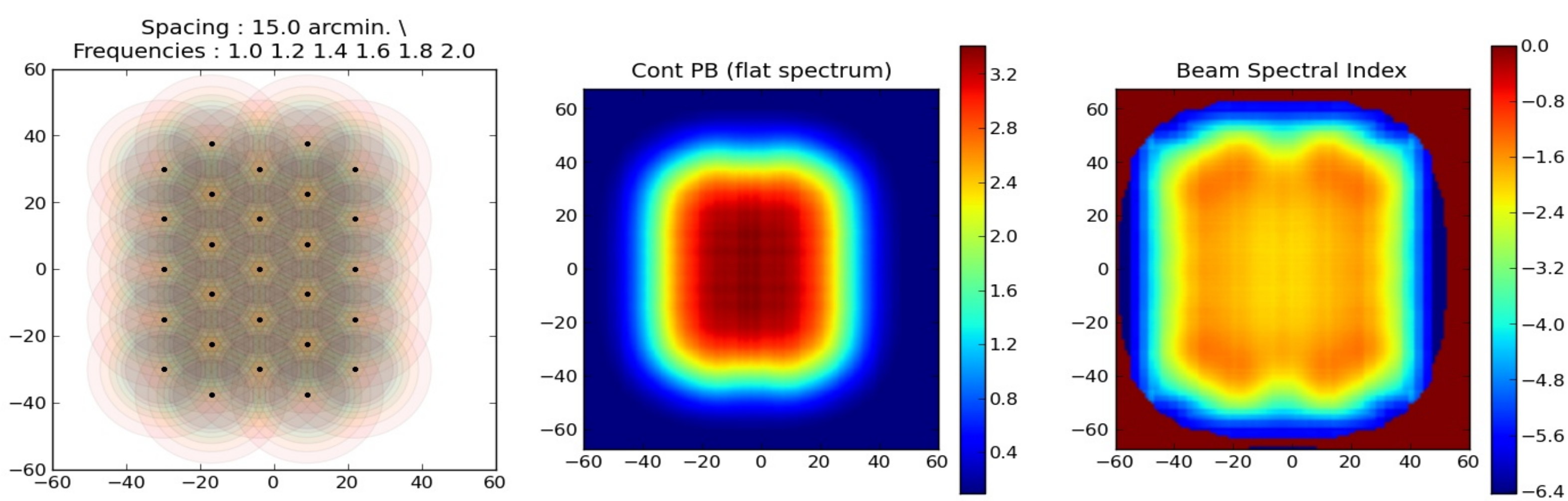
For single pointings, the wideband PB spectrum is relevant only away from the pointing center.

For mosaics, the wideband PB spectrum must be accounted-for <u>all over the mosaic field of view</u>

25 arcmin spacing (1-2 GHz)



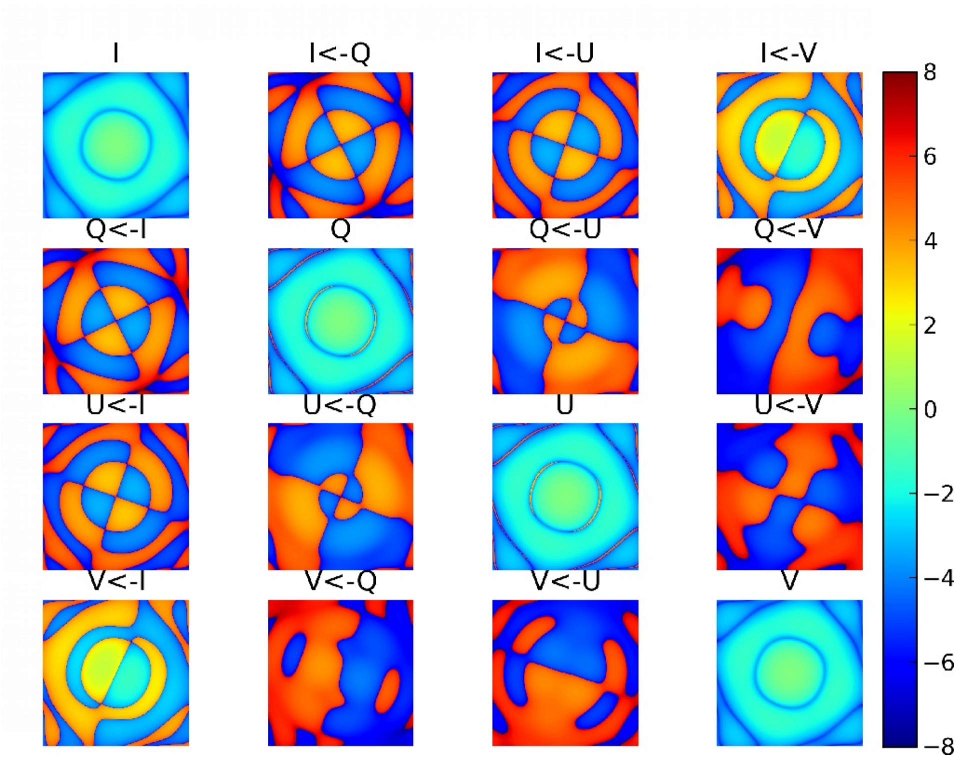
15 arcmin spacing (1-2 GHz)



DD Effects in Full-pol. Imaging

- DD “Mueller” matrix:

$$\begin{bmatrix} I_I^{Obs} \\ I_Q^{Obs} \\ I_U^{Obs} \\ I_V^{Obs} \end{bmatrix} =$$



Stokes Basis

$$\cdot \begin{bmatrix} I_I^o \\ I_Q^o \\ I_U^o \\ I_V^0 \end{bmatrix}$$

- Affects DR at the $10^{3\text{-}4}$ level
- PB Stokes-Q, -U is few% of Stokes-I

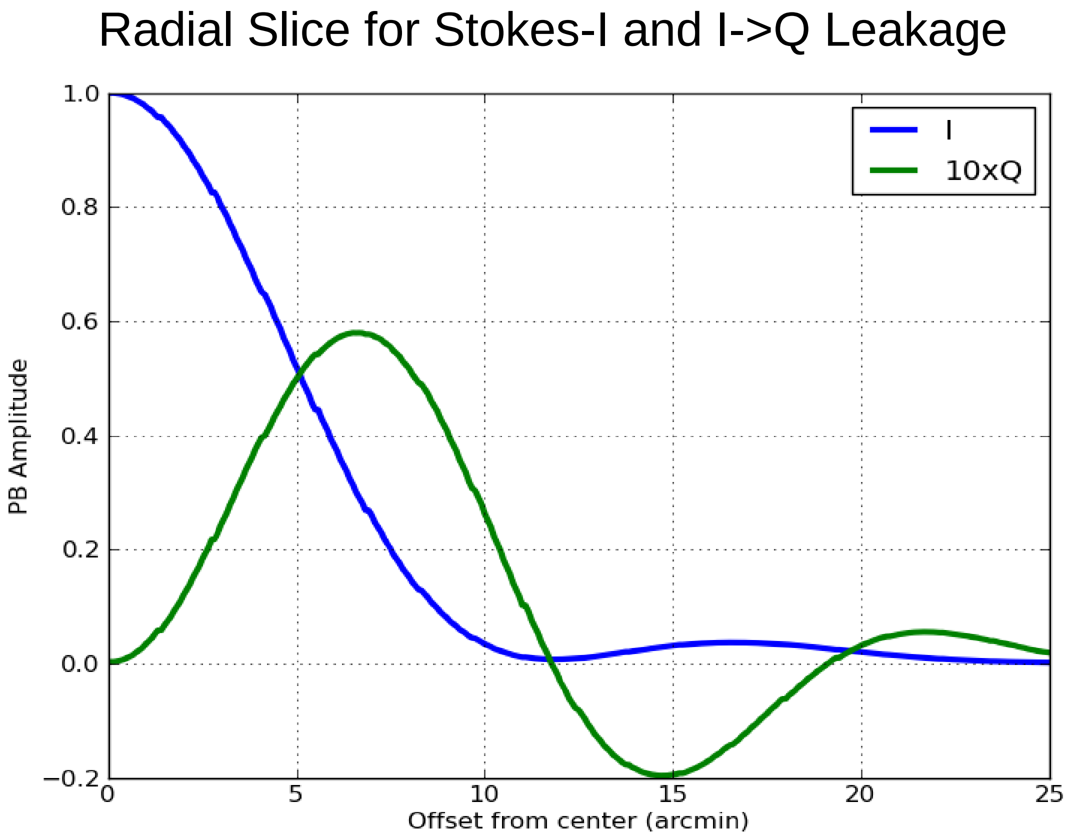
[Jagannathan et al., AJ, 2017; PhD Thesis]

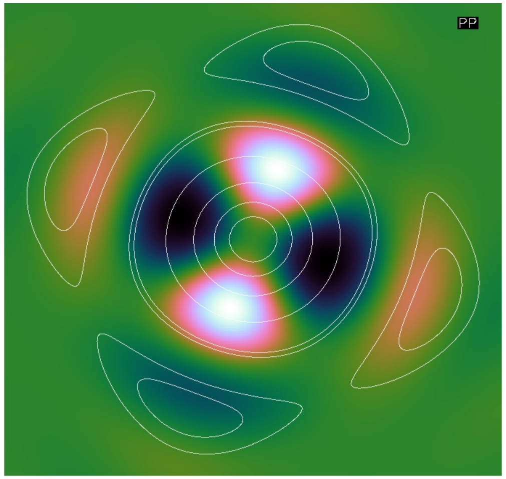
Issues in Wide-field Wide-band Full-Pol. Imaging

- PB Effects
  - In-beam effects : DD Leakage
  - Parametric Aperture Illumination model (Holographic measurements not sufficient)
  - Pointing Errors
  - Mosaic patterns
- Variations with frequency
  - Frequency dependence of intrinsic Q and U
  - Frequency dependence due to PB
- Computing load
  - Larger CF for wide-field imaging: Fundamentally more expensive
  - Larger memory footprint: Fundamentally required, any which way you cut it

Full-pol. Imaging: In-beam Leakages

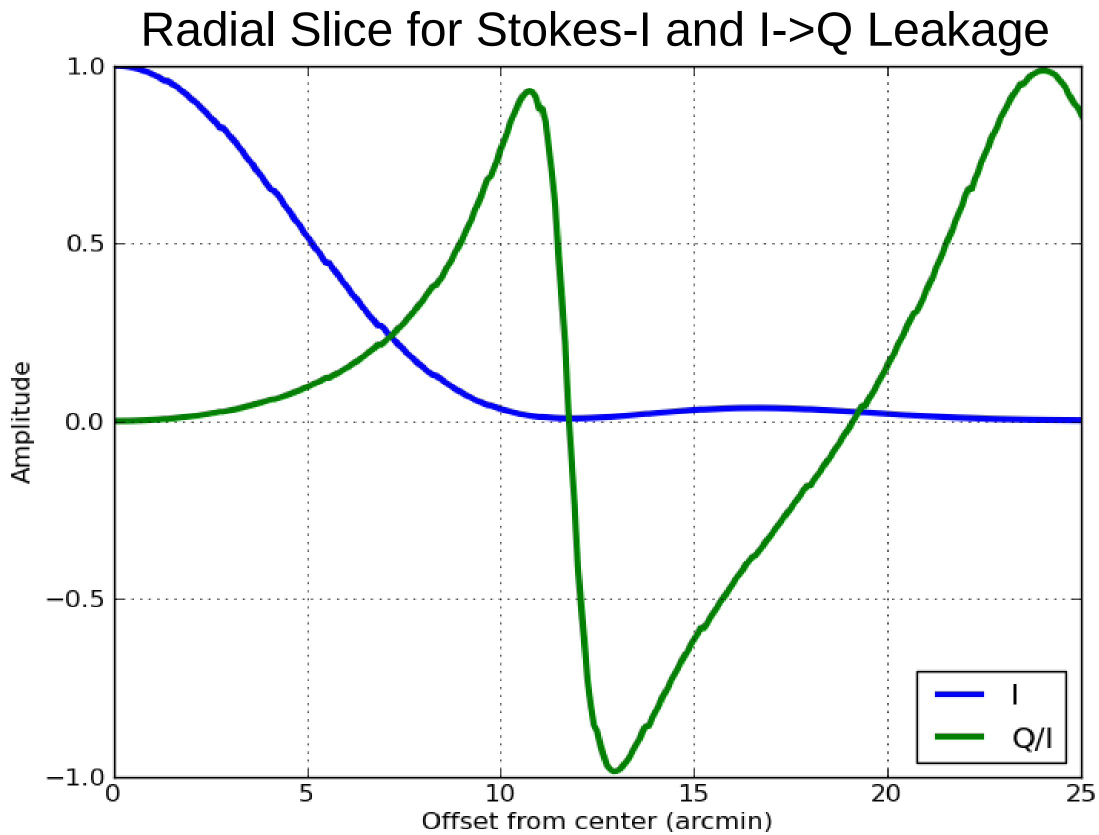
- Leakage (Off-diagonal elements of the Mueller matrix)
  - Vary with direction (position in the beam), Parallactic Angle (time) and frequency

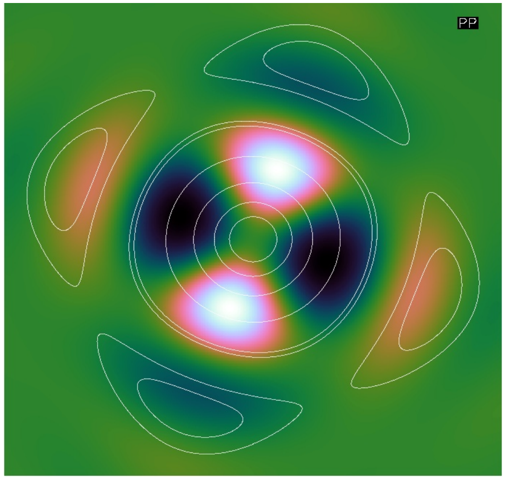




Full-pol. Imaging: In-beam Leakages

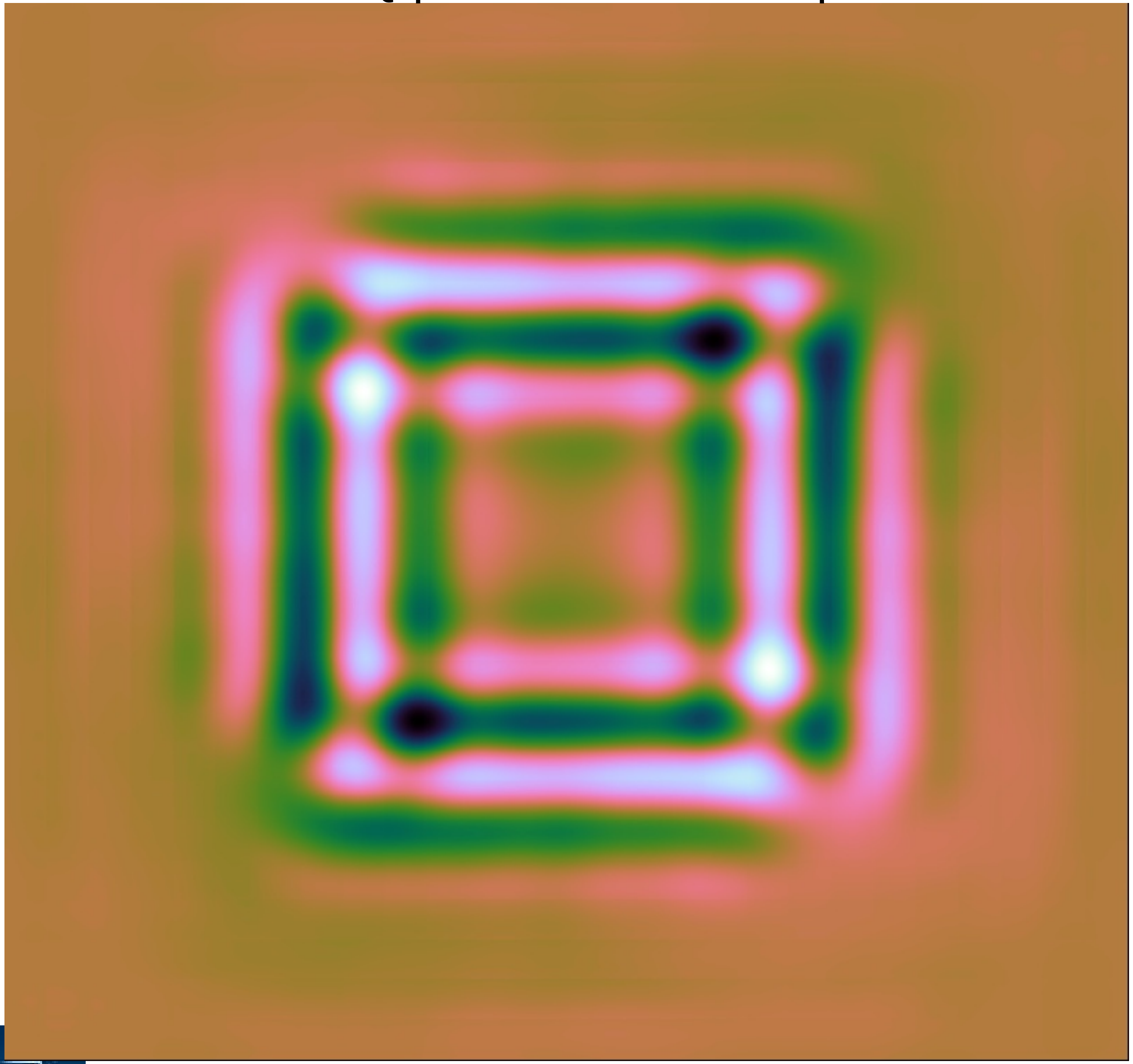
- Leakage (Off-diagonal elements of the Mueller matrix)
  - Vary with direction (position in the beam), Parallactic Angle (time) and frequency





Full-pol. Imaging: Mosaic Sensitivity Pattern

In-beam Stokes-Q pattern for a 11x11 point mosaick

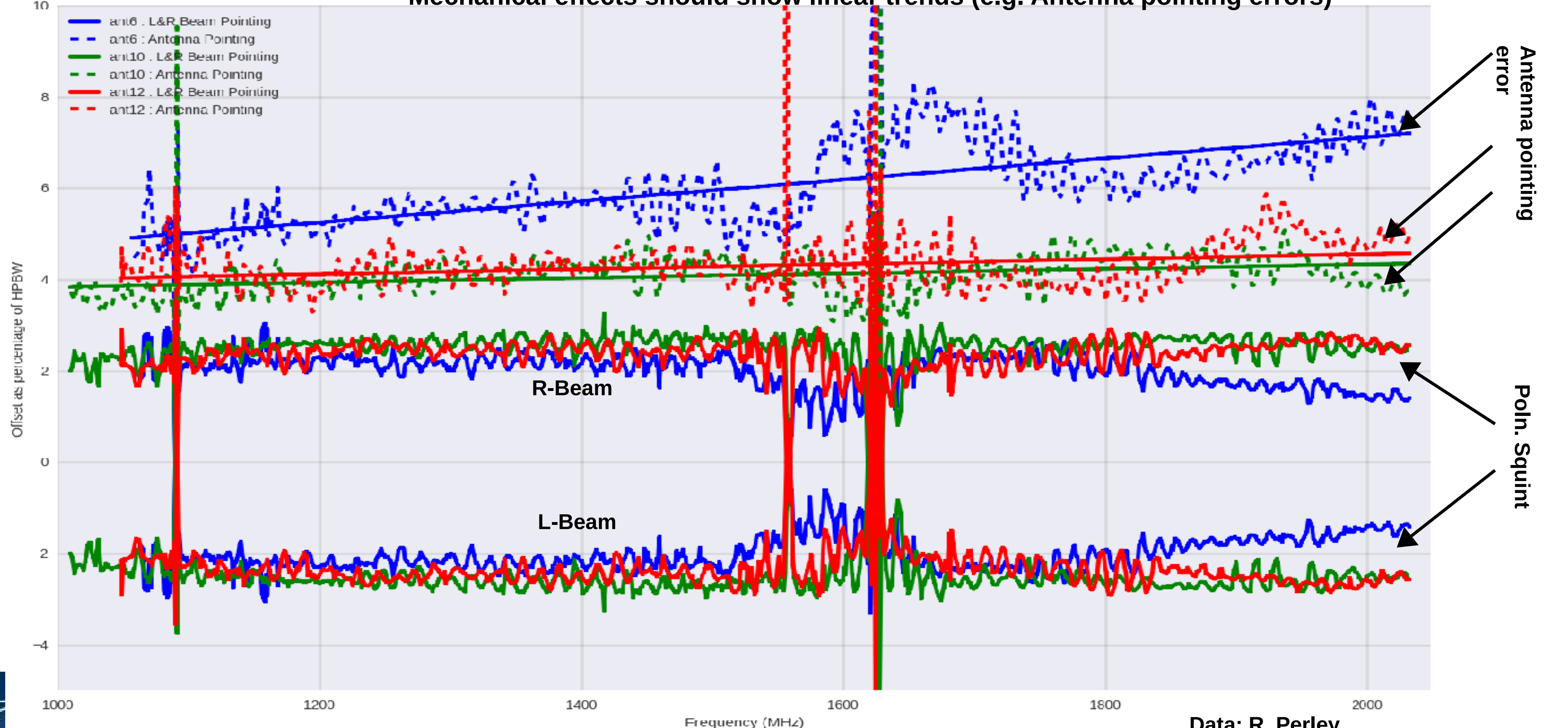


- Heterogeneous case; rotation due to PA change also ignored
- The resulting pattern is combination of overlapping Clover-leaf pattern of each pointing
- In-beam DD leakage spreads all across the mosaicked region.

Full-pol. Imaging: PB Effects

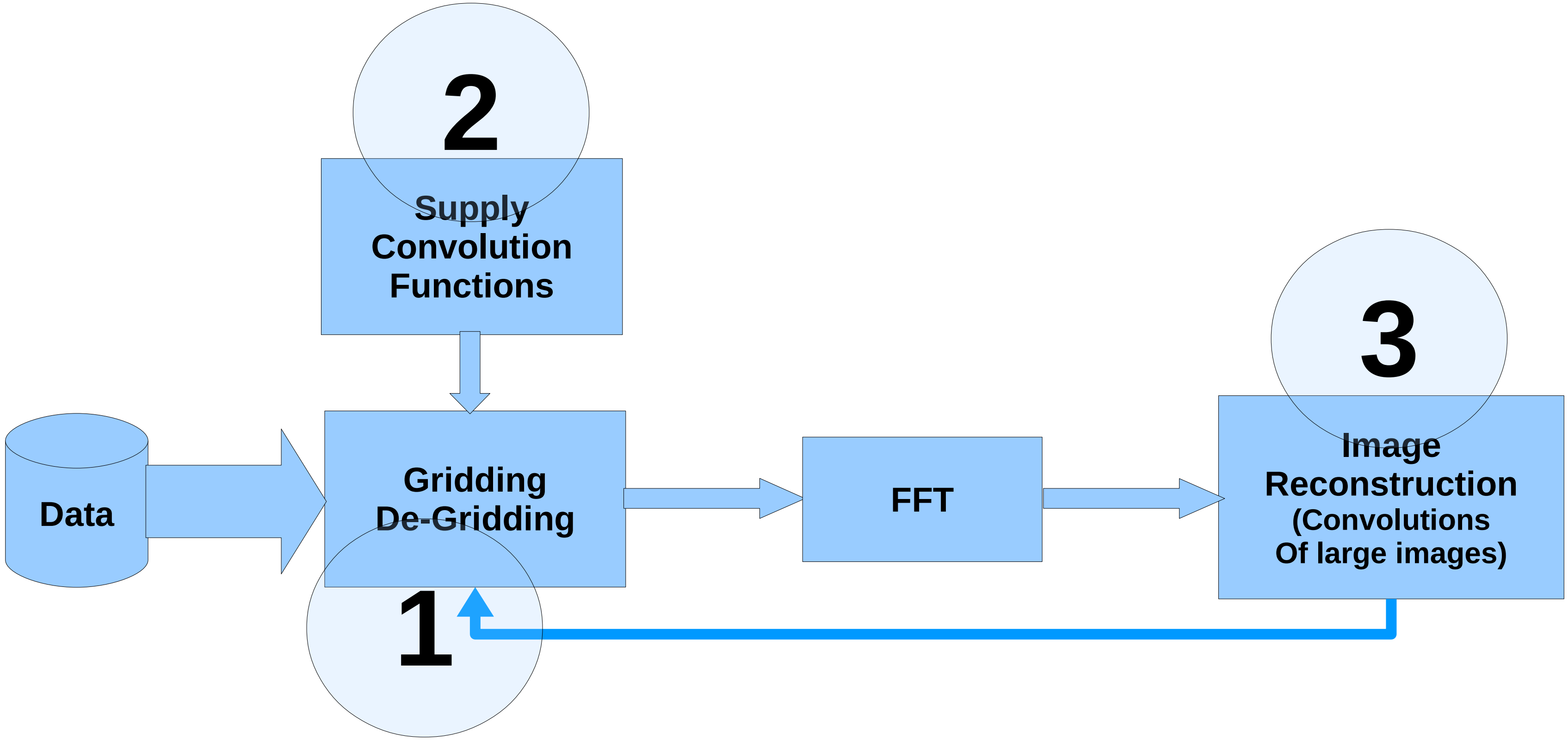
- Parametric model of antenna Aperture Illumination
  - Difference between Ant6 and Ant10 in “homogeneous array”

In the graph below: Optical effects should be independent of frequency (e.g. Poln. Squint)
Mechanical effects should show linear trends (e.g. Antenna pointing errors)

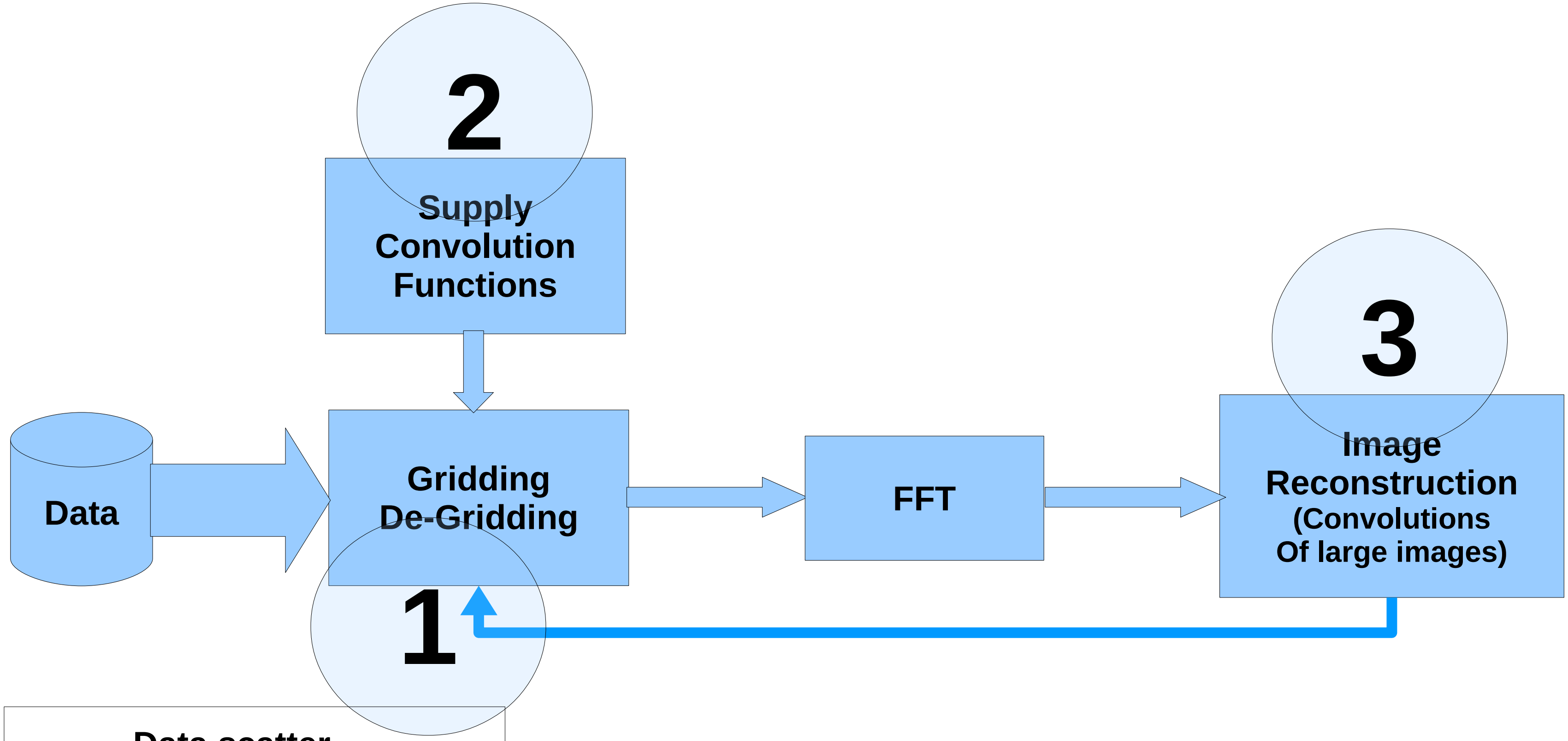


Data: R. Perley
Analysis: P.Jagannathan, S.Bhatnagar

The compute hot-spots



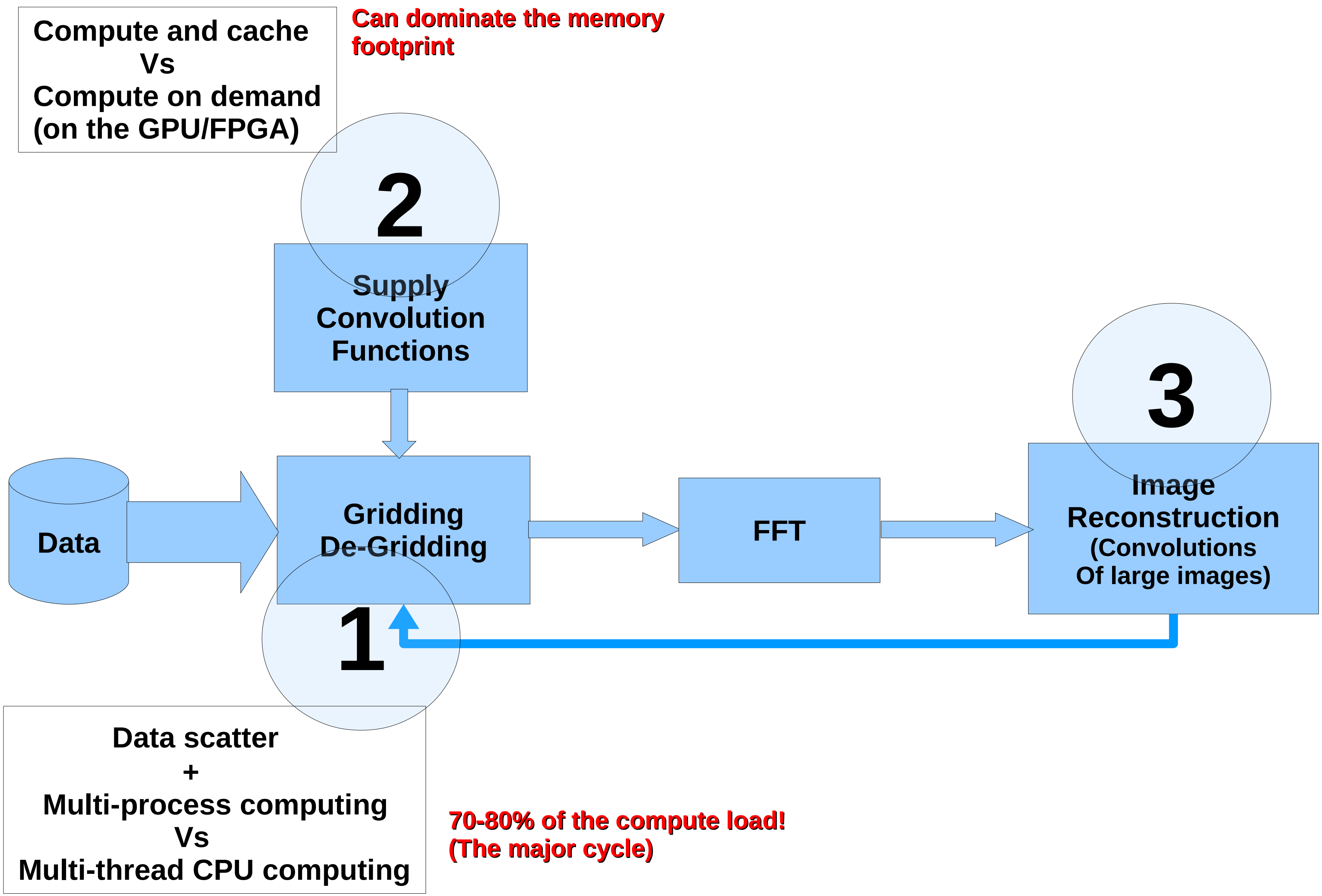
The compute hot-spots



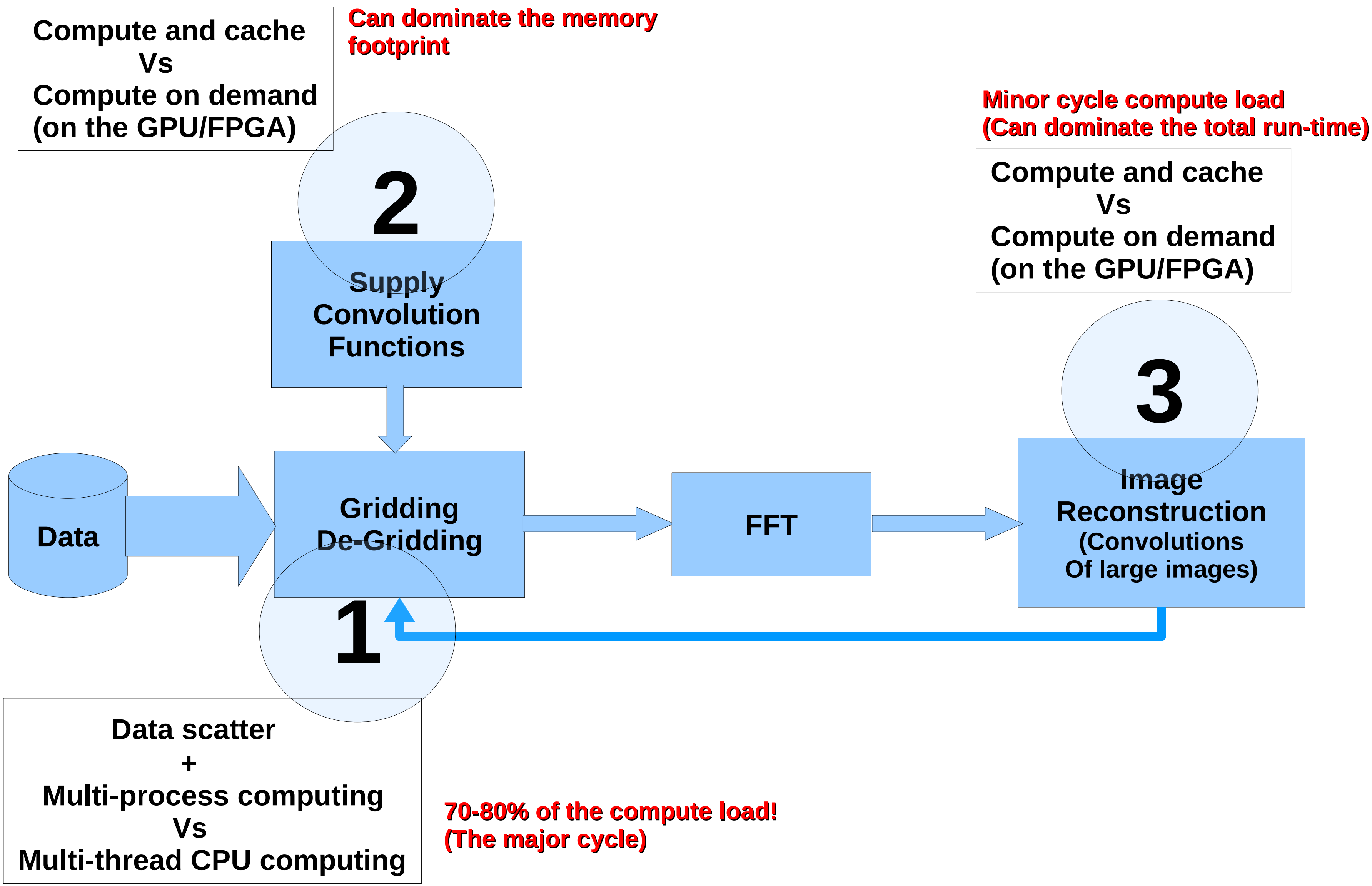
Data scatter
+
Multi-process computing
Vs
Multi-thread CPU computing

70-80% of the compute load!
(The major cycle)

The compute hot-spots

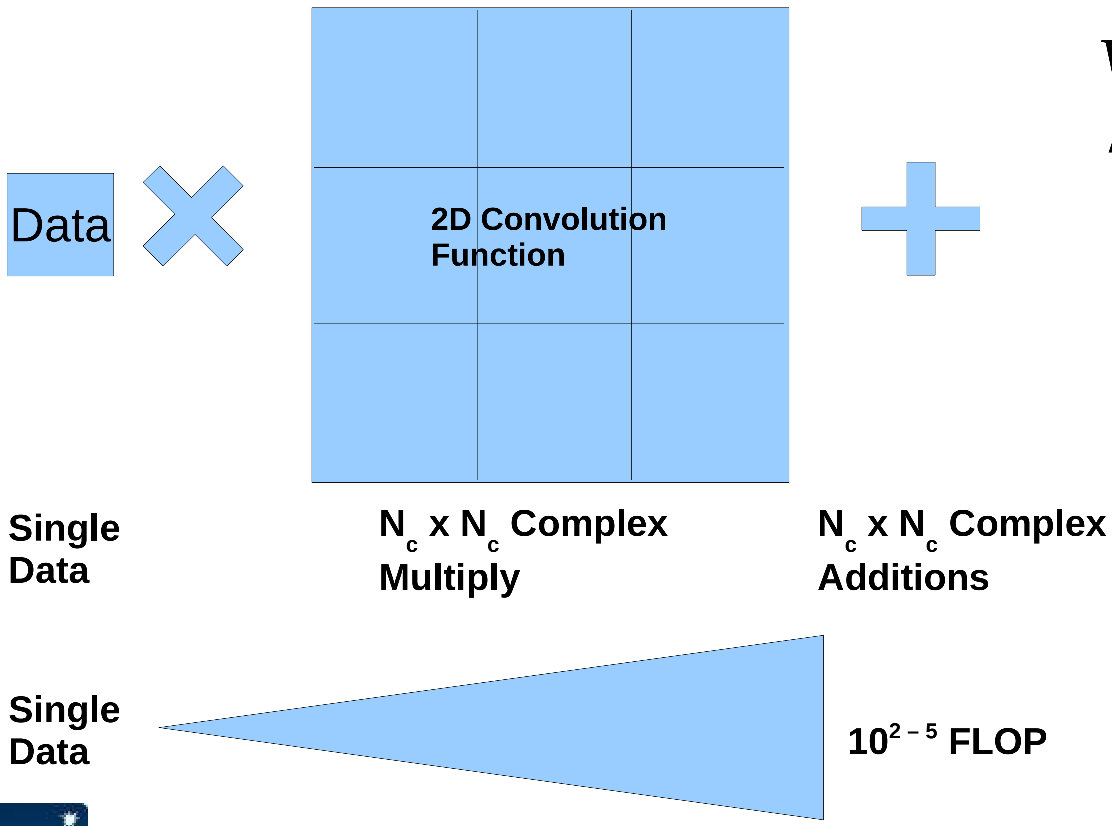


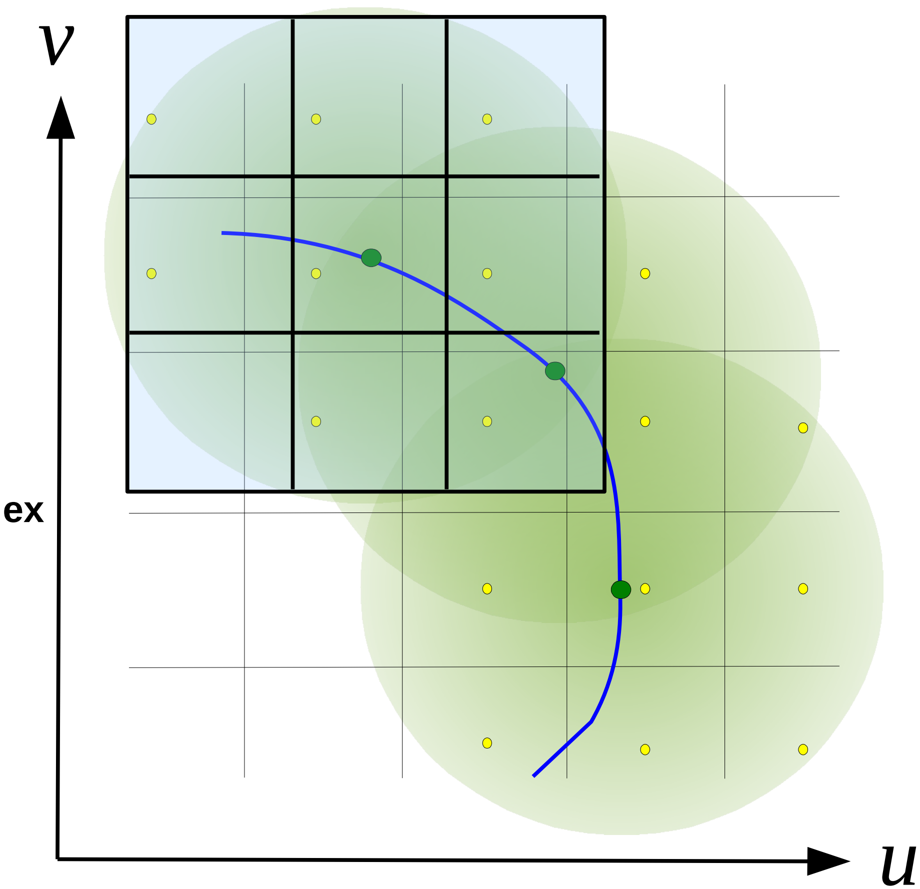
The compute hot-spots



Gridding: Computations

- Gridding/de-gridding: 2D interpolation via convolutional resampling
- 2D convolution functions ← → 2D weighting functions

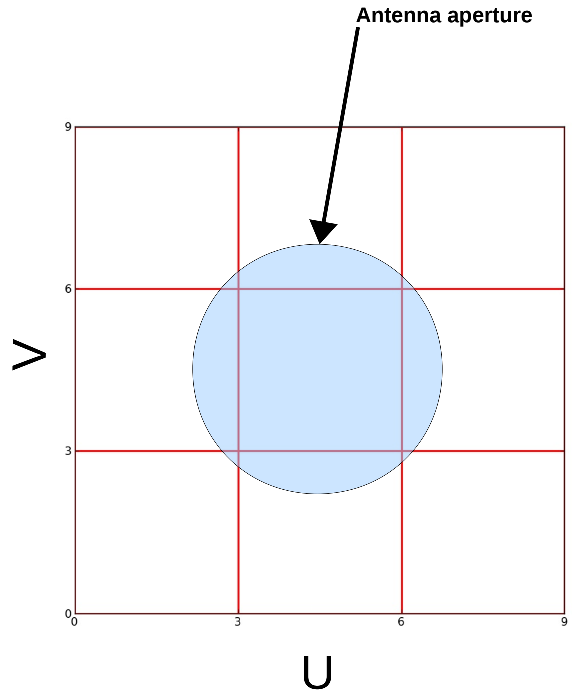




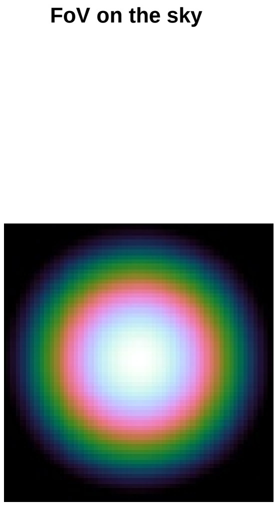
N_{data} x 12 N_C^2 N_{data} = $O(10^{10 -- 12})$

WF imaging: A-Projection

- WF imaging needs larger convolution functions (CF)



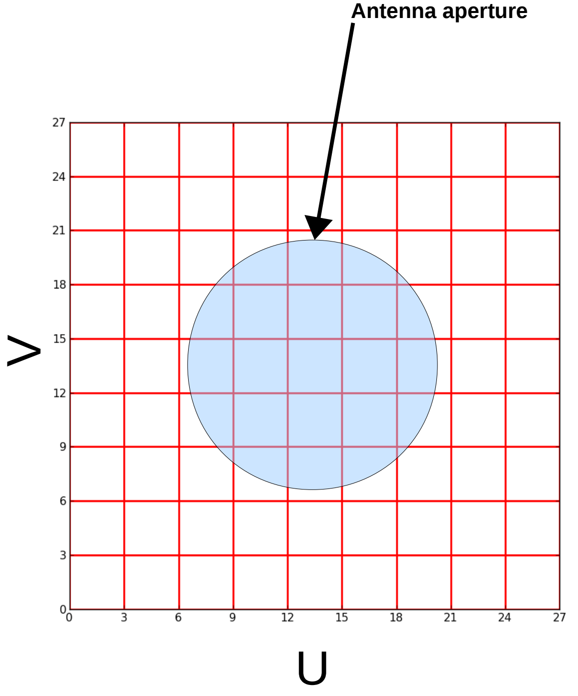
Number of uv-pixel
across antenna aperture



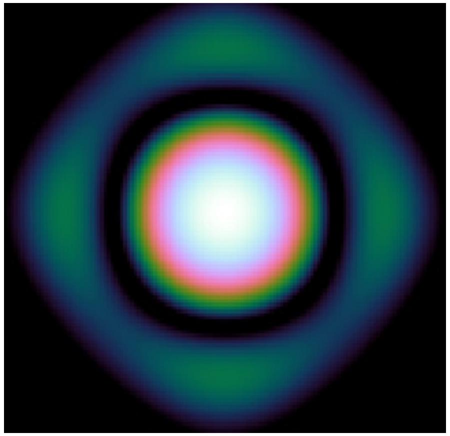
Just the main lobe (20% point)

WF imaging: A-Projection

- WF imaging needs larger convolution functions (CF)



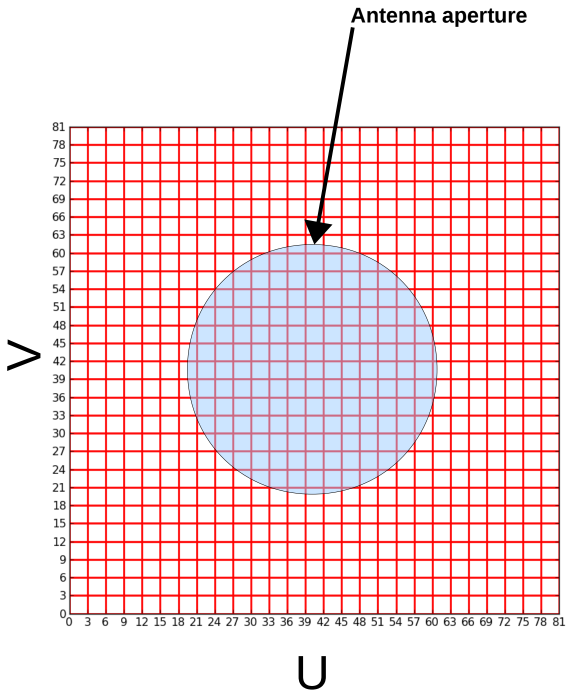
Number of uv-pixel
across antenna aperture



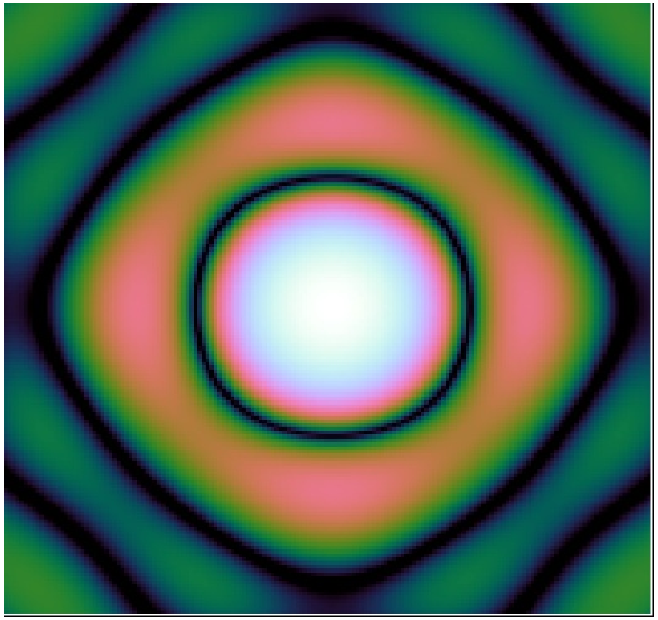
Include the first sidelobe (few%)

WF imaging: A-Projection

- WF imaging needs larger convolution functions (CF)



Number of uv-pixel
across antenna aperture



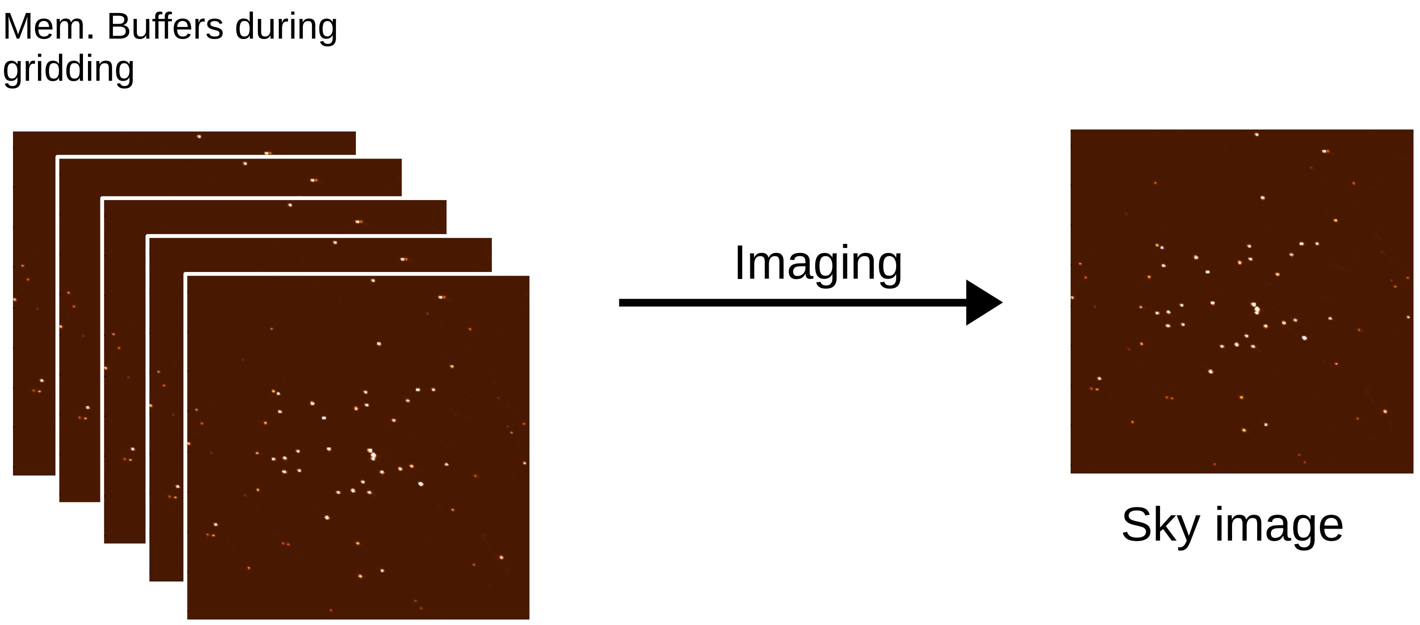
..beyond the first sidelobe

Imaging Memory footprint

- Each sky-image of size N_x x N_y requires
  - 2 x Complex x (N_x x N_y)+ (N_x x N_y) = **5 x (N_x x N_y) floats**

Imaging Memory footprint

- Each sky-image of size $N_x \times N_y$ requires
  - 2 x Complex x $(N_x \times N_y)$+ $(N_x \times N_y)$ = **5 x $(N_x \times N_y)$ floats**

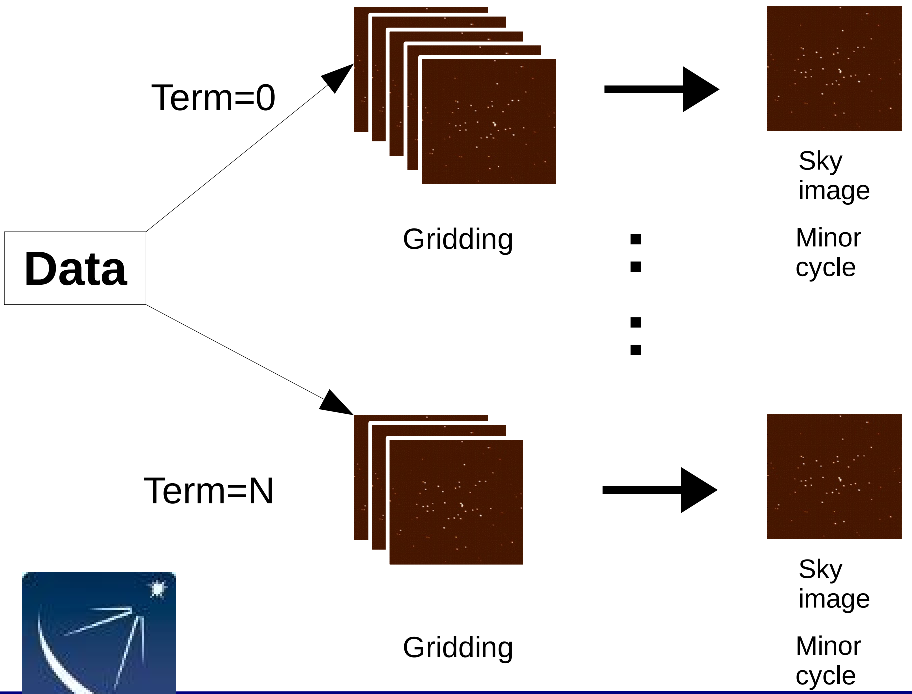


Major cycle

Minor cycle

MT-MFS: Higher memory footprint

- WB A-Projection: N_A x N_{SPW} (order 10x increase in CF memory footprint)
- MS-MFS
  - Compute load: Gridding for N_{terms} images+ Convolution of large images
  - Memory: Multiple minor-cycle images (N_{scales})
  - Total images (each of size N_x x N_y) : N^2_{terms} x N^2_{scales}



Memory storage for: N^2_{terms} x N^2_{scales}

Compute convolutions of images

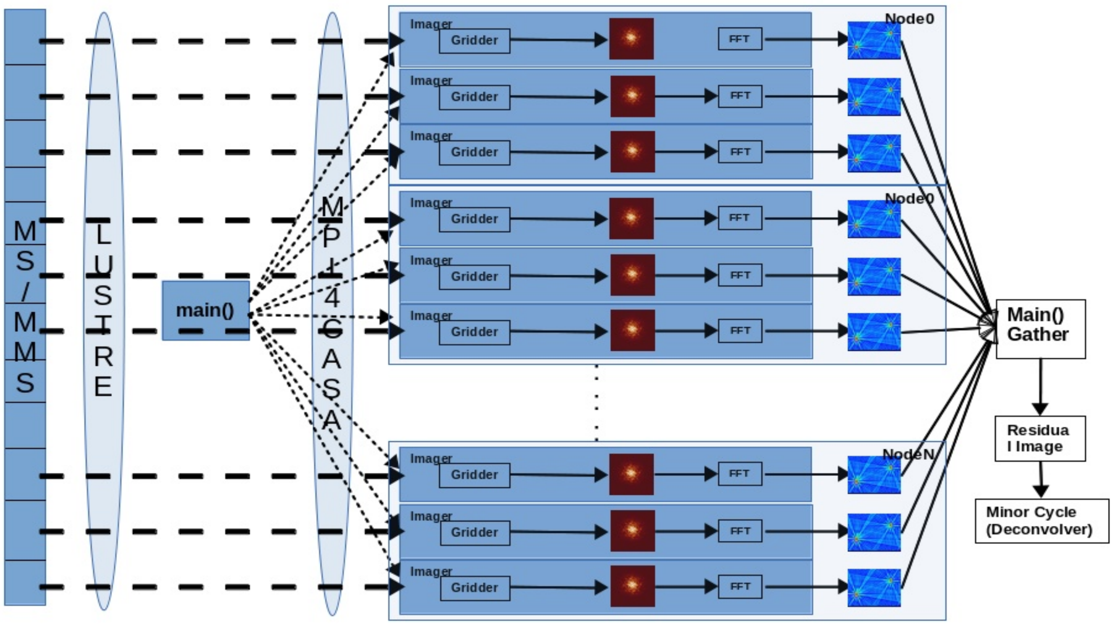
Gridding Parallelization (HS 1) - I

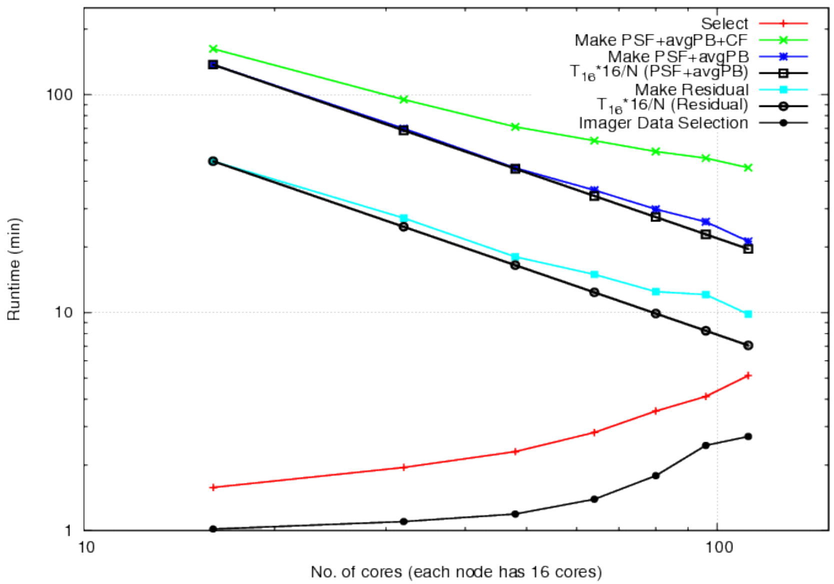
- Compute load: $N_v \times 12\ N_c^{\,2}$
- Scatter along data axis

Memory footprint increases
Linearly with no. of procs.

Too high for A-array imaging

$N_{proc} \times 5 \times (N_x \times N_y)$





Computing models

- Software options
  - Run multiple CASA-instances by-hand
  - Use CASA parallel computing framework via *mpicasa*

- Hardware options:
  - Multi-core (desktop) machines
  - Cluster : For multi-node multi-core computing
    - @NRAO or @Your home institutions
  - Amazon Web Services (AWS):
    - Amazon's cloud computing platform
    - Largest collection of compute resources in the world (!)
    - Use based cost structure
  - XSEDE: Extreme Science and Engineering Disocvery Environment
    - NSF Funded Supercomputing facilities
    - Free to use but only availableto United States PIs

AWS and XSEDE Overview

- What is AWS?
  - Collection of physical resources and tools to enable ad hoc creation of a computer
  - Create single workstations or arbitrarily large clusters:
    - Rich choise of storage (RAM, disk), compute cores, GPUs,...
  - Create shareable from Gbytes to Pbytes
  - Pay as you go model

- What is XSEDE?
  - 16 Major facilities nation wide.
    - But available only to United States Pis
  - Resources granted via quarterly proposal/review process
  - Primarily focused on broad parallel large jobs (1000’s of cores)
  - Typically memory and storage limited compared to NRAO clusters

CASA on AWS or XSEDE

- Installation effectively as “tar -xz <tarball>”
- NRAO supported CASA installation
  - As AWS public machine images
  - As SDSC/Comet (XSEDE)
- Growing documentation repos., shared scripts to follow
  - DMSD HPC Group at NRAO will assist the community
  - Driven by community interest, limited by competing projects

- Most of the labour is in automation (non-interactive use)
- Highest utility for both AWS and XSEDE is in automated batch processing; not efficient for interactive processing