Wide-field Imaging

Feb. 24th, 2012

S. Bhatnagar

Why new algorithms?

- Instantaneous wide-band capability of the EVLA is a the single dominant parameter that enables new scientific capabilities
  - Instantaneous sensitivity improvements by $\sqrt{BW}$
  - Better imaging performance due to improved uv-coverage
  - More instantaneous information about the emission
    - Spectral Index, RM, non-monochromatic polarization,...
  - Hardware that allows new possibilities

- "...my scientific inquiry was limited by instrument capability..." or "...I have this scientific question that needs the EVLA..."
  - Enjoy :-)...

However...

- Know thy needs
  - As always, new algorithms have limits of applicability, limits of new science they enable
  - Have associated costs: higher computing, higher i/o
    - Translates to longer run-time, grief, length of PhD time-scale,...
- Technical solution
  - Have as wide a range of tools available in a flexible framework
  - In using software packages slightly beyond the "tasking level only", it is possible to more creatively combine software tools/techniques to enable the capabilities you need and keep the complexity in control
    - MakeMyTask in casapy can be very useful
      - At personal as well as at a community-contributions level
- Moral: Don't use things blindly (no silver bullet)
  - Know what you need

Plan

- What do we mean by wide-field?
- Projection algorithms to correct for various wide-field effects
    - Relation with minor cycle algorithms
- Algorithms “unification scheme” :-)
    - Similarity between various wide-field algorithms
- Algorithms
    - For W-term correction
        - W-Projection, Multi facet Imaging
    - For PB corrections
        - A-Projection: Low and high frequency
    - AW-Projection at low frequency bands
- Connection with Mosaicking:
    - Generalization of single pointing

What do we call Wide-field?

- Imaging that requires invoking any of the following:
  - Corrections for non co-planar baseline effects
    - Errors due to planar geometry assumptions > Thermal noise (L/S/C-bands)
  - Corrections for the rotational asymmetry of the PB
    - Imaging beyond 50% point, mosaicking
  - Corrections for the frequency or polarization dependent effects
    - PB, ionosphere/atmosphere

- Noise limited imaging at "low" bands (L, S and probably C Band)
  - Because of the radio brightness distribution
- Noise limited imaging of structure comparable to the PB beam-width

$$I_{Continuum} = \int PB(\nu)\left[I_o(\nu/\nu_o)^{\alpha(\nu)}\right] d\nu\, dt = \int I_o(\nu/\nu_o)^{\alpha_{pb}(\nu,t)+\alpha(\nu)} d\nu\, dt$$

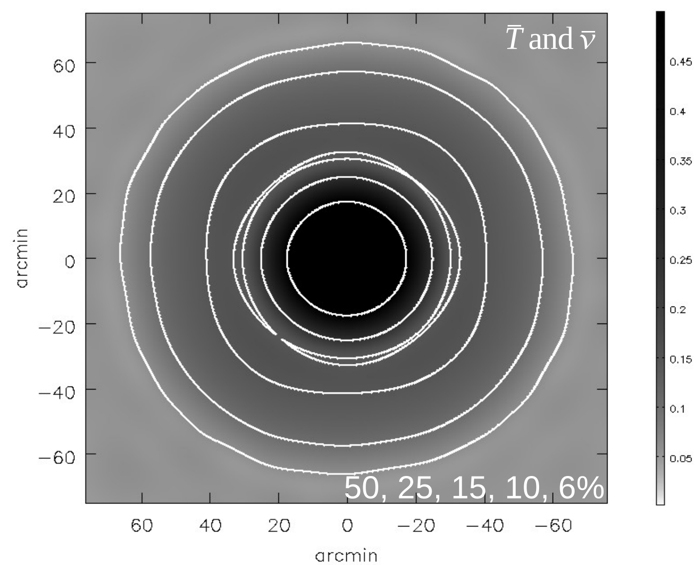
- Mosaicking
  - By definition, imaging on scales larger than the PB beam-width

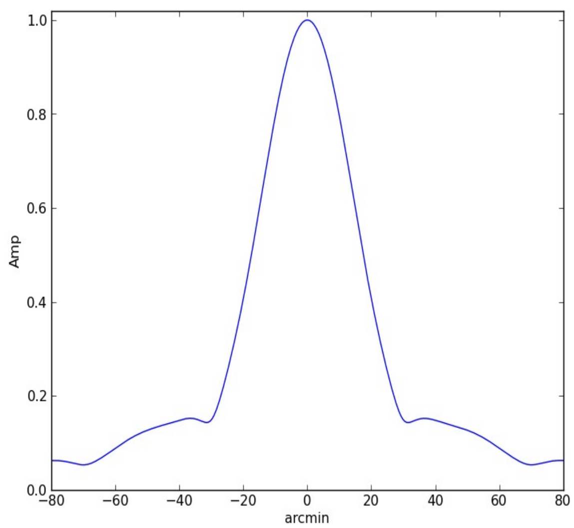
Why wide-field?

- Primarily due to improved continuum sensitivity
- 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
- Due to large bandwidth, EVLA is sensitive father out in the FoV
- E.g. @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 VLA-null!
  - No null in the EVLA sensitivity pattern

Wide-field Issues

- For the same integration time, EVLA is sensitive to emission farther out





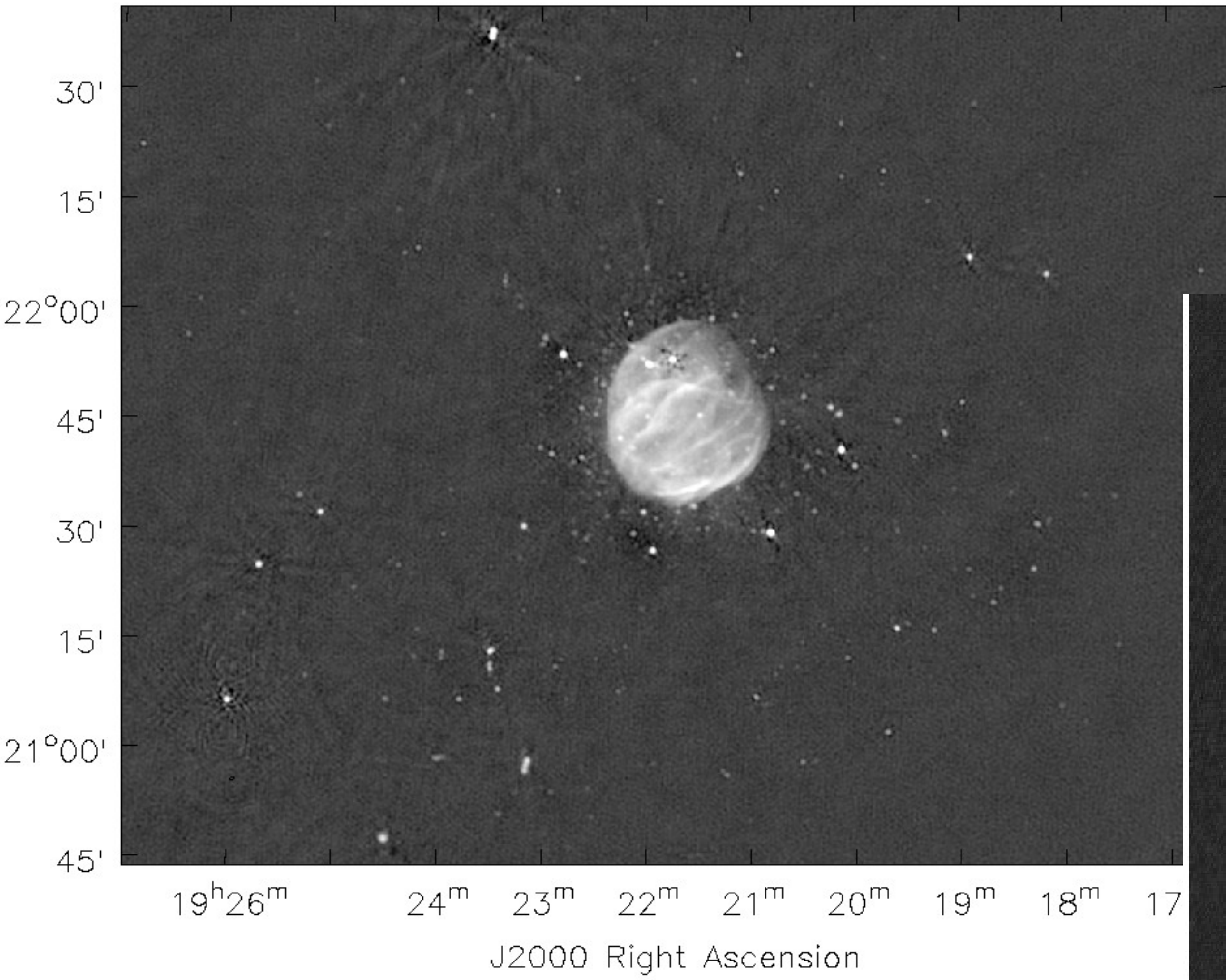
[Bhatnagar, Rau, Green & Rupen, 2011, ApJL, 739, L20]

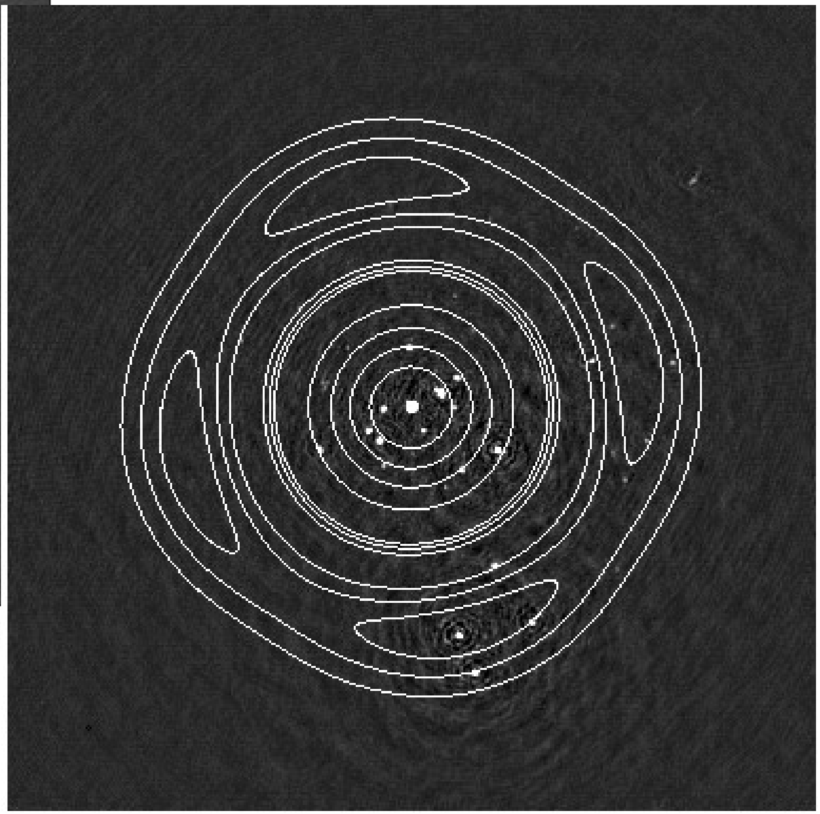
- Error at the center of the image due to a source at a distance R

$$\Delta S = S(R) \times PB(R) \times PSF(R)$$

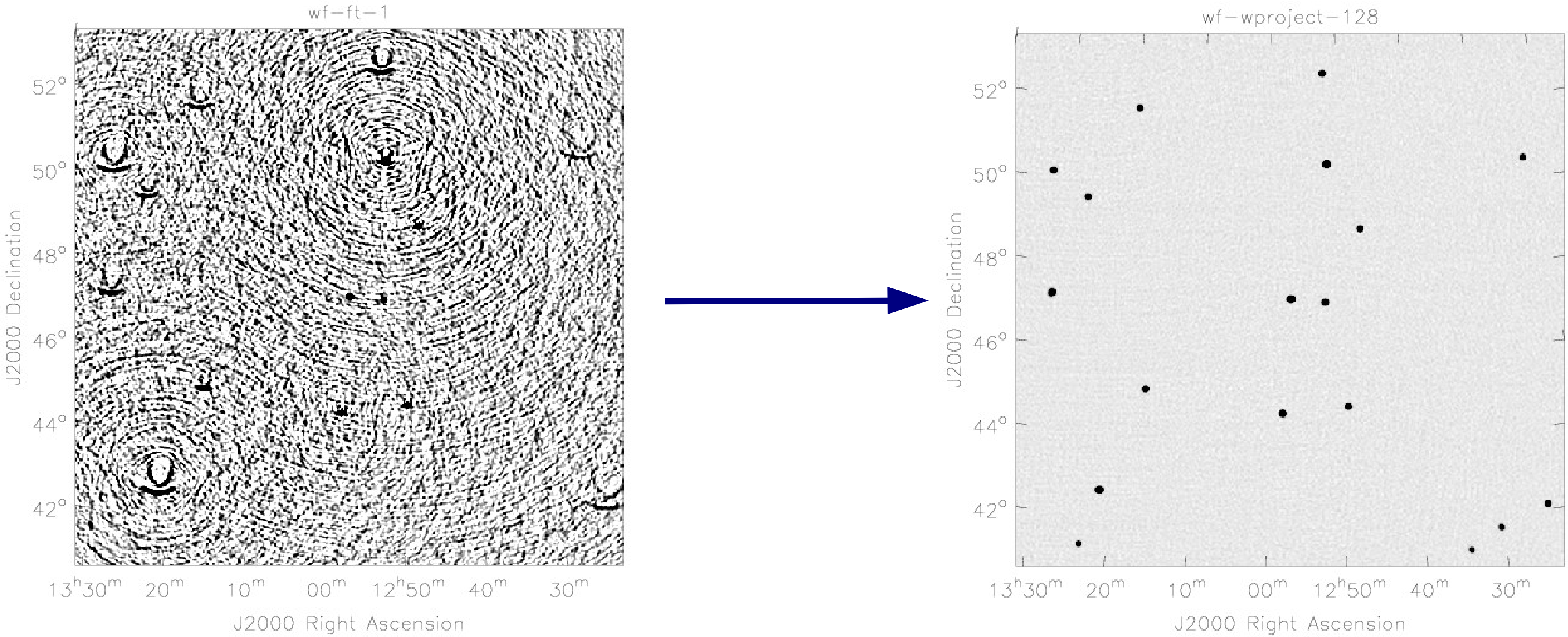
  - R = 1°, S(R)=1Jy, $\Delta S = 1 mJy - 100 \mu Jy$

Wide-field Issues



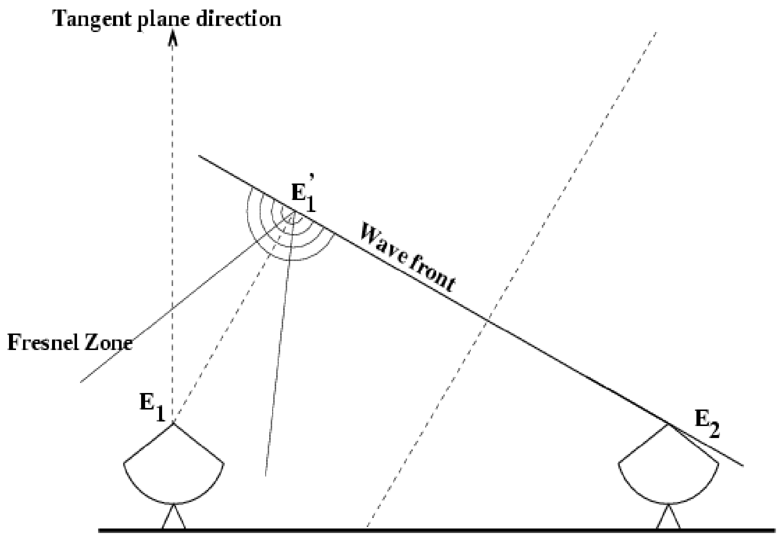


Effects of the W-Term



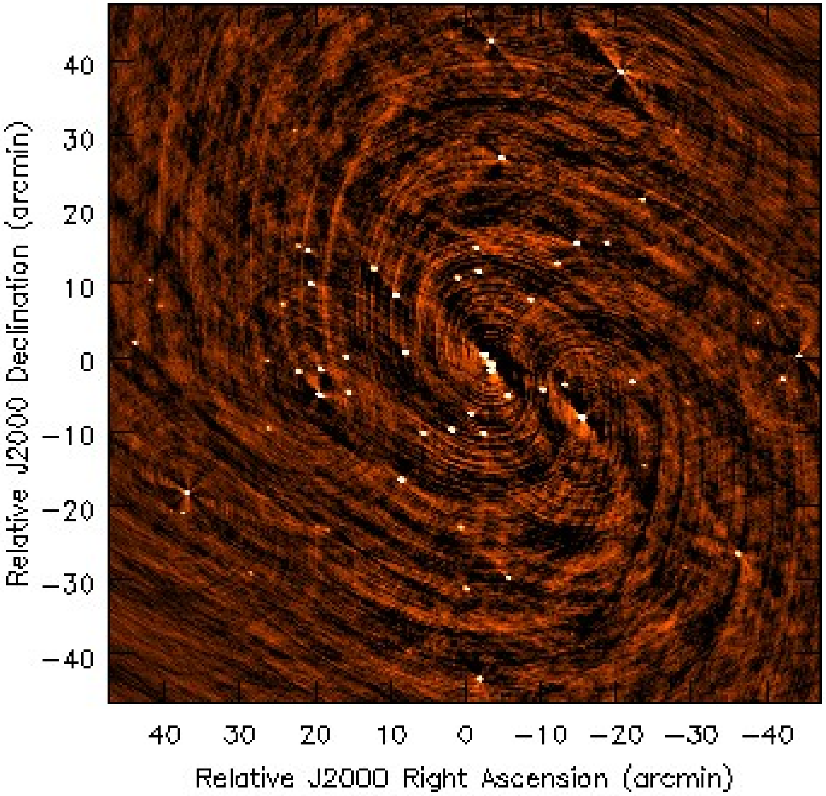
Non co-planar baseline: The W-term

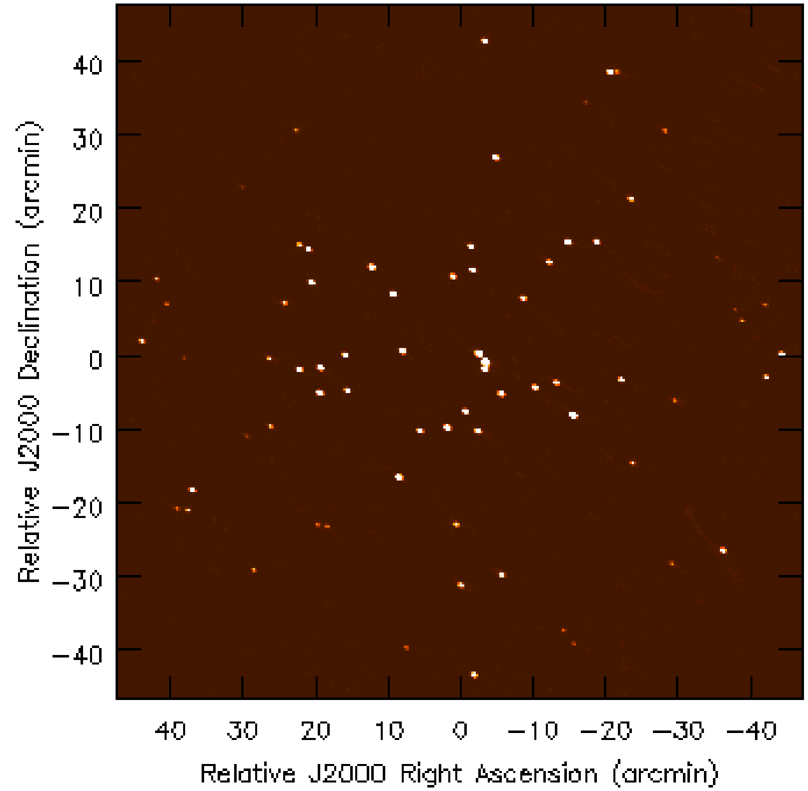
- 2D FT approximation of the Measurement Equation breaks down
  - $\frac{\lambda}{B_{max}} \leq \theta_f^2$ $\theta_f = Angular\ distance\ from\ the\ phase\ center$



- 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

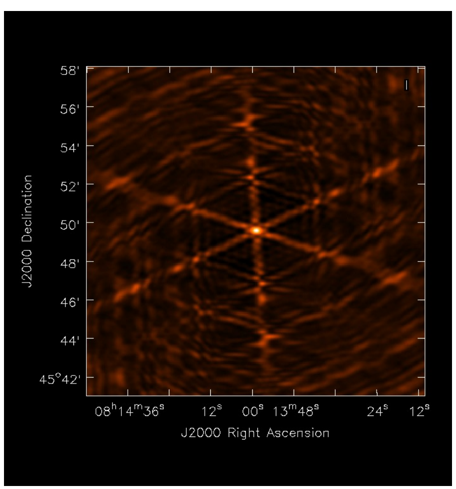

Relative J2000 Declination (arcmin)
40
30
20
10
0
−10
−20
−30
−40
40 30 20 10 0 −10 −20 −30 −40
Relative J2000 Right Ascension (arcmin)

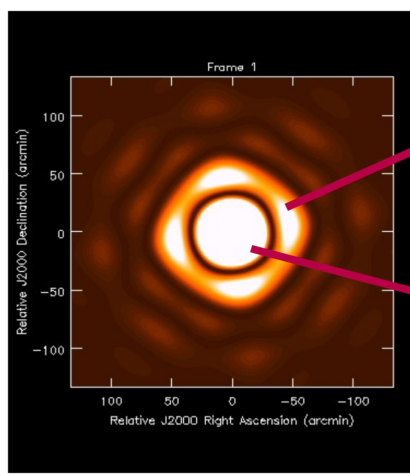

Relative J2000 Declination (arcmin)
40
30
20
10
0
−10
−20
−30
−40
40 30 20 10 0 −10 −20 −30 −40
Relative J2000 Right Ascension (arcmin)

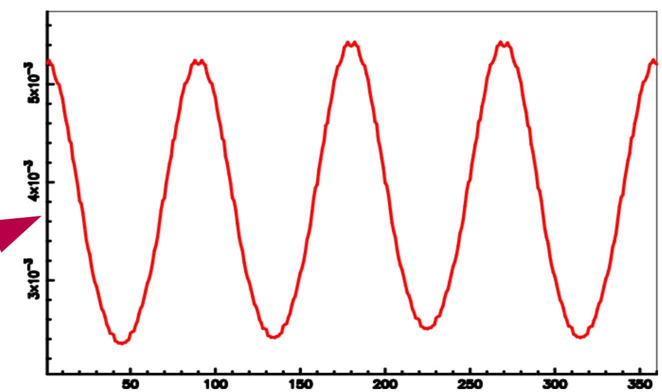
PB Effects: Rotation asymmetry

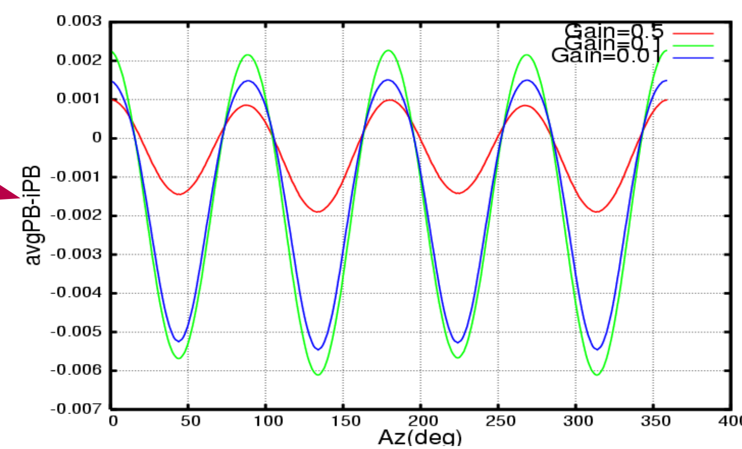
- Only average quantities available in the image domain
- Asymmetric PB rotation leads to time and direction dependent gains

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





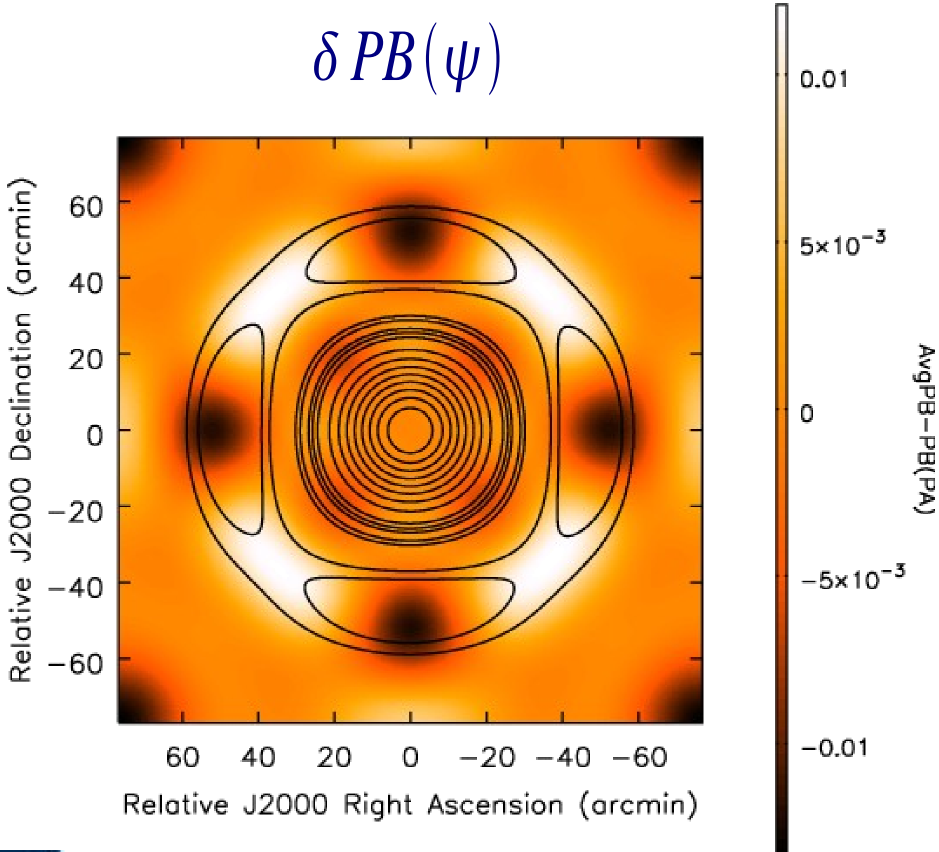




PB Effects: Error Propagation

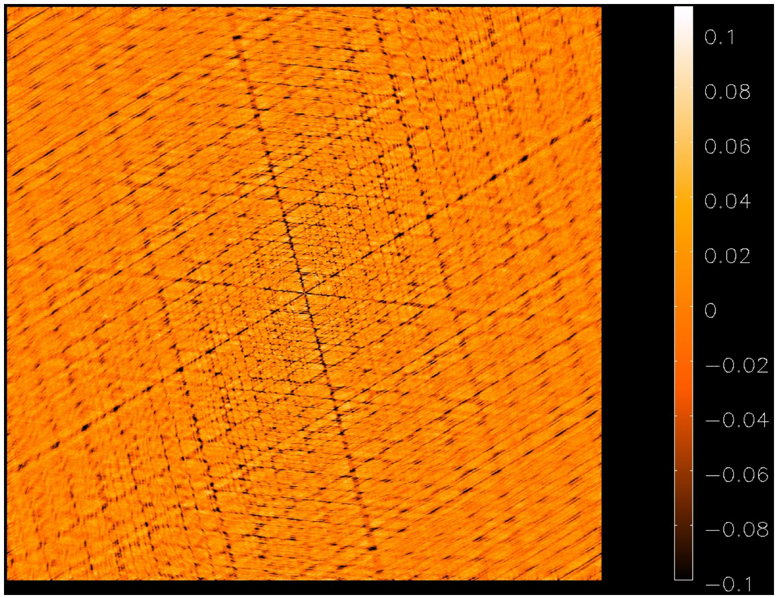
$$\Delta I^R = \sum_{\psi} \delta PSF(\psi) * [\delta PB(\psi) I^o]$$

$\delta PB(\psi)$



E.g. 5×10^{-3}

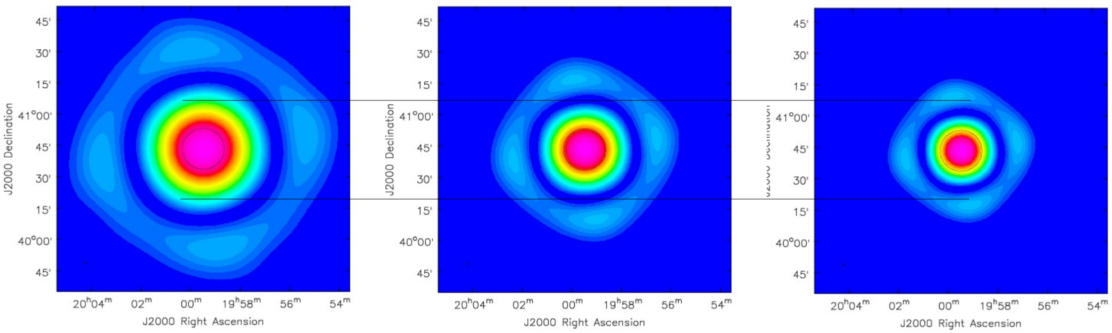
$\delta PSF(\psi)$

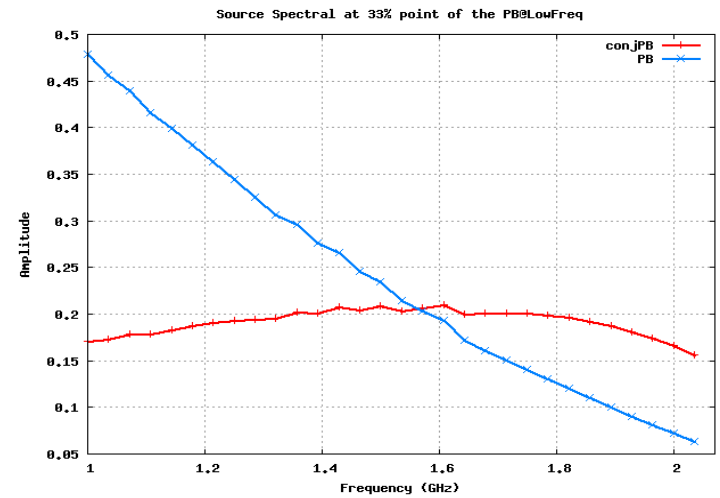


E.g. 2×10^{-2}

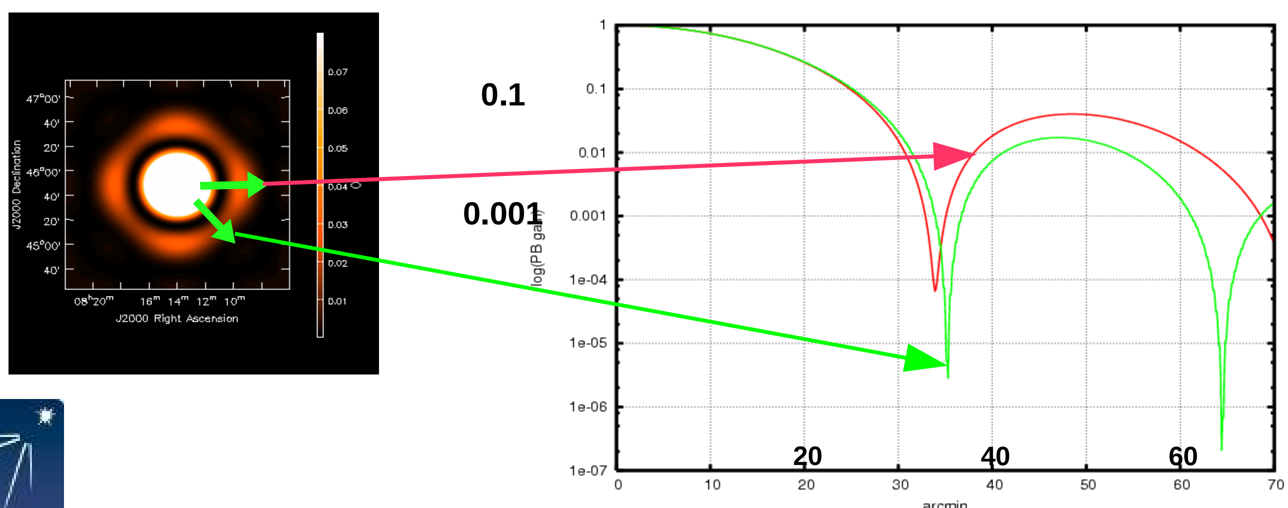
Frequency dependence of the PB

- Assume linear scaling with the frequency





- Frequency- and direction-dependent gains

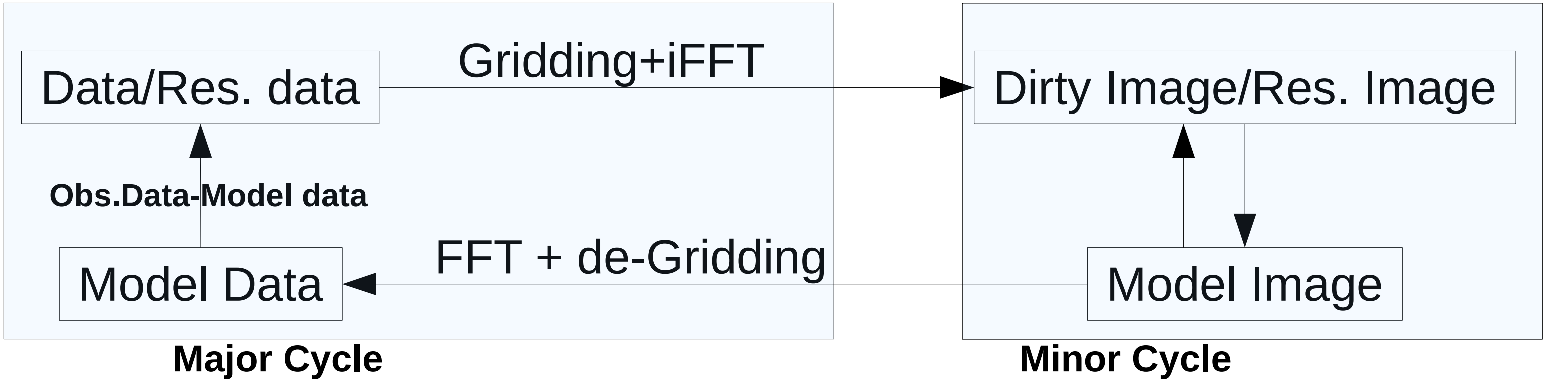


Effect antenna pointing errors

- Typical EVLA antenna pointing errors 10-15 arcsec (w/o reference pointing)
- Can limit imaging at low bands because of stronger, more complex sky emission at low frequencies (LSC-bands)
  - Flux at half-power point maximally contributes to increased noise floor
- Does limit imaging at high frequencies due to smaller PB
  - Solution: Reference pointing every ~30 min. (due to time dependence)
- It's a time and direction dependent effects, but antenna based
  - New possibilities: Pointing SelfCal (EVLA Memo #84, 2004)
    - Possible due wide-band capability of the EVLA
  - Why? Better time utilization, allows pointing selfcal and related improvements

Algorithms: CS Clean recap

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



- Most commonly used algorithm
- Every major cycle access the entire data base
  - Significant increase in I/O and computing load
- Assumes, co-planar, time- and freq-independent Measurement Equation
- Cannot account for wide-field wide-band and time variability issues

Deconvolution as ChiSq Minimization

- $V^M = A I^M + A N$ $\qquad V_{ij} = deGrid_{ij} FT(I)$
- Non-linear solver, to solve for the Model Image
  - Compute residuals: $\quad$ $\mathbf{V^{Obs} - AI^M}$ (data domain)

    $\mathbf{I^d - BI^M}$ (image domain)
  - Make Residual Image $\mathbf{I^{res}}$
  - Find update direction: Steepest Descent Algorithm

    $$I^c = max\left(-2[I^{Res}]\frac{\partial \chi^2}{\partial\, Param}\right)$$
  - Update model: $I_i^M = T(I_{I-1}^M)$ $\qquad$ $for\ our\ discussions\ this\ is = I_{i-1}^M + \alpha * I_i^c$
  - Since Major Cycle does model subtraction without averaging, variable terms can be included in that step

Major Cycle
(always expensive)

Minor Cycle
(can be expensive)

Algorithms "unification scheme"

- Incorporates direction dependent effects as part of the gridding function
  - ME: $V_{ij} = A_{ij} I^o + N_{ij}$
  - Construct D, such that $\frac{D_{ij}^T A_{ij}}{D_{ij}^T D_{ij}} \approx 1$
  - Compute residuals (major cycle): D_{ij} for forward and D_{ij}^T for reverse transform

- W- and A-Projection construct ***D*** differently
  - A-Projection has additional normalization issues:
    - Flat-noise vs. flat-sky normalization
- Mosaicking: (more in K. Golap's talk in Thursday Lecture Series)

[https://safe.nrao.edu/wiki/pub/Software/Algorithms/WebHome/Mosaicking_aoc.pdf]

$$I^{Mosaic} = \sum_k I(l_o - l_k)$$

Use $D_{ij} e^{\iota[(l_o - l_k).u_{ij}]}$ where D_{ij} can be A_{ij}, W, or $A_{ij} * W$

- The Fourier transform shift theorem

Projection algorithms

- Direction-dependent ("image plane") effects as convolutional terms in the visibility domain
- ME entirely in the visibility domain: $\boldsymbol{V^{O}_{ij} = A_{ij} I^{M} = M_{ij} F I^{M} = M_{ij} \left[V^{M} \right]}$

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

- Diagonal: "pure" poln. products
- Off-diagonal: Include poln. leakage

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

- $\boldsymbol{V^{O}_{pp}} = M_{pp} * V^{M}_{pp} + M_{p\ \ p2q} * V^{M}_{pq} + M_{q\ \ p2q} * V^{M}_{qp} + M_{p2q\ \ p2q} * V^{M}_{qq}$
- Generalization of the direction-independent ME
  - Replace functions by complex numbers $\quad M_{ij} = g_i g^{*}_{j}$
  - Replace convolution ('*') by complex product

Algorithms "unification scheme"

- "Single polarization" case: Single element of the Mueller Matrix

- Imaging

$$\boldsymbol{V}^{\boldsymbol{Grid}} = \boldsymbol{CF} * \boldsymbol{V}^{\boldsymbol{obs}}$$

$$\boldsymbol{I}' = FFT[V^{Grid}]$$

- Prediction (de-gridding):

$$\boldsymbol{V}^{\boldsymbol{Grid}} = FFT^{-1}[I^{M'}]$$

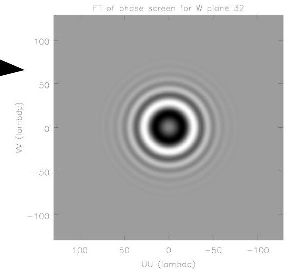
$$\boldsymbol{V}^{\boldsymbol{M}} = CF^{T} * V^{Grid}$$

- CF can be A-term, W-term, AW-term, wide- or narrow-band

W-Projection

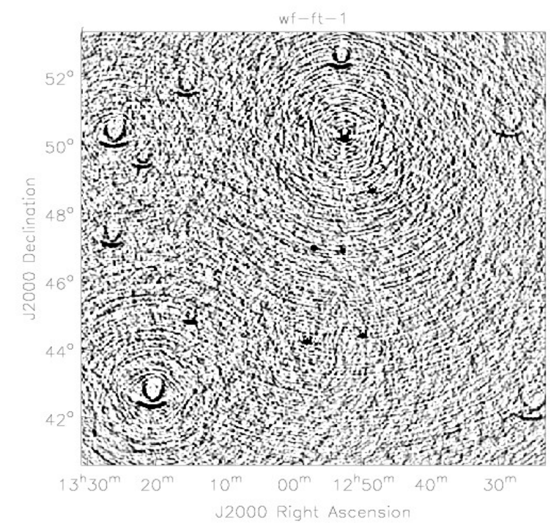
- W-Projection: (CASA Imager: ftmacine=”wproject”)

$$D = FT\left[e^{2\pi\iota\sqrt{w-1}}\right]$$



- Potentially fully corrects for the effects of the W-term
  - In practice, D is computed at a finite w-resolution, with interpolation in between
  - gridmode='widefield'; wprojplanes=N; facets=M

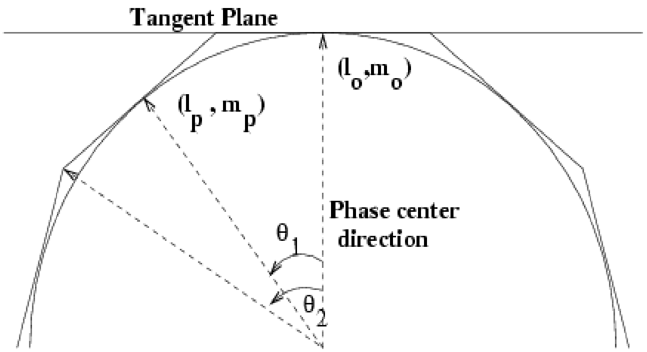
- D is non-hermitian
  - Post deconvolution correction is not possible
  - Same as: “corrections for antenna based phase errors cannot be corrected for post-deconvolution”



[Cornwell, Golap & Bhatnagar, 2008]

W-Projection + Multi-faceting

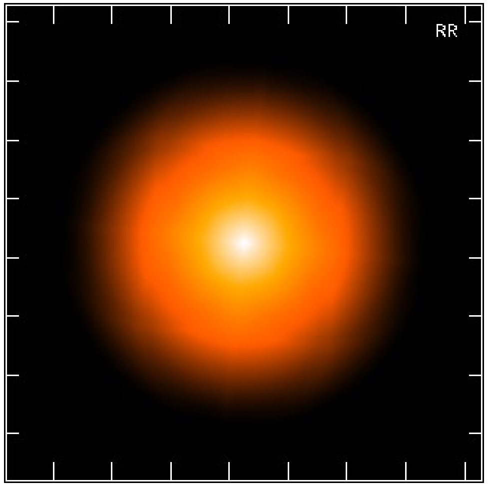
- Multi-facet imaging (CASA Imager: facets > 1)

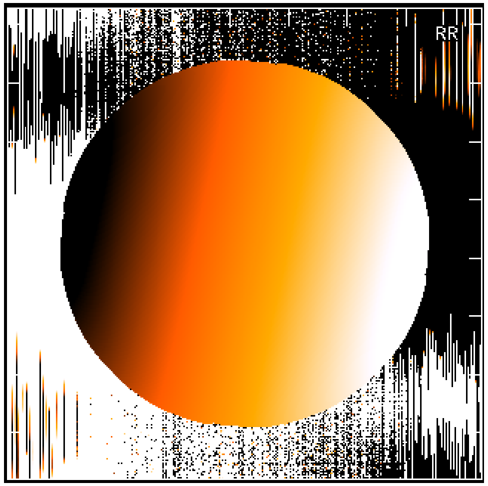


- Split the sky into multiple, smaller tangent-plane images
- A linear approximation of this image-plane operation is possible in the visibility plane: $\boldsymbol{I}(\boldsymbol{C}\,l) \rightarrow |det(\boldsymbol{C})|^{-1}\boldsymbol{V}(\boldsymbol{C}^{-1^T}u)$
  - Advantage: leads to a single combined image in the minor cycle
- Combination of W-Projection and Multi-facet imaging possible:
  - Reduces the no. of w-planes and number of facets

A-Projection

- A-Projection: $\boldsymbol{D}$= Auto-correlation of Aperture illumination function
  - Function of time, frequency and polarization
- Since image is averaged over time and frequency, time- and frequency-dependence cannot be corrected post-deconvolution
  - Same issue as non Hermitian nature of antenna based phase, W-term

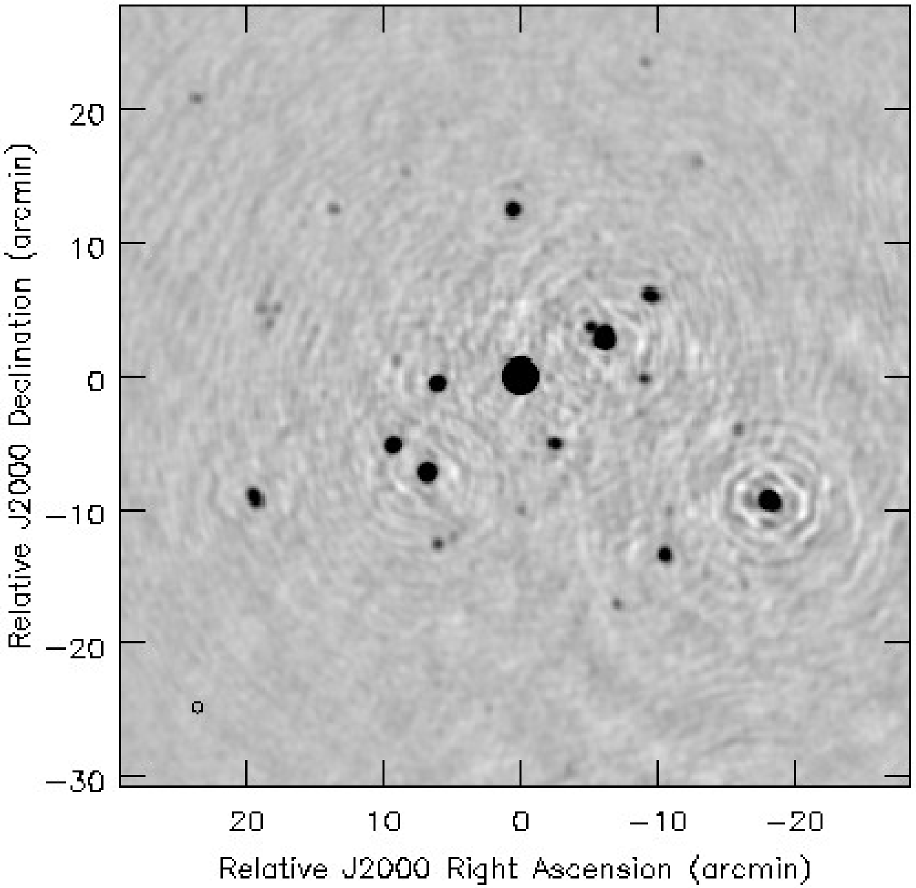




- Effective PB is time- and polarization-independent

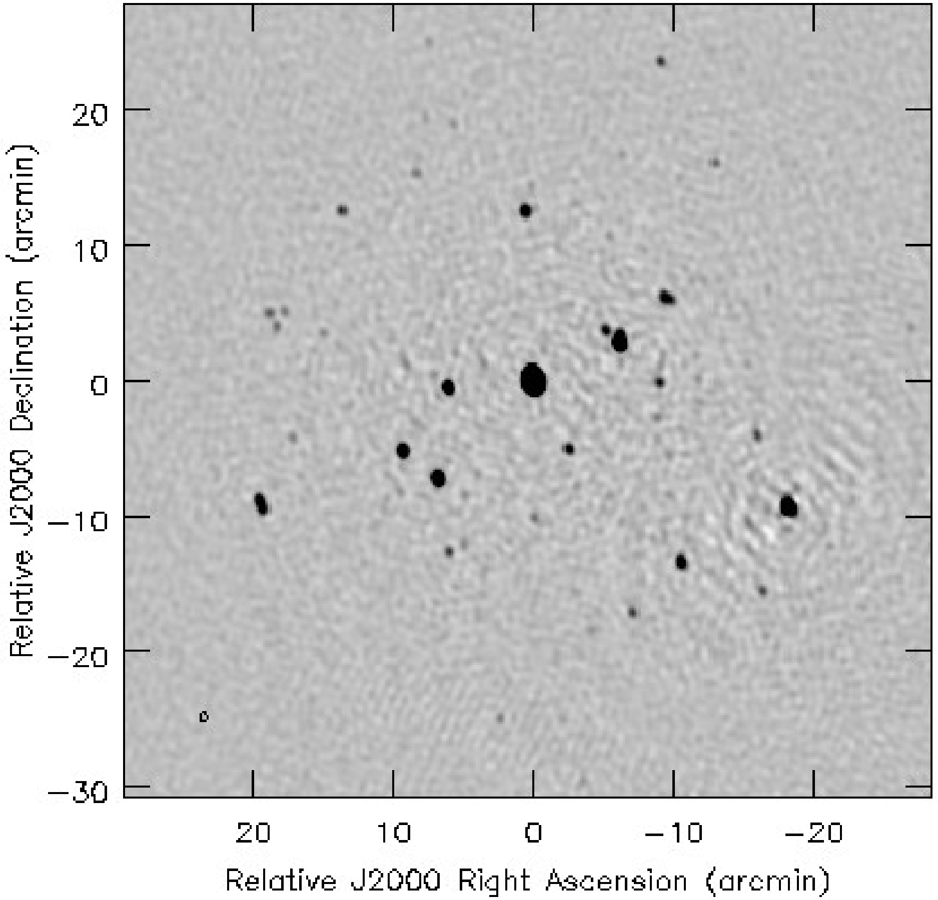
[Bhatnagar, Cornwell, Golap & Uson, 2008]

A-Projection: Stokes-I Before



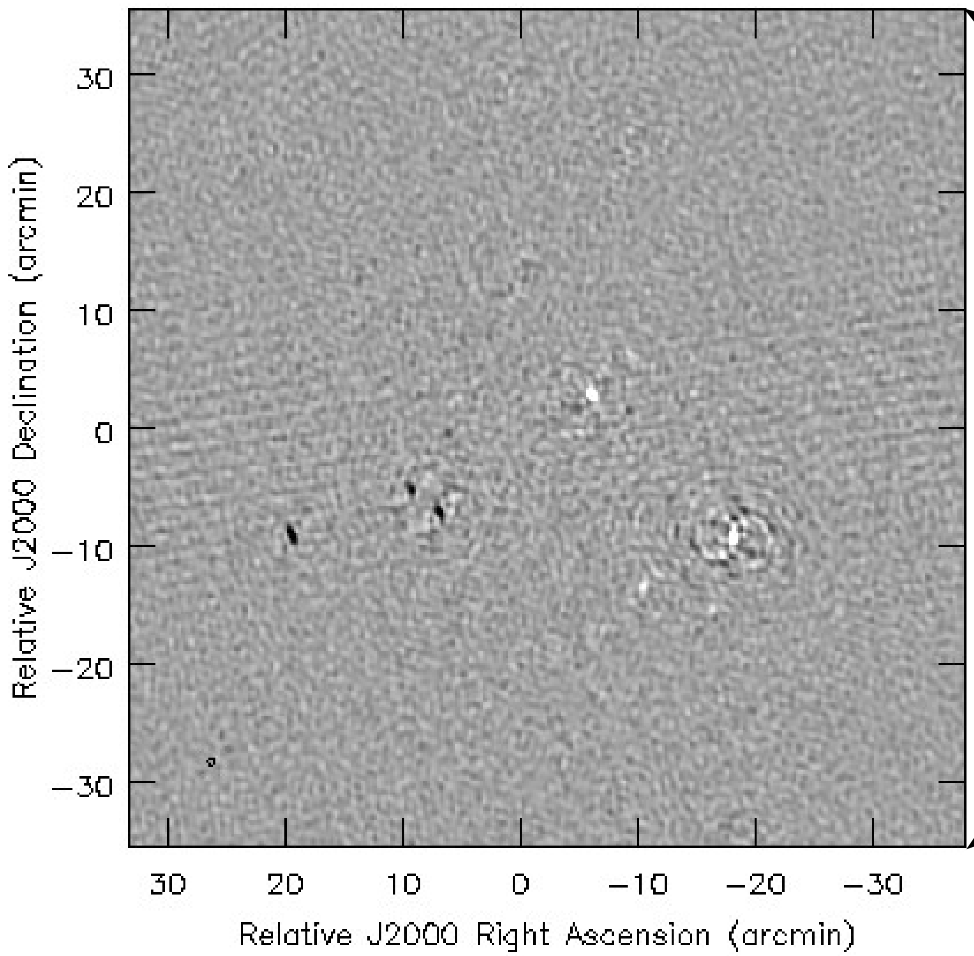
Effective PB is time-variant

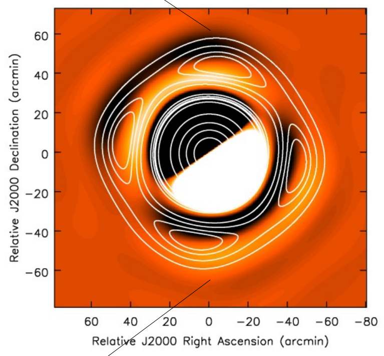
A-Projection: Stokes-I After



Effective PB is time-invariant

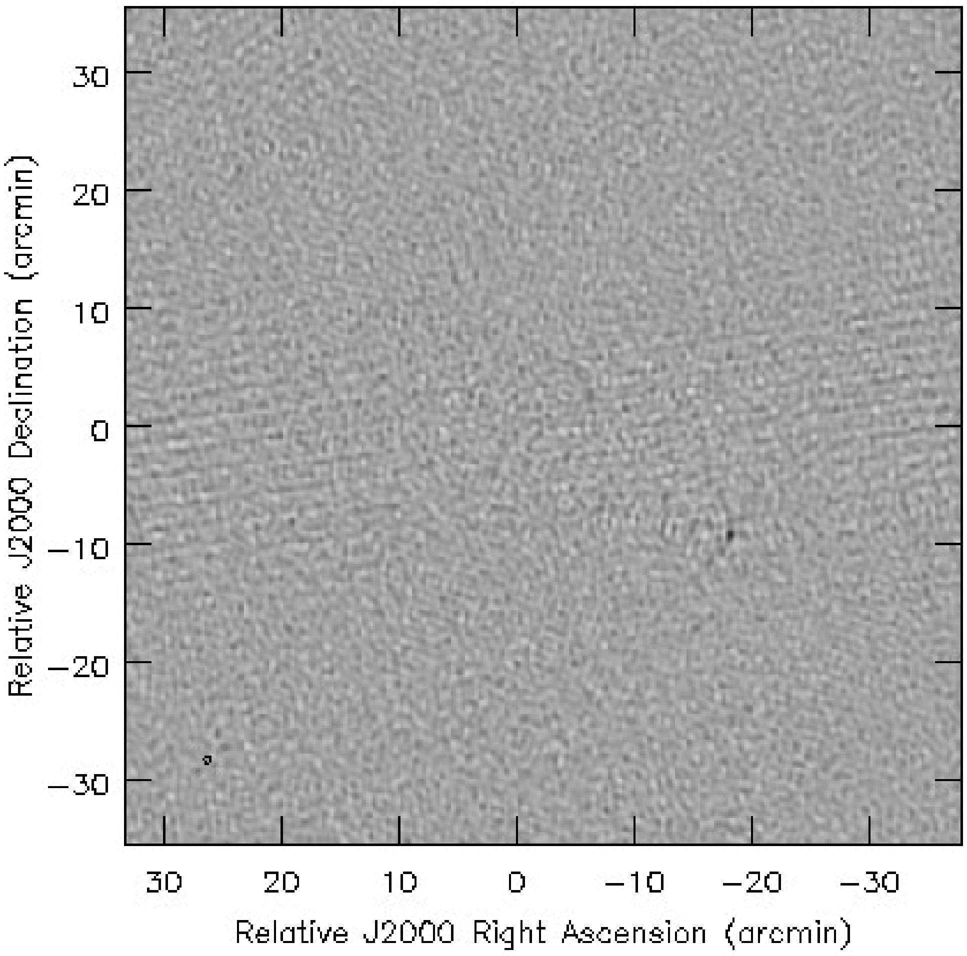
A-Projection: Stokes-V Before





Effective PB is polarization-variant

A-Projection: Stokes-V After



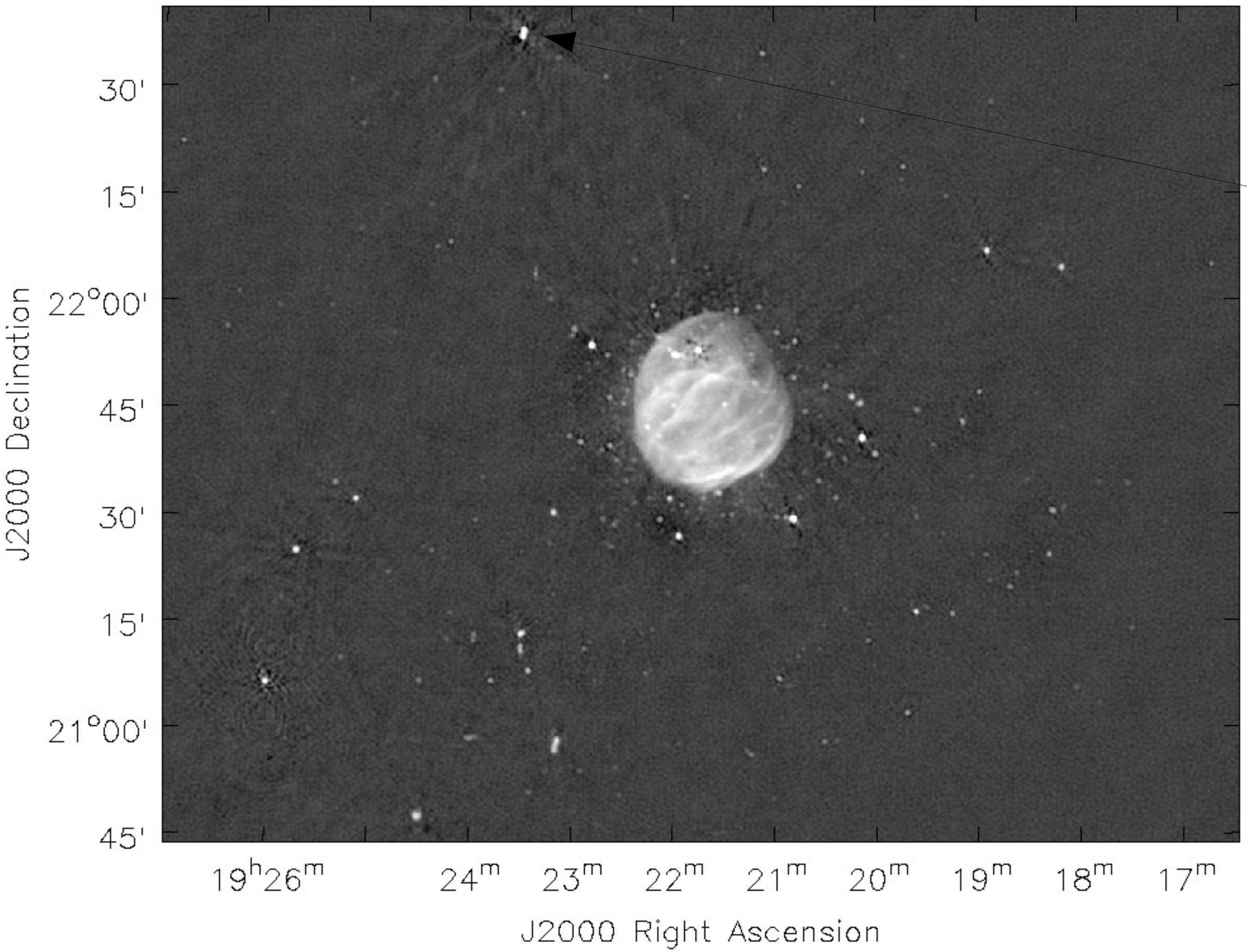
Effective PB is polarization-invariant

Imaging at high frequencies

- Definition: Frequency at which the array is co-planar for the required FoV
- To the first order, aperture illumination may linearly scale with frequency (or at least with in a certain range in frequency)
- Wavelength much smaller than the physical reflecting structures
  - Geometrical ray-tracing models might be sufficient
- Can be computed once per SPW, rotated in time, and scale in frequency during imaging
  - Significantly reduces memory foot print, at the cost of computing
  - Can be computed efficiently on GPUs

Imaging at low frequencies

- Definition: Frequency at which the array is non co-planar for the required FoV

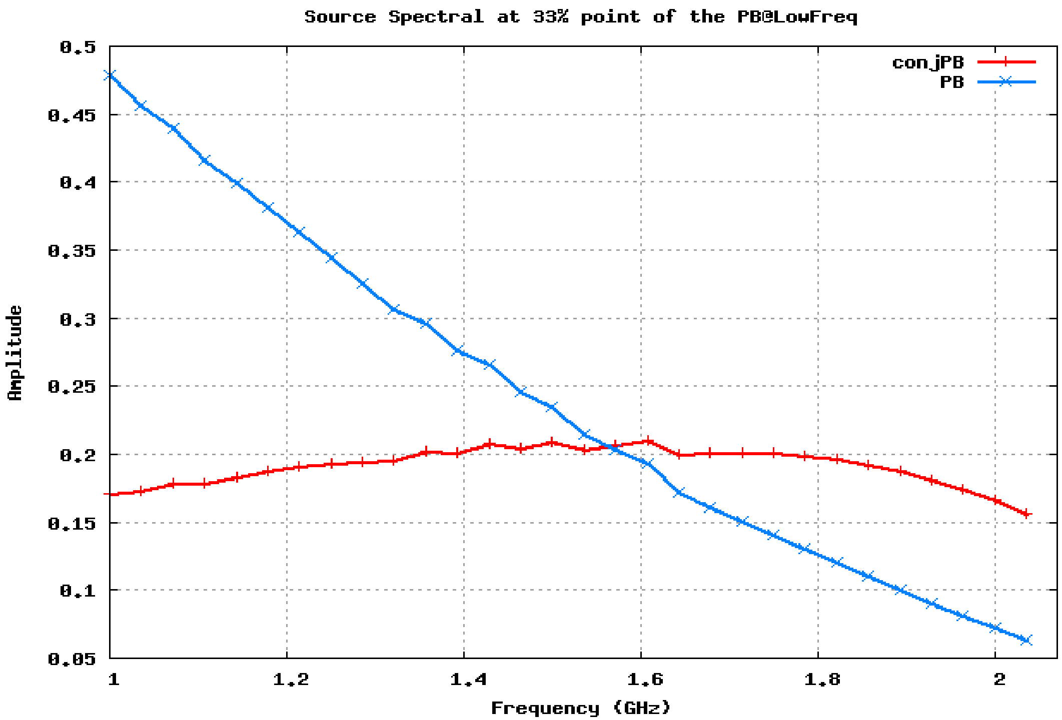


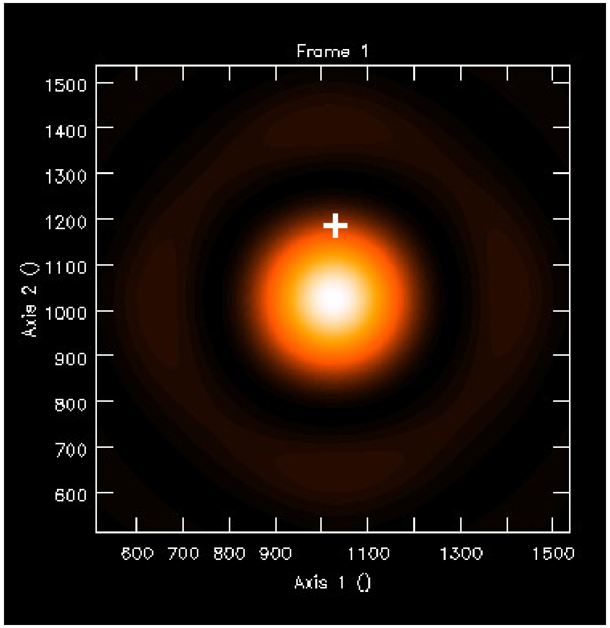
- PB variations with time
- Even D-array is non co-planar
- BW ~400 MHz

- Need:
  Wide-band AW-Projection

Wide-band A-Projection

- PB gains vary with
  - Time: Rotation with PA
  - Frequency: Mostly optics (D/λ)
  - Direction: Rotationally asymmetric PB, non-uniform frequency dependence

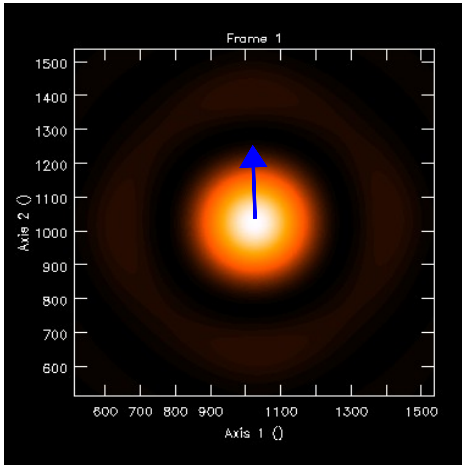




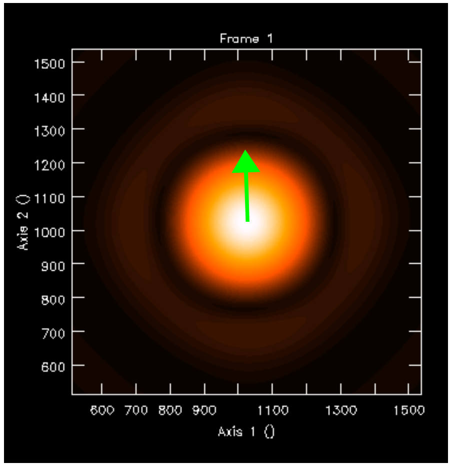
Frequency along the animation axis (1-2 GHz)

Wide-band A-Projection

- The effective PB is frequency independent
  - Rotationally asymmetric PB, non-uniform frequency-dependence

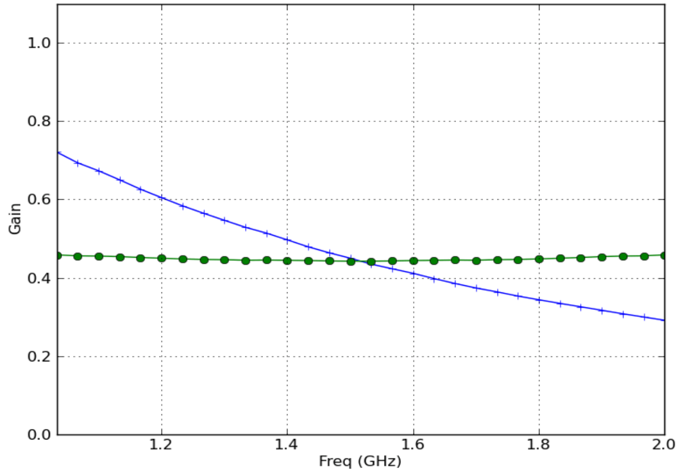


Classical imaging



WB A-Projection imaging

Frequency along the animation axis (1-2 GHz)

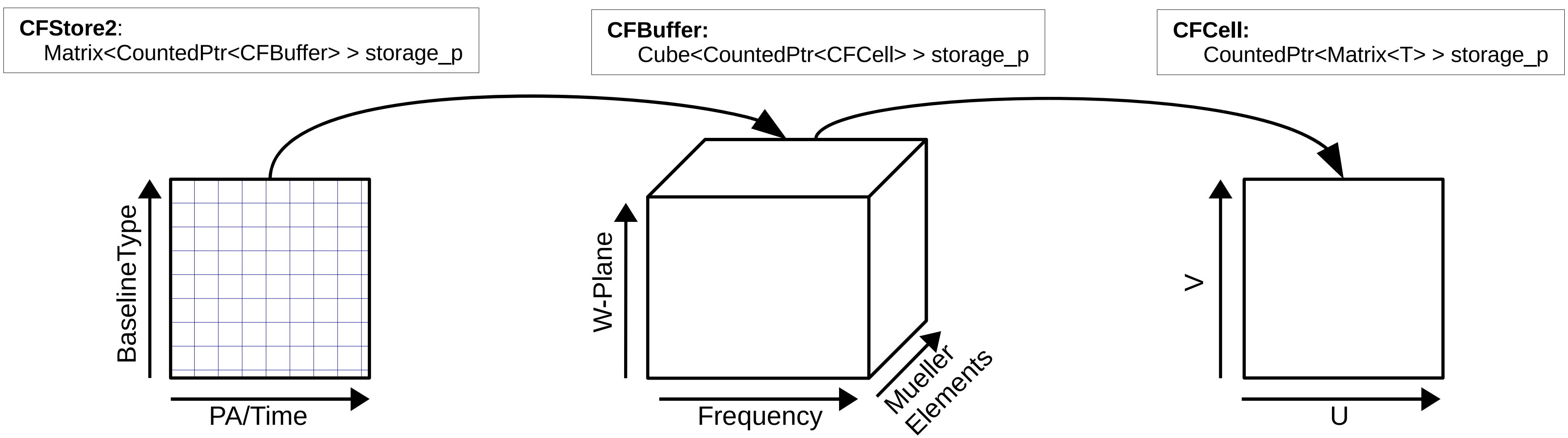


Blue curve: Classical imaging
Green curve: WB A-Proj. imaging

From the center of the beam, moving outwards along the animation axis.

Wide-band AW-Projection

- $\boldsymbol{D}(\nu) \neq (A * W)(\nu/\nu_o)$
- Full-polarization case requires:

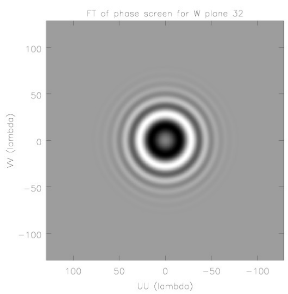


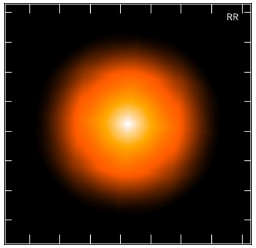
- Can be configured for optimal usage for:
  - High frequency: A-Projection, scaling with frequency
  - Low frequency: AW-Projection
  - Heterogeneous array

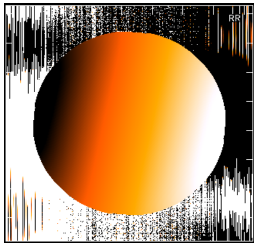
Physics of “unification”

- Physics of DD terms go into the construction of D
- Multiple DD terms become “convolution of convolution functions”

W <convolved> A







- E.g. form of the phase of A-term accounts for mosaicking, pointing corrections, etc.
- Wide-band, full-pol., low-freq. Mosaic can be done naturally
  - Complexity goes in the construction of the CFs
  - Rest of the imaging / calibration framework remains oblivious