Imaging in Radio Astronomy in the presence of Direction Dependent Effects

BASP Frontiers Workshop, Sept. 6th 2011

S. Bhatnagar

National Radio Astronomy Observatory
USA

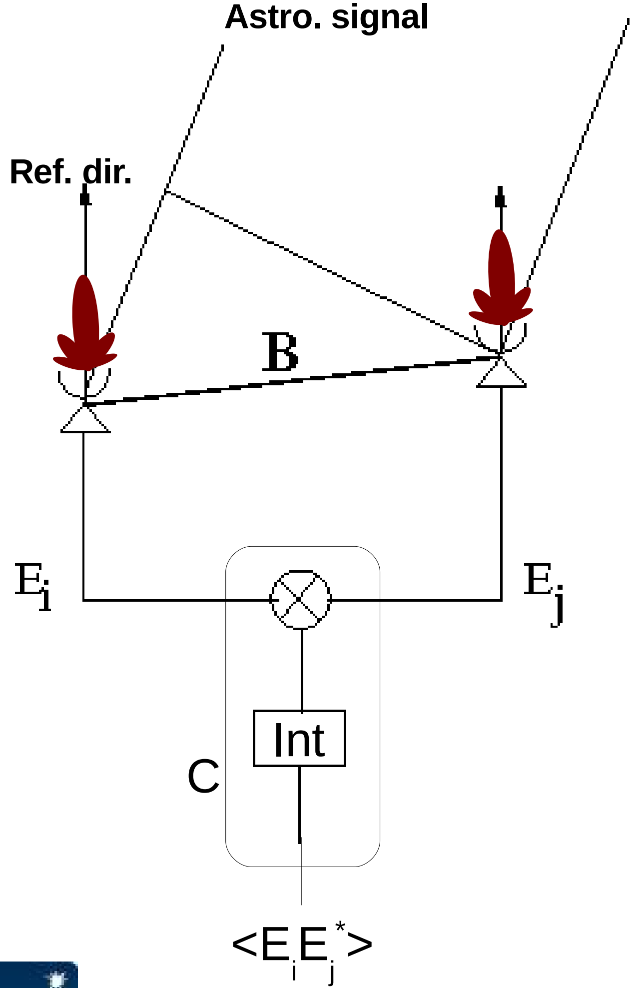
Motivation

- Astronomical studies require high resolution, high sensitivity imaging devices
- In radio bands, wavelengths of interest range from meter to millimeter
- Largest collecting elements (typically antennas) that are practical to build range from 10m – 100m in diameter.

- Problem:
  - Resolution of single elements is often too poor (arcmin)
  - Single elements are not imaging devices (mostly)
- Imaging at sub-arcsec resolution require imaging devices with apertures of 10 – 1000s Km.

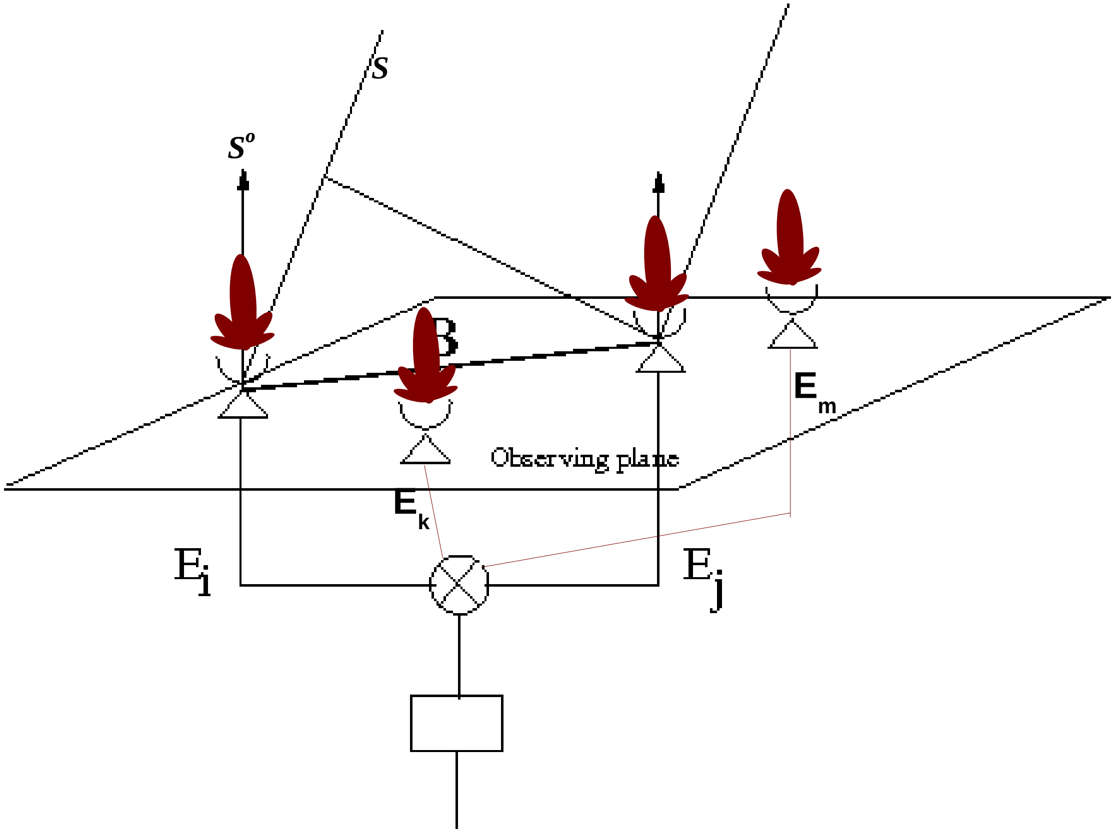
- Solution: Aperture Synthesis/Interferometric Telescopes

Basic set-up: Two element interferometer



- A pair of steerable antennas, separated on the ground
- Signals from each antenna (E_i) are multiplied
- The complex product is averaged in time and frequency (E_{ij}) and recorded for offline processing
- Difference between the time of arrival of the plane wavefront at the two elements w.r.t. to the reference direction is proportional to the projected separation (B) between the antennas
- Terminology:
  - B: Baseline vector
  - C: The Correlator (typically a dedicated HPC digital Machine)
  - $<E_i E_j^*>$: The Visibility from baseline i-j
  - Ref. Dir.: “Phase Center” – delays between the signals from this direction are electronically compensated.

Basic set-up: Aperture Synthesis



$\langle E_iE_j\rangle, \langle E_iE_k\rangle, \langle E_iE_m\rangle, \langle E_jE_m\rangle, \ldots$

$$V_{ij} = A_{ij} e^{\iota B_{ij} \cdot (\sigma - \sigma_o)}$$

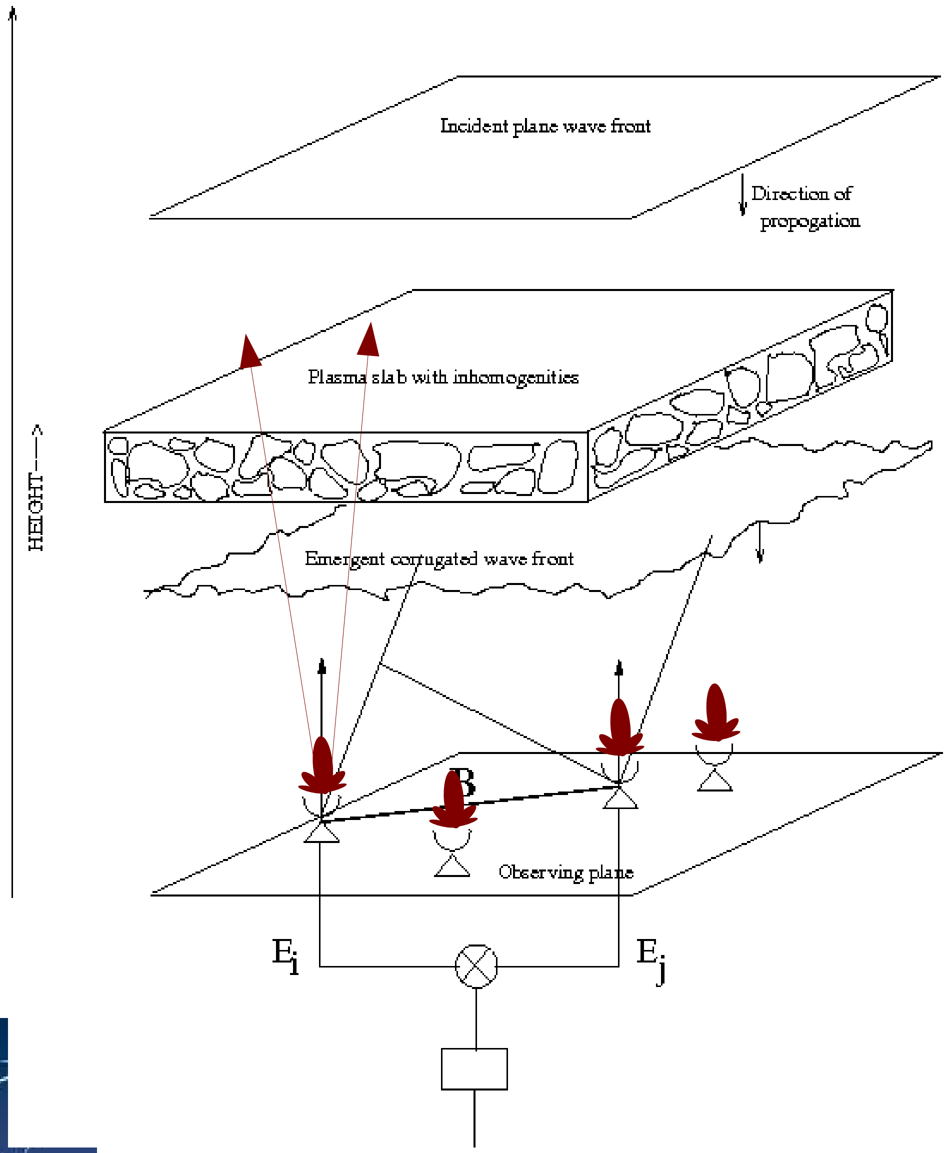
- An *N*-element array instantaneously measures *N(N-1)/2* baselines (complex values)
- All antennas track the ref. Dir. to compensate the earth rotation
- Typical imaging observations track for several hours
- More terminology
  - Baseline vector measured in a frame with the uv-plane tangent to the sky
  - Baseline co-ordinates

$$B_{ij}(t) = (u_{ij}(t), v_{ij}(t), w_{ij}(t))$$

$$u_{ij}(t) = u_i(t) - u_j(t)$$

  - Only relative separation between antennas matter
- Max. baseline corresponds to the size of the "aperture"
- Baseline vector changes as earth rotates filling the aperture
  - Aperture Synthesis/Earth Rotation Synthesis

Aperture Synthesis



- Cosmic signals interact with Earth's ionosphere (low freq.) and atmosphere (high freq.) resulting in image degradation
- Further degradation due to non-ideal antenna far-field patterns (Primary Beam), pointing errors, etc.
- Such effects are in general
  - Direction Dependent
  - Time varying
  - Frequency and polarization dependent
  - Fundamentally antenna-based
- Signals from the sky are in general
  - Also obviously direction dependent
  - Freq. and polarization dependent
  - Fundamentally not antenna-based

Aperture Synthesis: Theory

- Basis of imaging: van Cittert-Zernike Theorem:

  For small field-of-view (FoV) or for n<<1, image is 2D Fourier Transform of the Visibility (Coherence Function) (Ref: Born&Wolf)

- With finite number of antennas, the uv-plane is not fully sampled:

$$V_{ij}^{Obs} = S_{ij}\left[\int_{\sigma} I^{M}(\sigma)\, e^{2\pi\iota(u_{ij}l + v_{ij}m)}\, d\sigma\right] = \sum_{t}\left[S_{ij}(t)\,.\,V_{ij}^{Sky}(t)\right]$$

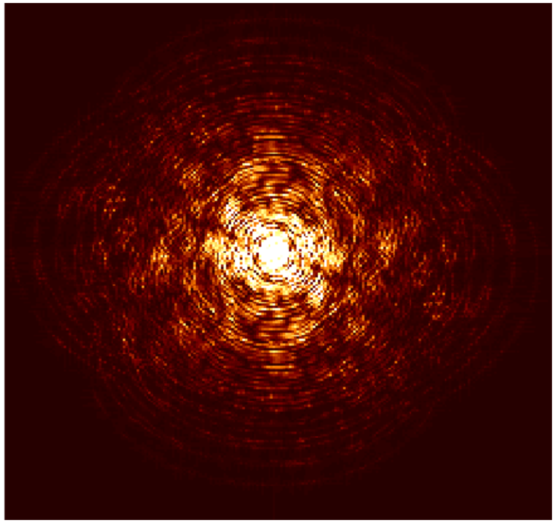
- Small FoV Measurement Equation:

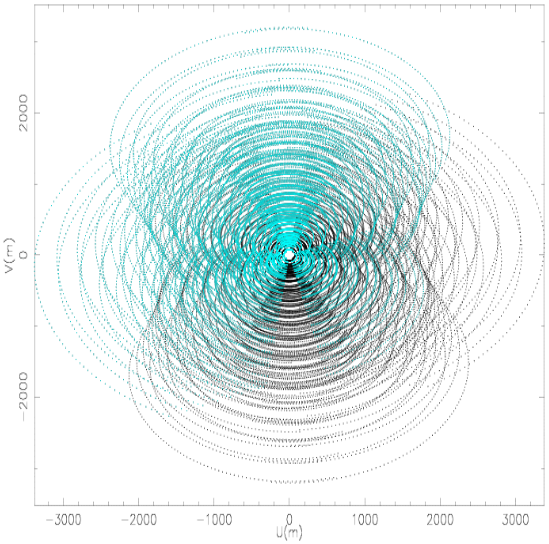
$$V^{Obs} = [S]\,V^{Sky} = [S]\,[F]\,I^{sky}$$

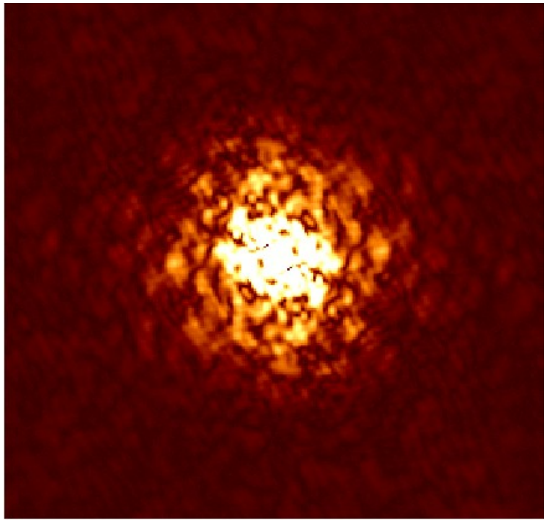
- S: The uv-coverage, Sampling Function (Transfer Func.)

  PSF: Fourier Transform of S (Impulse Response)

Synthesis Imaging: Data Domain





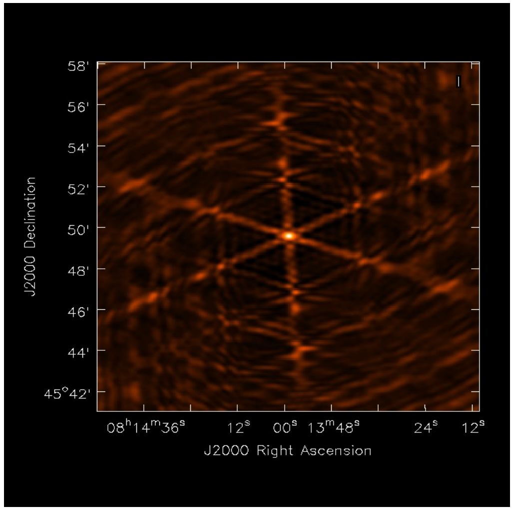


$$V^{Obs} = S \cdot V^{o}$$

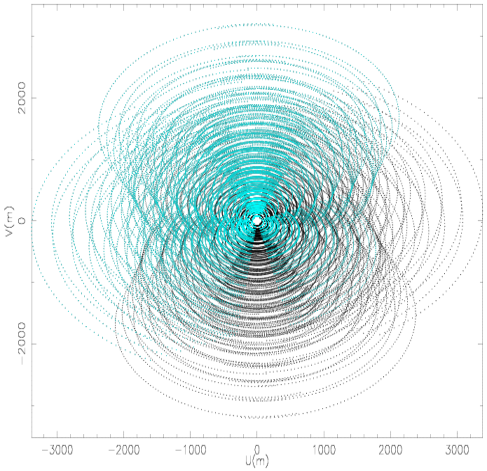
[Complex] [Real] [Complex]

- Incomplete sampling of the data domain.
- Related to s2MRI (talk from yesterday)?

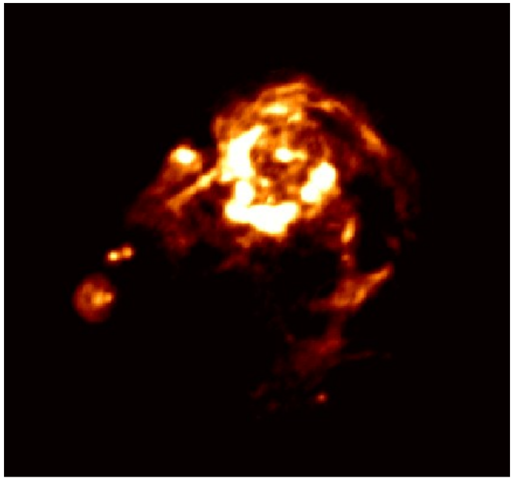
Synthesis Imaging: Image Plane



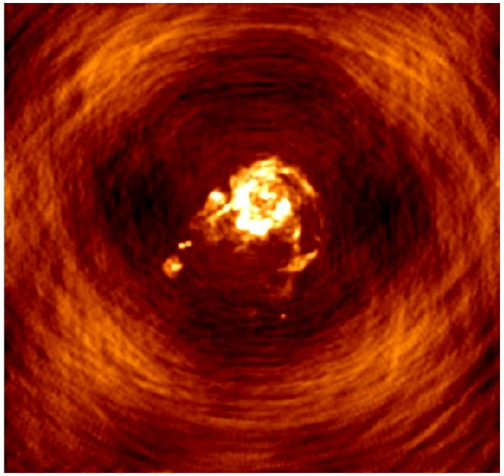
$PSF = FT[S]$



Sampling function

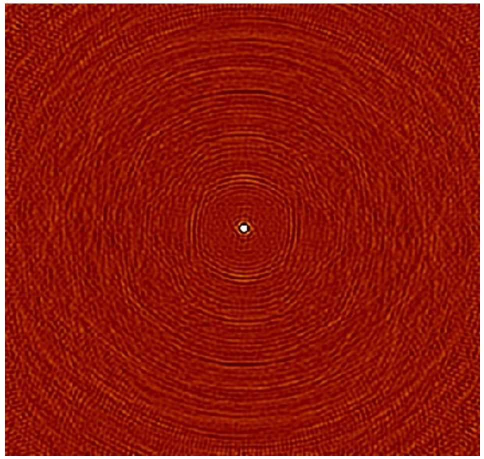


$I^o = FT[V^o]$

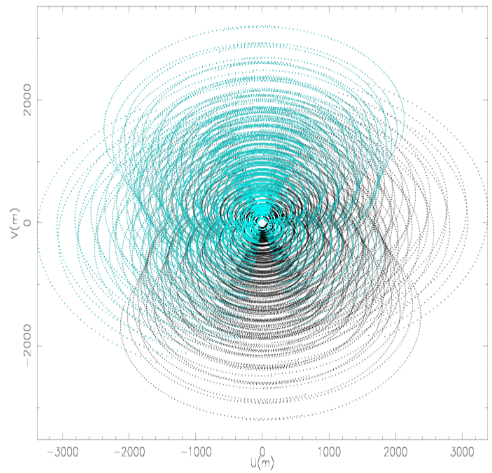


$I^d = FT[V^{Obs}] = PSF * I^o$

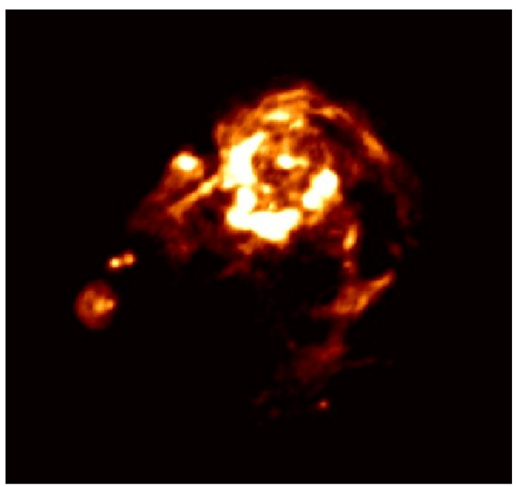
Synthesis Imaging: Image Plane



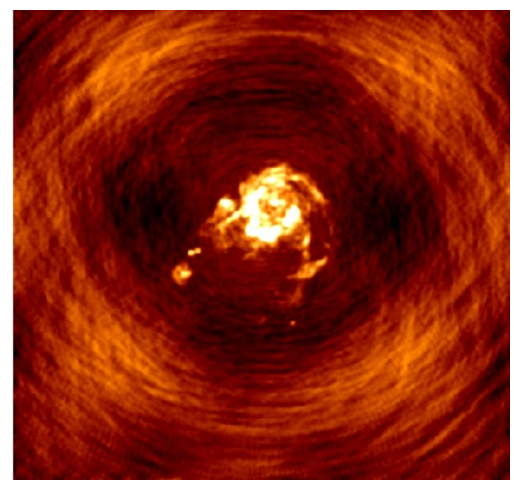
$PSF = FT[S]$



Sampling function



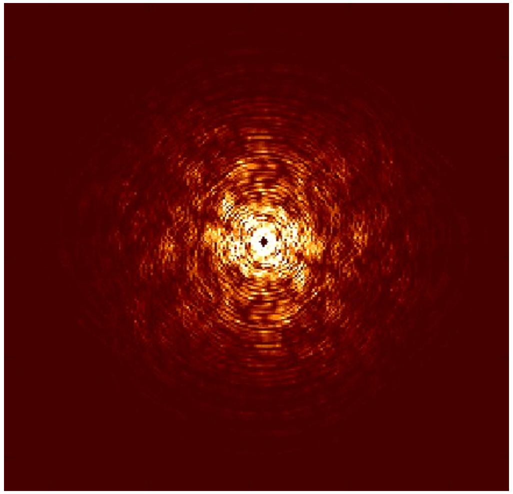
$I^o = FT[V^o]$



$I^d = FT[V^{Obs}] = PSF * I^o$

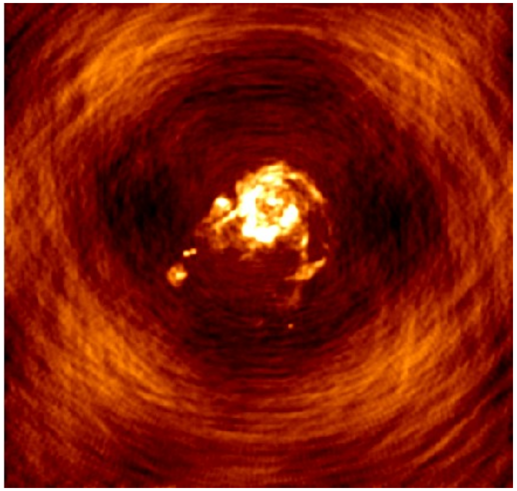
Imaging and Image Reconstruction

- Imaging: Transform the Visibility data to the image domain
  - True sky convolve with the PSF
  - The “Dirty Image”
- Visibility data is not on a regular grid
  - Needs re-sampling on a regular grid to utilize the computational advantages of the FFT algorithm
- Re-sampling done via convolutional interpolation
  - The “Gridding” operation $V_{ij}^{G} = \left[C * V^{Sky} \right]_{ij}$



FFT

↔

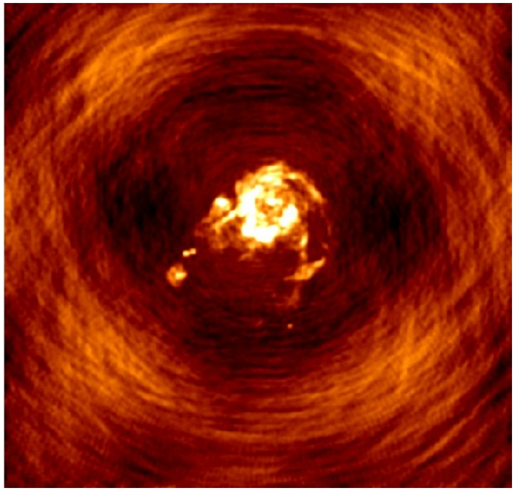


Imaging and Image Reconstruction

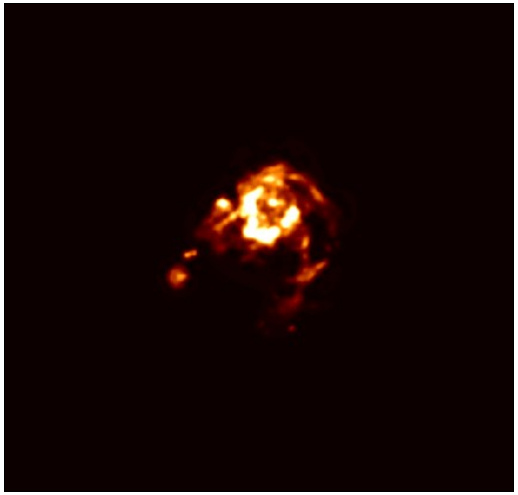
- Image dynamic range of Dirty Image: few x 100: 1
  - Typical instrumental dynamic range: 10^6: 1
- Image reconstruction:

$$Minimize: \chi^2 = \left| V^{Obs} - A\, I^M \right|^2 \; where \; I^M = \sum_k P_k ; \quad P_k \equiv Pixel\ Model$$

- In general, ***A*** is singular (is rectangular): image reconstruction algorithms are non-linear and iterative in nature:
  - Most commonly used algorithm: CS-Clean



→



Deconvolution

- Process of removing the effects of the emission in one part of the image on another part of the image
  - PSF sidelobes couples distant, otherwise independent pixels
  - Mathematically, even for an image with only multiple point sources, the Hessian is not diagonal (or diagonally dominant)
- Only average quantities are available in the image domain
  - Time and frequency averaging to realize higher sensitivity
  - Averaging across uv-plane
- Purely image-plane based deconvolution applicable only for the static case (along time, frequency and polarization axis)
  - **Hogbom Clean**: Static case, limited by quantization errors
  - **Clark Clean**: Static case + partially handle quantization errors
  - **Cotton-Schwab (CS) Clean**: Static case + handle quantization errors
  - **Multi-Term MFS (minor) + CS-Clean (major)**: Time-static, Freq-dynamic case
  - **Projection (major cycle) + MT-MFS (minor)**: Time- and Freq-dynamic case

Deconvolution as ChiSq Minimization

- $V^M = A\, I^M + N$ $\qquad$ N is a Gaussian random process (from Physics)
- Linear equation, parametrized by I^M
  - However, A is singular
- Need non-linear solvers to solve for I^M ("two level" iterations)
  - Compute residuals $V^R = V^{Obs} - AI^M$
  - Make residual Image $I^R = [F]V^R$
  - Find update direction: Steepest Descent

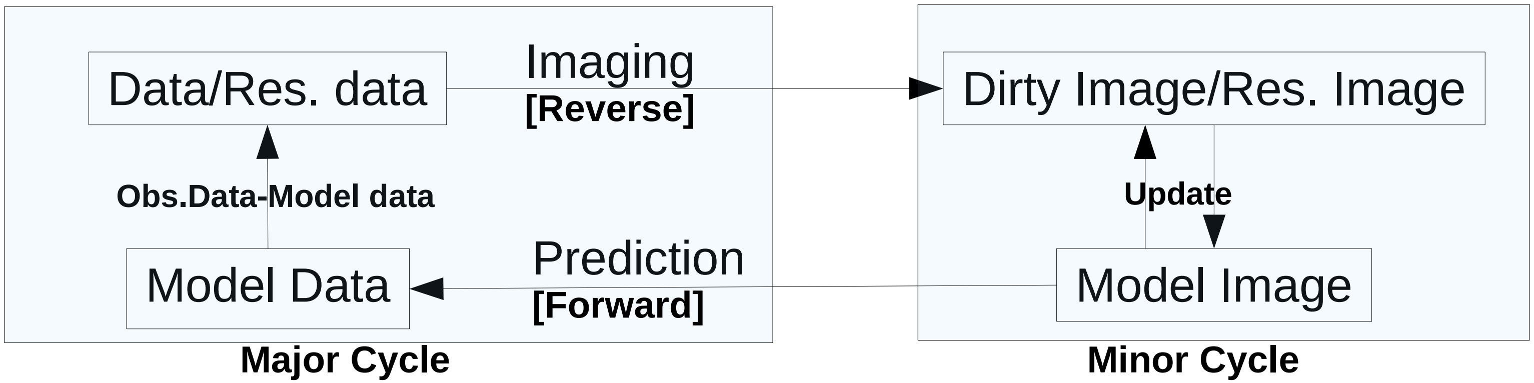
$$I^c = max\left(-2[I^{Res}]\frac{\partial \chi^2}{\partial\, Param}\right)$$

  - Update model: $I_i^M = T(I_{I-1}^M)$ $\qquad$ $for\ our\ discussions\ this\ is = I_{i-1}^M + \alpha * I_i^c$
  - α Is the loop-gain/step-size

Major Cycle (always expensive)

Minor Cycle (can be expensive)

Image Reconstruction



- Data prediction (predict data from a given image model)

  $\boldsymbol{V} = \boldsymbol{A}\,\boldsymbol{I}^{o} + \boldsymbol{N}$ $\qquad V_{ij} = deGrid_{ij}\, FT(I)$

- Imaging: $\boldsymbol{A}^{T}\boldsymbol{V} = \boldsymbol{A}^{T}\boldsymbol{A}\,\boldsymbol{I}^{o} + \boldsymbol{A}^{T}\boldsymbol{N}$ $\qquad I^{Dirty} = PSF * I + PSF * Noise$

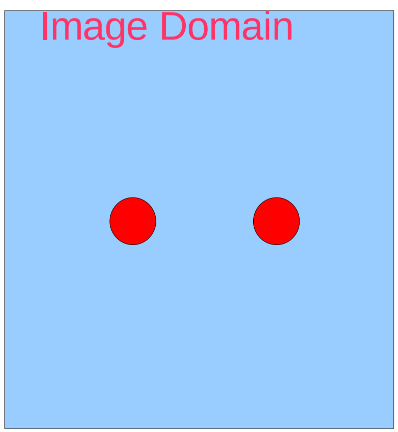
- Approximate reverse transform (derivative computation)

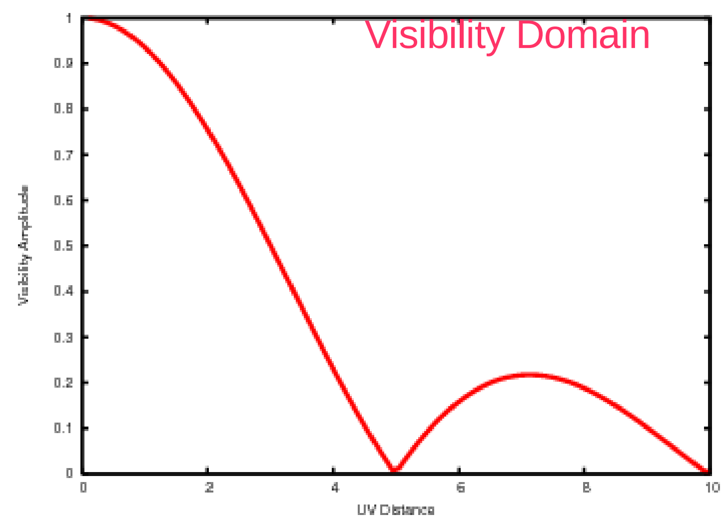
  Accurate forward transform
  - Construct approximate $\boldsymbol{A}^{T}$ to include DD effects

- Noise in the image plane is not independent per pixel
  - Correlated at the scale of the PSF

Natural domain of parameters

- $$V_{ij}^{Obs}(\nu) = M_{ij}(\nu,t) W_{ij} \int M_{ij}^{S}(s,\nu,t) I(s,\nu) e^{2\pi\iota(b_{ij}.s)} ds$$
- Unknowns:
  - M^S_{ij}: Instrumental/atmospheric DD effects, time and freq. dependence
  - I(s): Complex structure of the source, frequency dependence
- Modeling the domain where the information one seeks naturally resides is the optimal/natural domain (maximizes the information content)

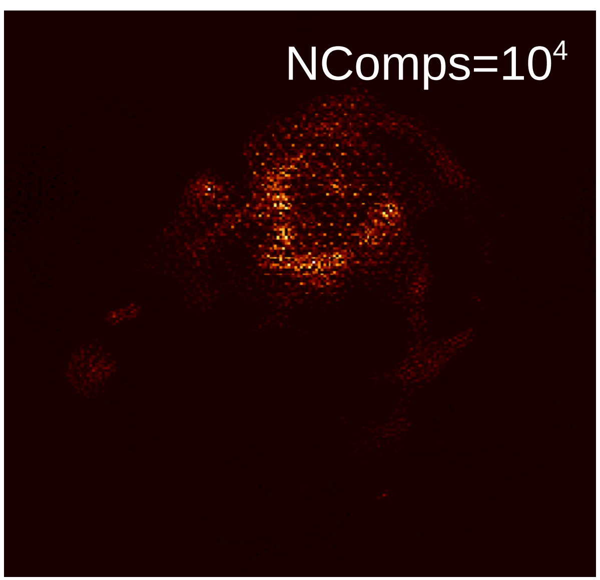


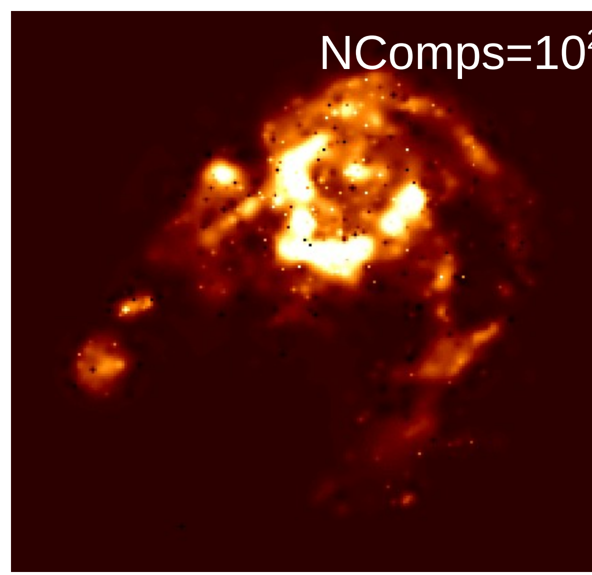


Parametrized model for sky emission

- $$V_{ij}^{Obs}(\nu) = M_{ij}(\nu,t) W_{ij} \int M_{ij}^{S}(s,\nu,t) I(s,\nu) e^{2\pi\iota(b_{ij}.s)} ds$$

- The function ***I(s)*** represent sky emission
  - Information it represents is inherently in the sky domain
    - – Parametrize structure: Asp-Clean, MS-Clean
    - – Parametrize frequency dependence: MT-MFS, MS-MFS

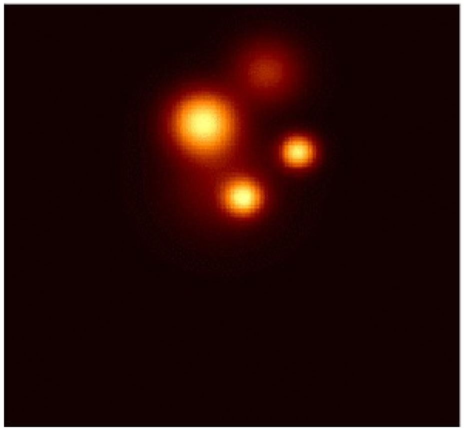




- Better parametrization in the Natural Domain

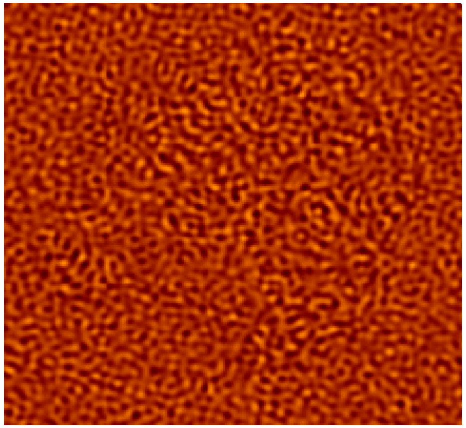
Image Reconstruction

- Iterative build the model image in minor cycle



- Stop when peak residuals are greater than the effects of the approximations
- Trigger reconciliation with the data (Major Cycle)

- Compute residuals in major cycle (expensive)



- Stop when convergence criteria is satisfied

DD Effects

- In principle $$V_{ij}^{Obs}(\nu,t)=s_{ij}\int_{\sigma} I(\sigma)\, e^{2\pi\iota(u_{ij}l+v_{ij}m)}\, d\sigma$$
- In practice:

$$\boldsymbol{V_{ij}^{Obs}(\nu,t)=S_{ij}G_{ij}\int_{\sigma} X_{ij}(\sigma,\nu,t)\, I(\sigma)\, e^{2\pi\iota(u_{ij}l+v_{ij}m+w_{ij}(n-1))}\, d\sigma}$$

- $\boldsymbol{G_{ij}}$ is the antenna-based Direction-Independent term
  - Solved using the SelfCal algorithm (early 1980s)
- $\boldsymbol{X}$ is the antenna-base DD term
  - Antenna pointing errors, Primary Beam, Geometrical effects
  - Rest of this talk + talks by Cotton (yesterday), Smirnov (next talk)
- $\boldsymbol{I}$ is the image-plane based DD term (non antenna-based)
  - Talk by Rau (last talk of this session)

DD Terms: Projection Algorithms

- Two fundamentally different approaches being pursued:
  - Projection Algorithms
  - Partitioning Algorithms

- **Projection Algorithms**: Model the DD-terms in the natural domain
  - Solve for the parameters of the models
- Domain of compact representation/sparse domain
  - Data domain for antenna-based terms
  - Image for non antenna-based terms
- Physical modeling of the effects
- Minimize the degrees-of-freedom (DoF)
  - Independent measurements: $O(N^2)$
  - No. of parameters: $O(\text{few x N})$
- Lower complexity
  - Complexity independent of the complexity of the source

DD Terms

- ***X*** is the antenna-base DD term

$$V_{ij}^{Obs} = \left[V^{Sky} * G_{DD} \right]_{ij} \qquad G_{DD} = FT\left[X \right]$$

- Sources of DD effects:
  - PB effects: $$G_{DD} = FT\left[PB \right]$$
  - Antenna Pointing Errors $$G_{DD} = FT\left[PB \right] e^{\iota(\phi_i - \phi_j)}$$
  - Effects of spherical geometry for wide-field imaging: The W-Term: $$G_{DD} = FT\left[e^{\iota w_{ij}(n-1)} \right]$$
- Precise PB shape, antenna pointing errors are unknown:
  - Need to solve for the appropriate parameters of X
  - E.g. Pointing SelfCal Algorithm
- W-term is known geometrical effects
  - Can be pre-computed

Advances in Calibration and Imaging Techniques in Radio Astronomy, Rau et al., Proc. IEEE, Vol. 97, No. 8, Aug.2009, 1472

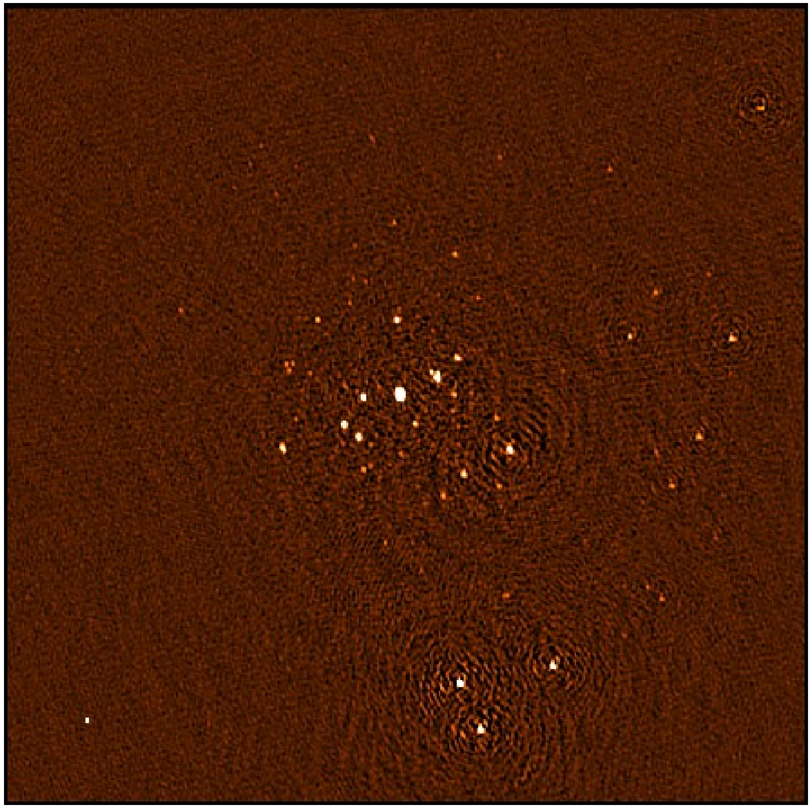
DD Terms: Partitioning Algorithms

- Apply classical Direction Independent techniques to solve for DD terms (piece-wise constant approximation)
  - Partition the image-plane such that DI assumption is valid in each partition
  - Apply DI techniques to each partition and stitch

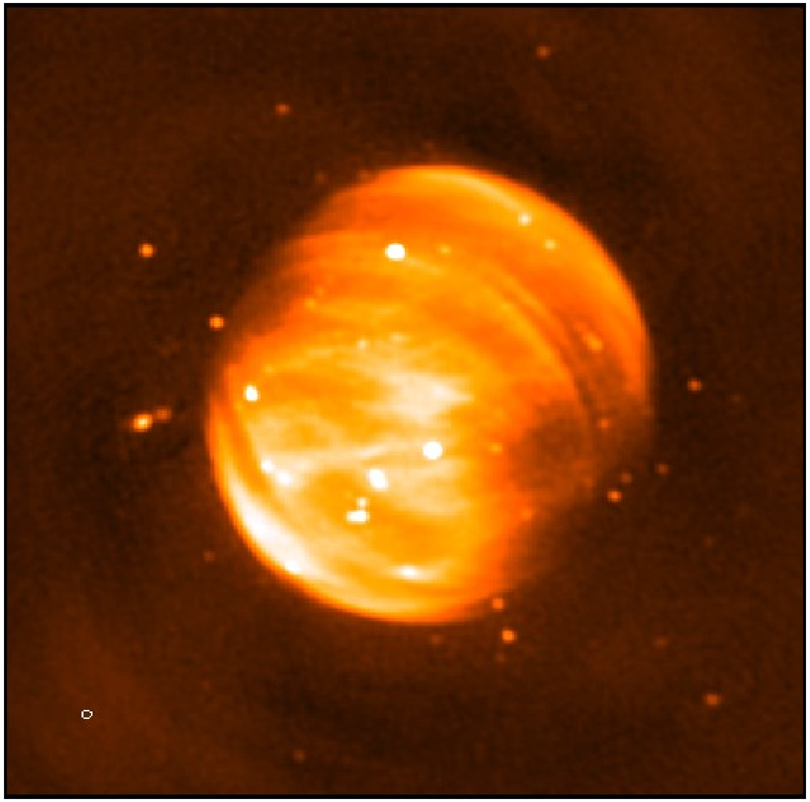
- Conceptually easier to understand
  - Possibly because classical understanding can carry-over
- Absorbs the combination of all effects
  - Phenomenological approach (sometimes useful)

- Has trouble at partition boundaries
- DoF: O(few x N x No. of partitions)
- Higher complexity
  - Complexity a strong function of complexity of the celestial source
  - Complexity increases when wide-band and polarization effects are included (required for modern telescopes)

Range of imaging challenges



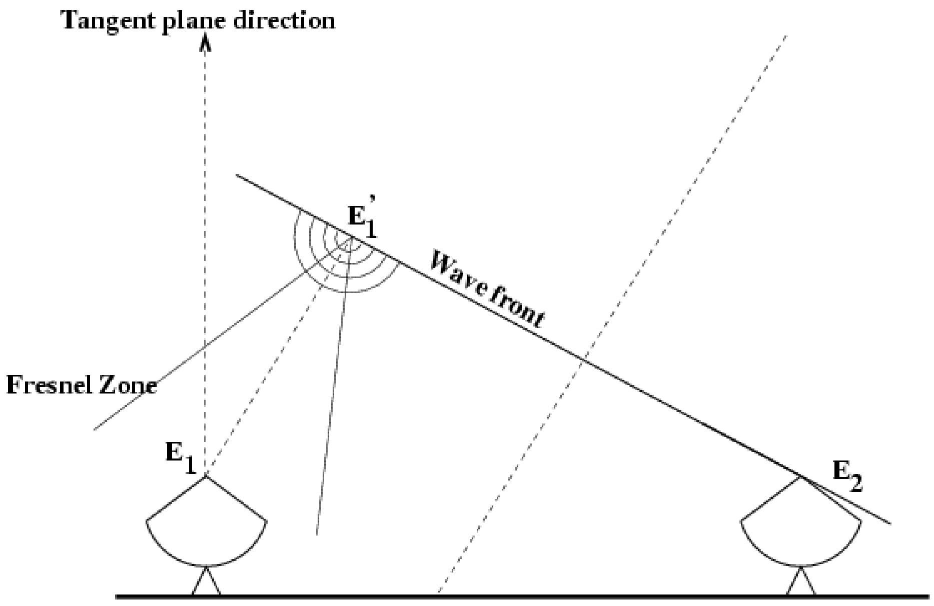
Field with compact sources filling the FoV



Compact + extended emission filling the FoV

- Useful algorithms must efficiently handle a large range of scales
- Deal efficiently with multiple iteration through TBytes of data

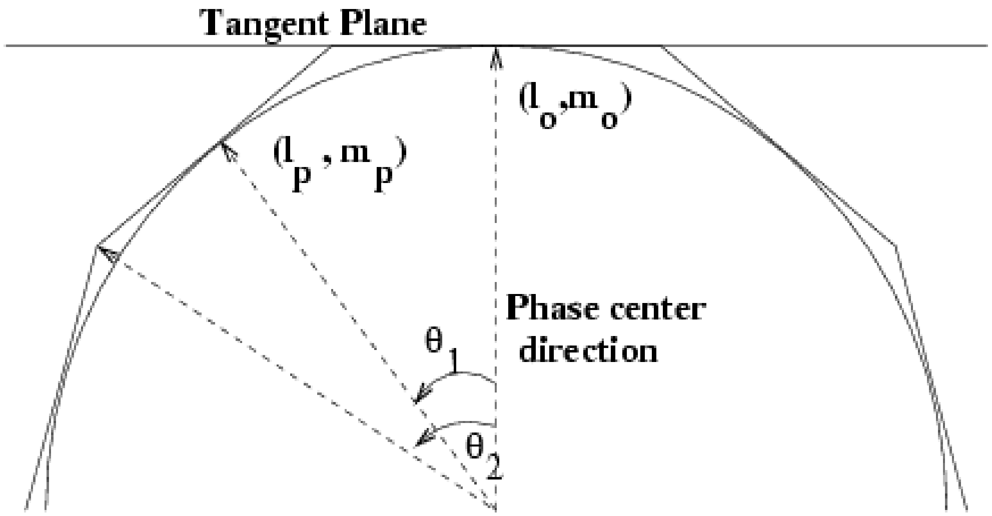
The W-Term: Projection Algorithm



- We measure: $V^o_{12} = \langle E_1 E_2^* \rangle$
- We interpret: $V^o_{12} = \langle E_1' E_2^* \rangle$
- We should interpret E_1 as [E_1' x Fresnel Propagator]
- Pre-compute $G_{DD} = FT\left[e^{\iota w_{ij}(n-1)}\right]$
- Imaging: $I^d = FT\left[V^{Sky} * G_{DD}\right]$

[Cornwell, Golap & Bhatnagar, A&A]

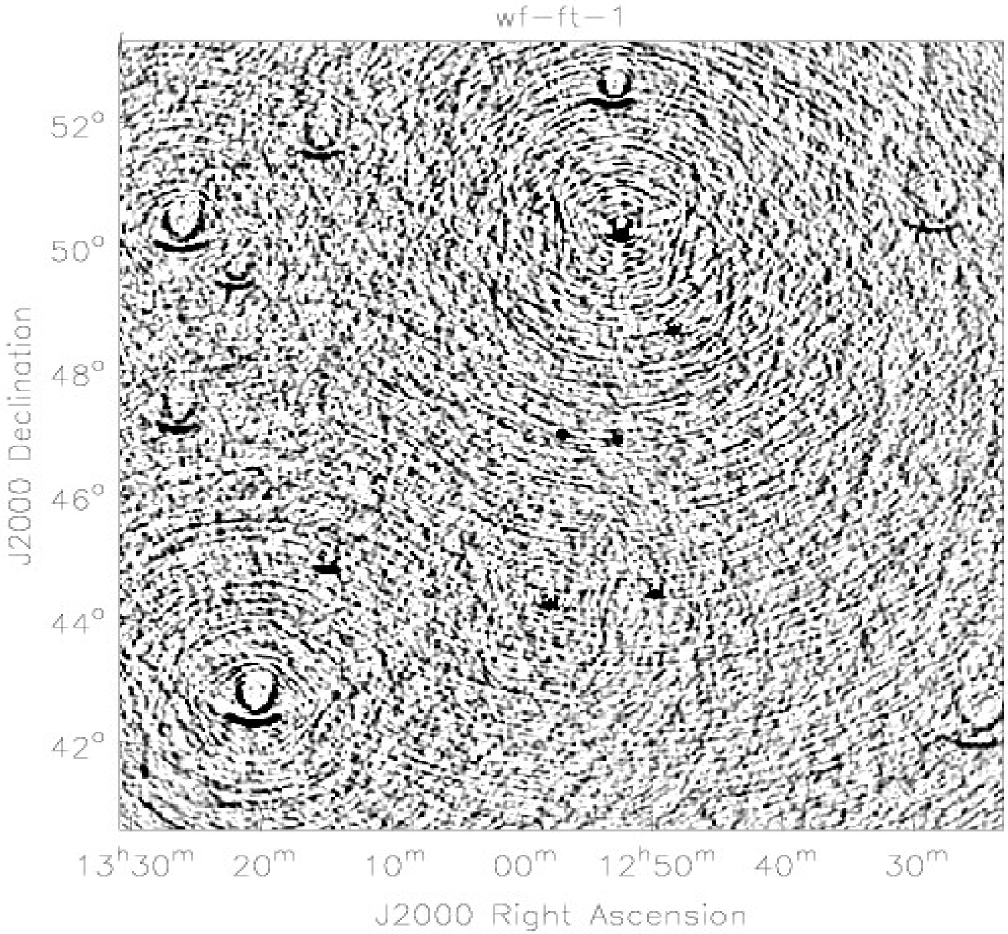
The W-Term: Partitioning



- Partition the sky into facets where 2D approximation is valid and classical techniques applicable
  - Equivalent partitioning in the data domain possible, but similar performance and computing requirements
- Stitch together the facets to make a single image
- Can be extended to also solve for DD effects, assumed constant across each facet ("Peeling", ref. Next talk)

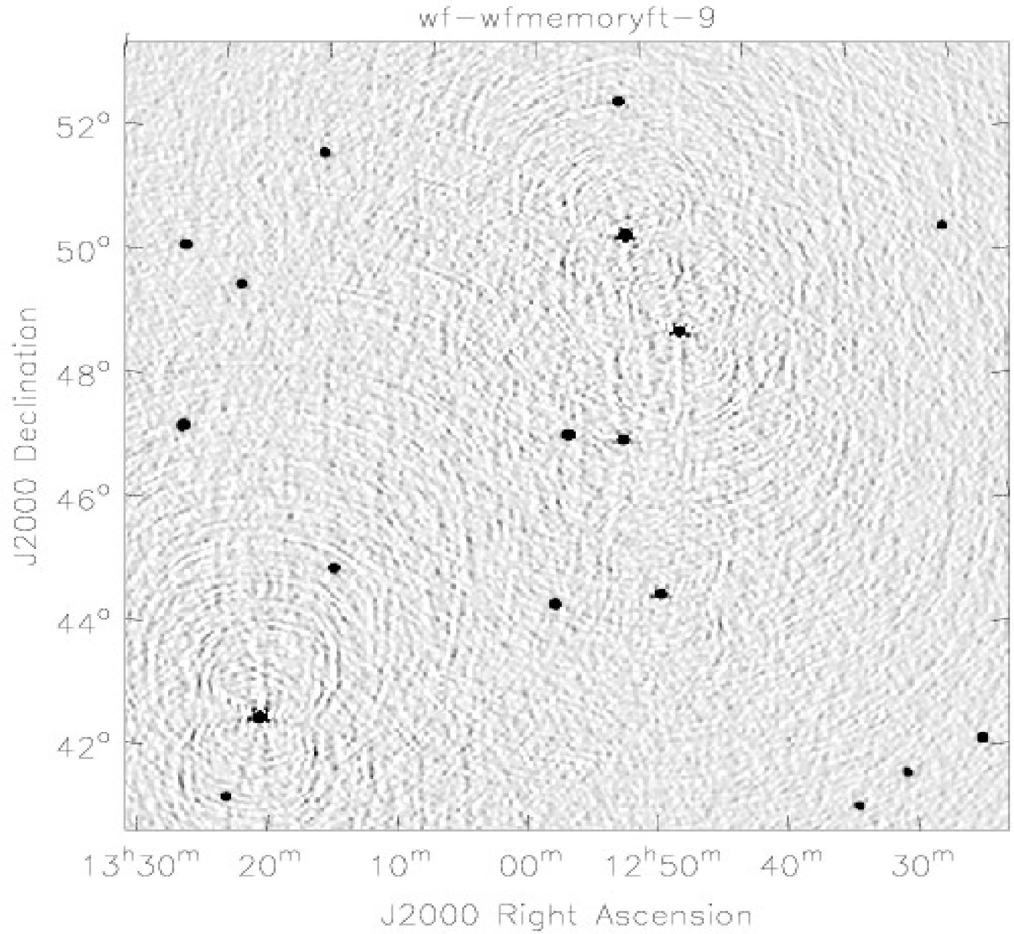
The W-Term: No correction

`

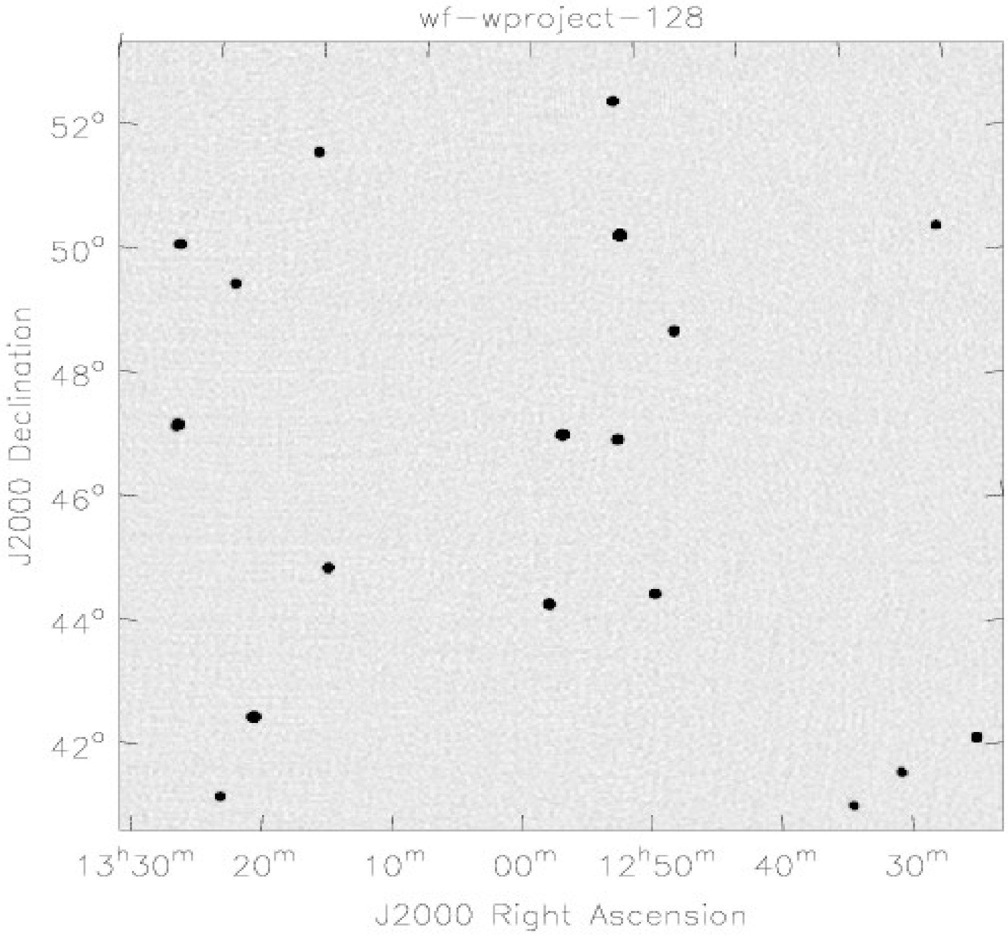


The W-Term: Partitioning

`

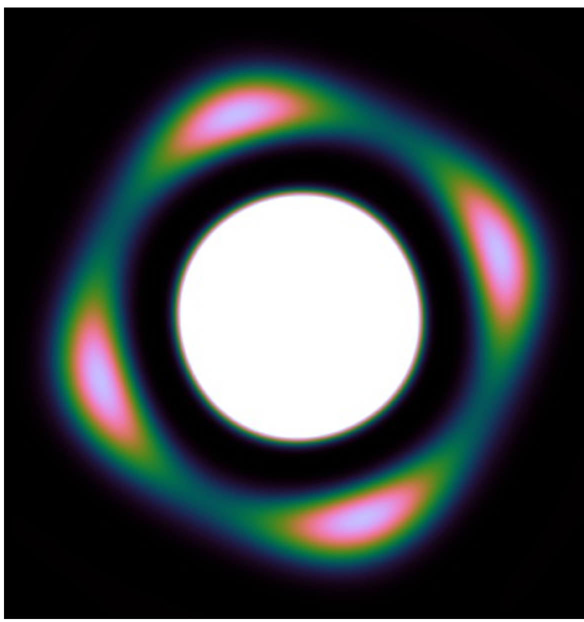


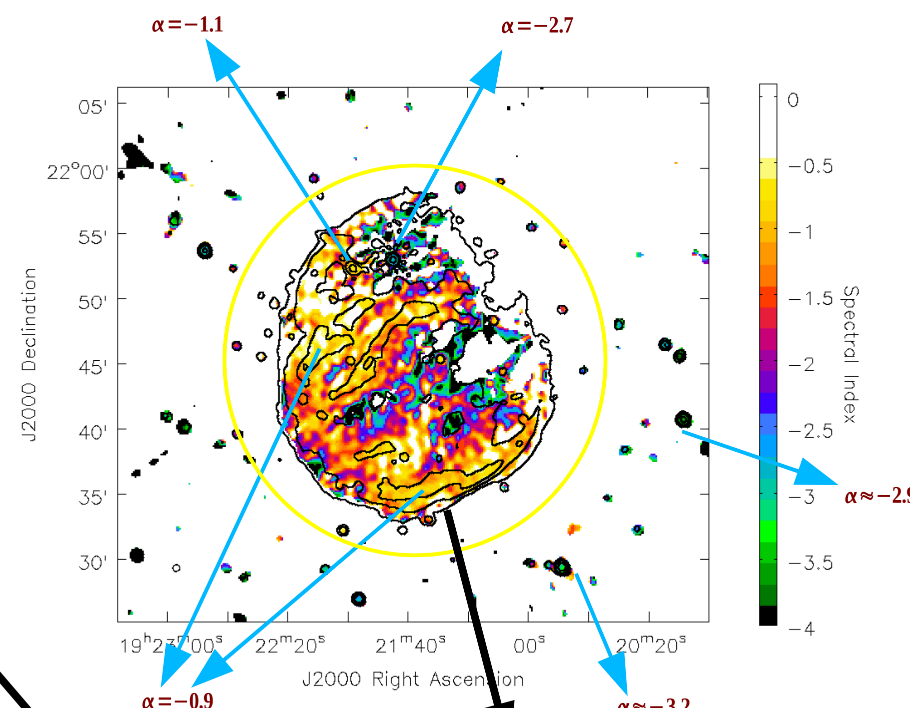
The W-Term: Projection



PB effects

- Time variability of the PB increases away from the center
- Frequency dependence increases with fractional bandwidth

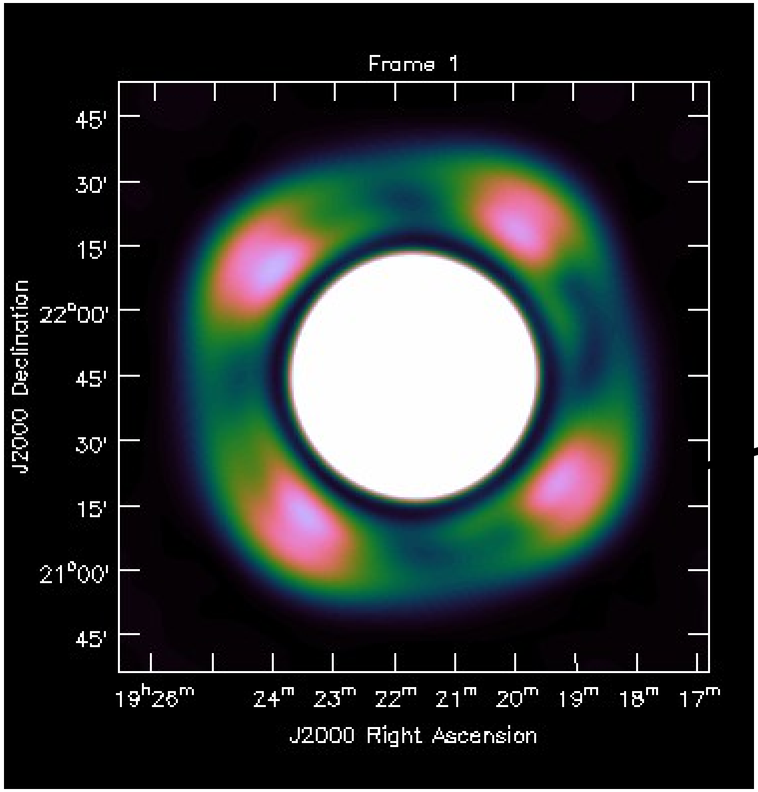




$$I^{Obs} = \sum_t \sum_\nu PSF(t) * [PB(s) I^{Sky}(s)(\nu/\nu_o)^{\alpha(s,\nu)}]$$

Wide-band PB effects

- To the first order, scaling of the PB with frequency

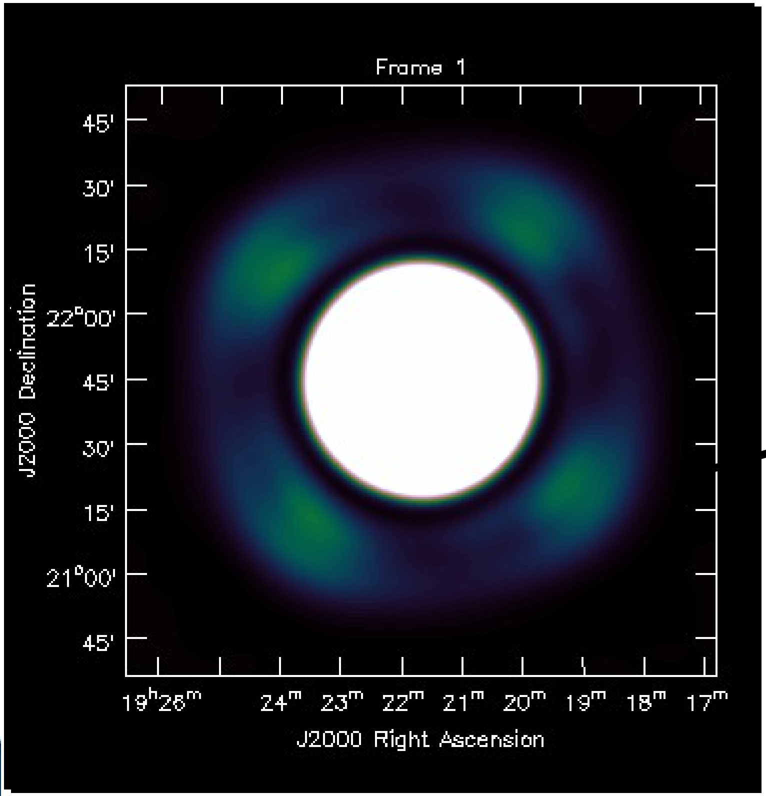


$$I^{Obs} = \sum_t \sum_\nu PSF(t) * \left[PB(s,\nu,t) I^{Sky}(\nu) \right]$$

$$\sum_\nu PB(s,\nu,t)$$

High sensitivity imaging

- Image corresponds to the sum of all the data.
  - Only average of antenna-based quantities are available in the image domain



$$I^{Obs} = \sum_t \sum_\nu PSF(t) * \left[PB(s,\nu,t) I^{Sky}(\nu) \right]$$

$$\sum_t \sum_\nu PB(s,\nu,t)$$

- Image domain corrections for time, frequency and antenna dependence is hard

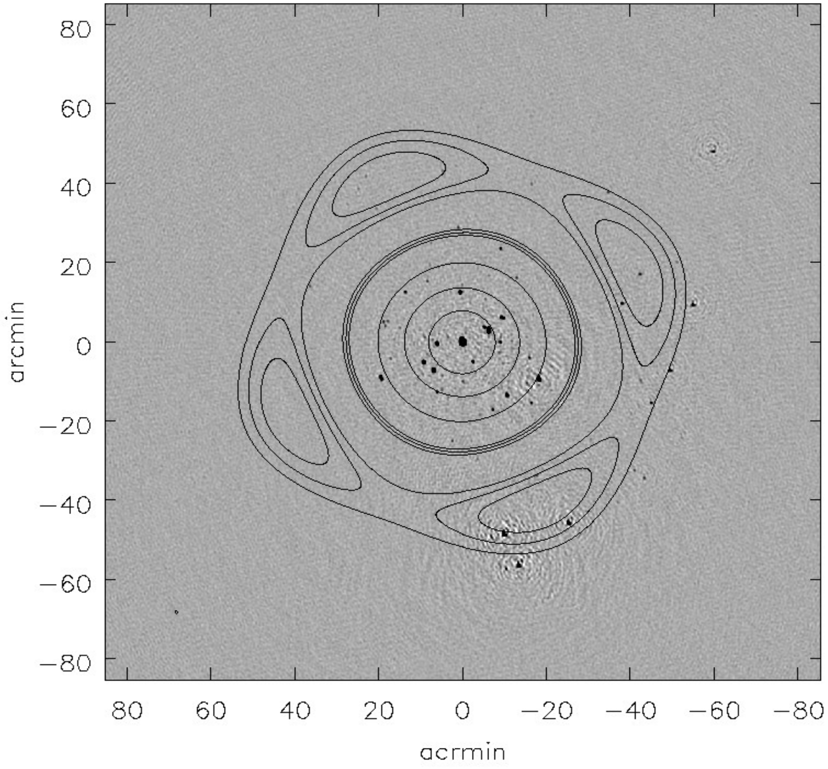
- Projection methods apply corrections in the Natural Domain
  - A-Projection for PB-corrections
  - W-Projection for W-term correction

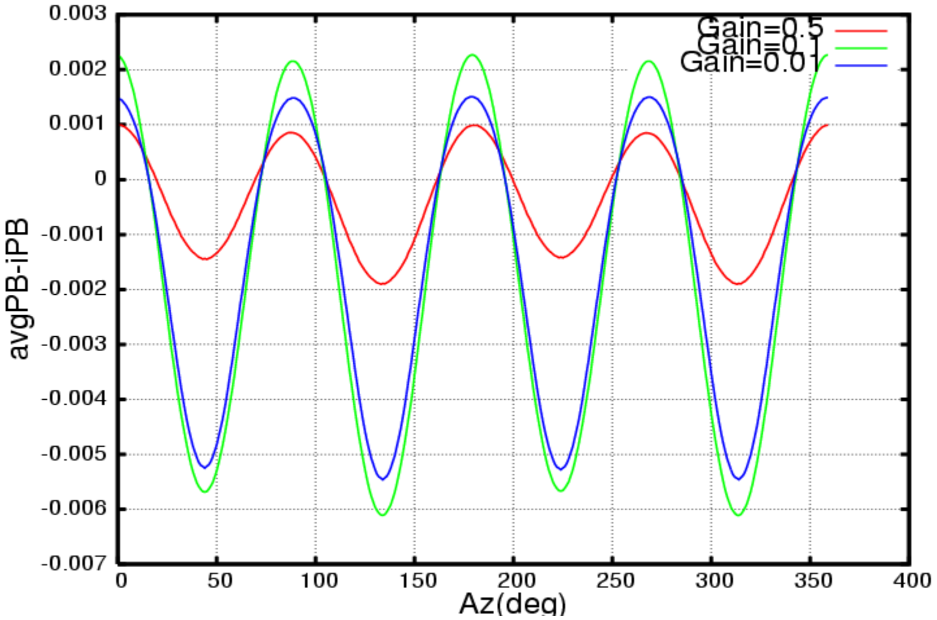
Wide-field Imaging: PB effects

- The observed data corresponds to I^{sky} multiplied by the antenna primary beam

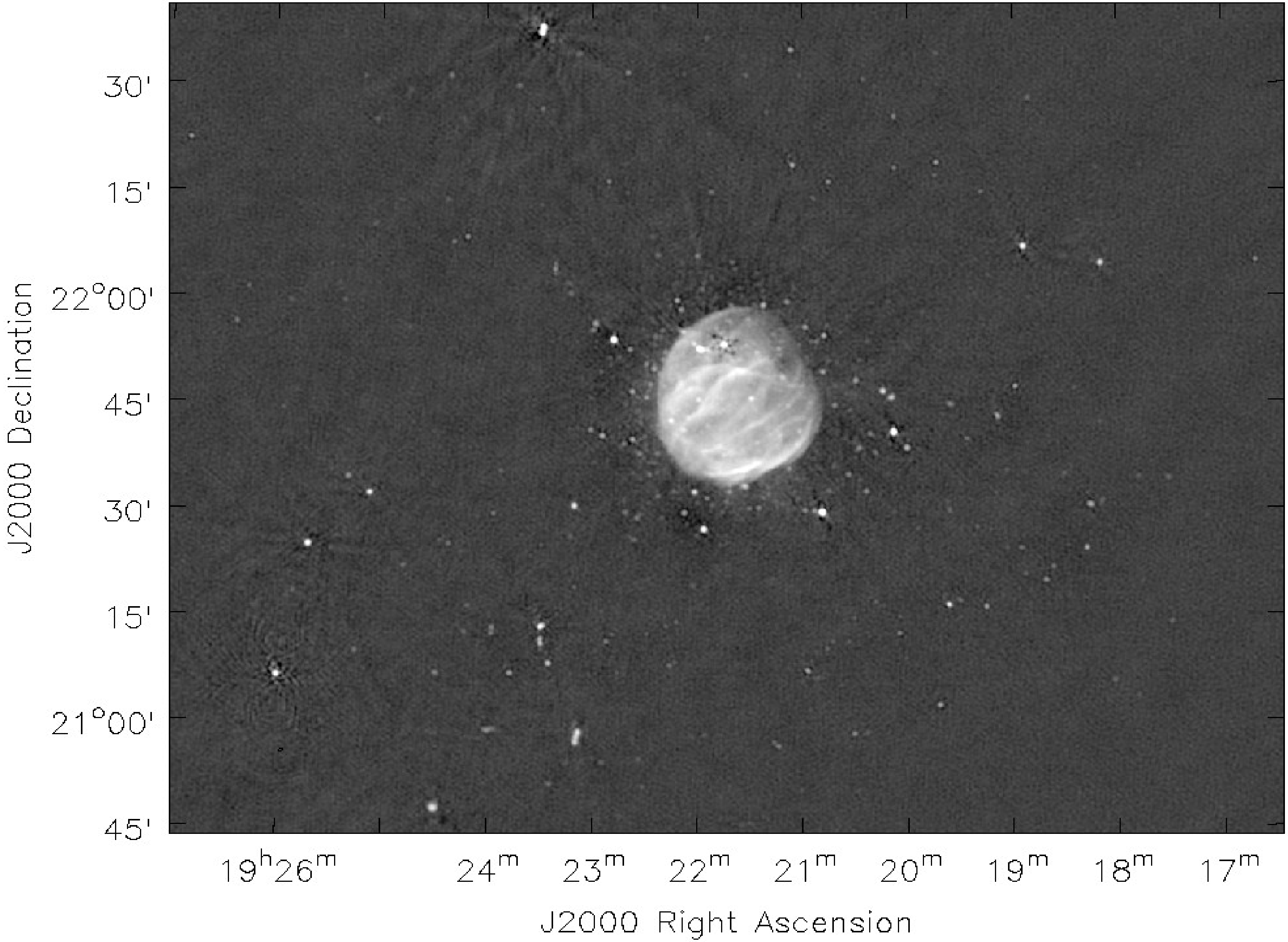
$$I^D = \sum_t \sum_\nu PSF(\nu, t) * \left[PB(s, t) \cdot I^{Sky} \right]$$

- PB varies with time due to rotation with PA and pointing errors.
- PB gain in general is also Directionally Dependent

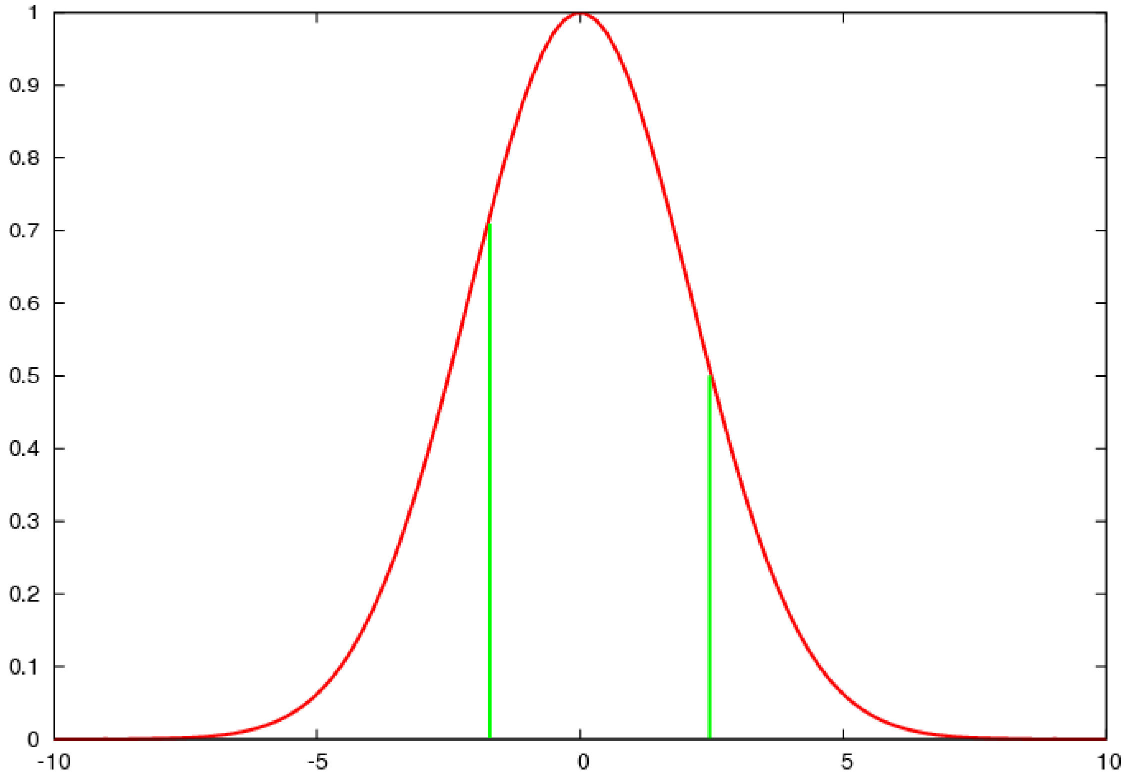




Wide-field Imaging: EVLA

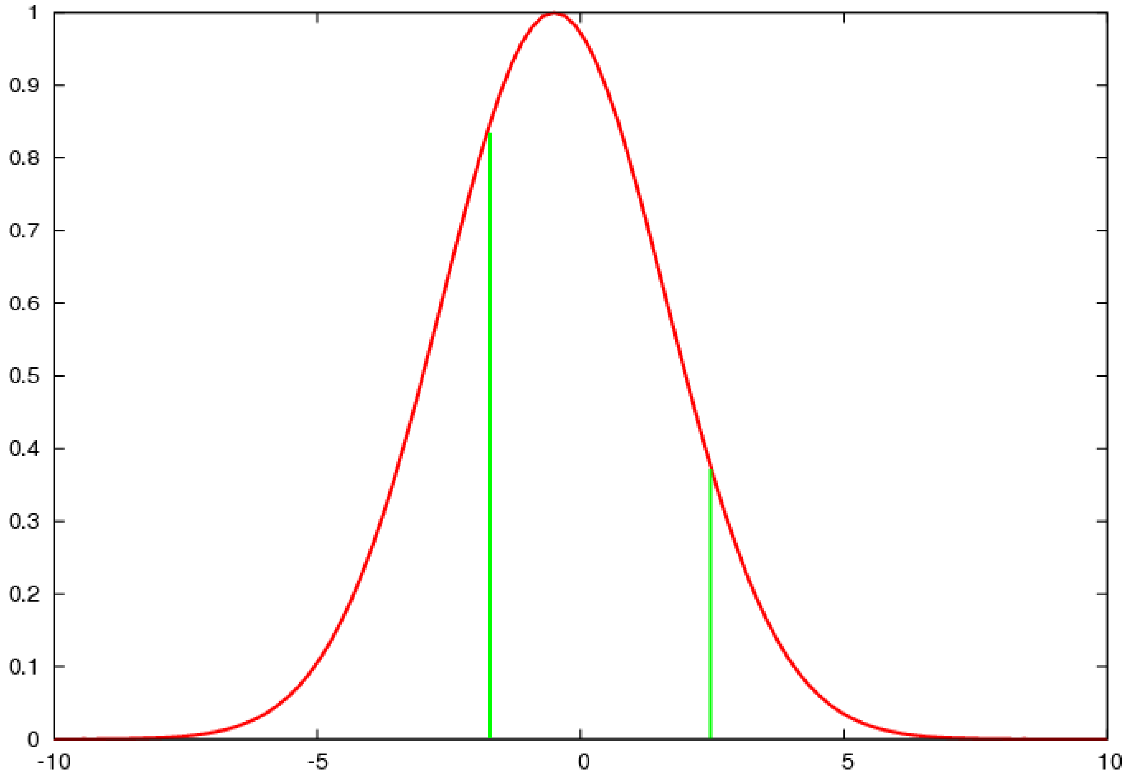


Wide-field Imaging: Pointing Errors



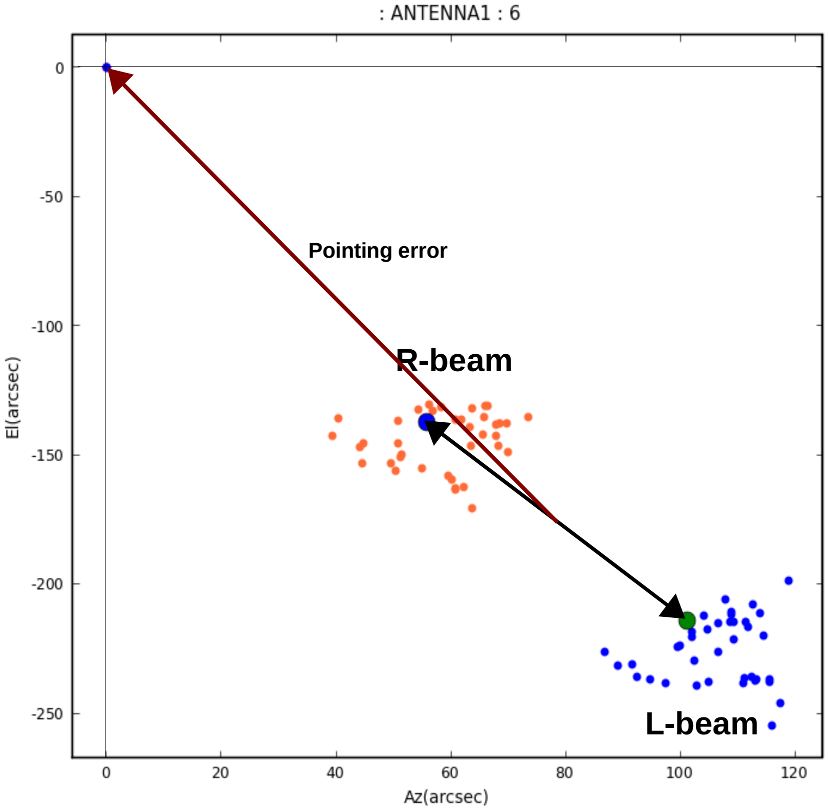
- Effect of antenna pointing error is a direction dependent effect
- A purely Hermitian effect in the data domain, in the absence of DI gains
  - To the first order, amplitude-only error in image domain
- However, there is significant in-beam phase structure –particularly for wide-field, full-Stokes imaging

Wide-field Imaging: Pointing Errors



- Effect of antenna pointing error is a direction dependent effect
- A purely Hermitian effect in the data domain, in the absence of DI gains
  - To the first order, amplitude-only error in image domain
- Faceting approach:
  - Solve for gains for A and B separately
  - Interpolate in between
- Pointing SelfCal
  - Use A-Projection with pointing terms
  - Solve for the shape of the function which best-fits the gain variations at A and B

Wide-field Imaging: Pointing Errors



- El-Az mount antennas
- Polarization squint due to off-axis feeds
  - The R- and L-beam patterns have a pointing error of +/- ~0.06

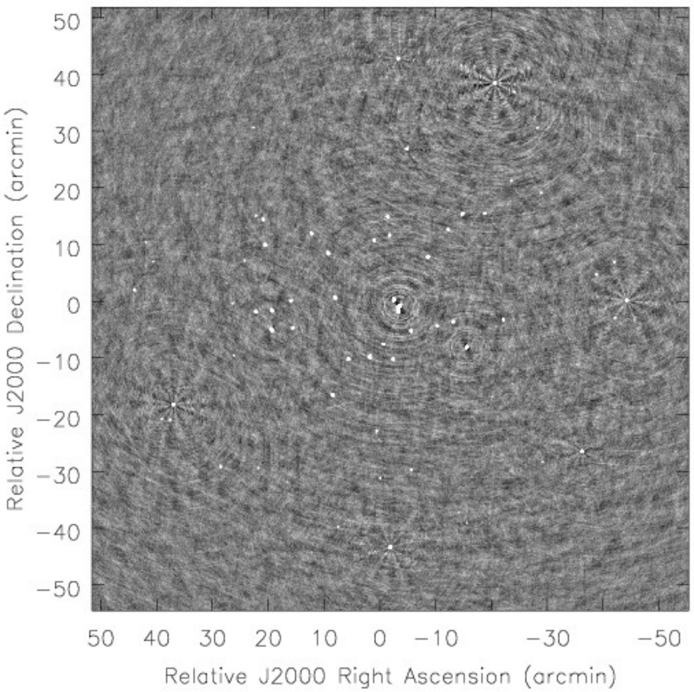
- DoF used: 2 per antenna
- SNR available for more DoF to model the PB shape

- EVLA polarization squint solved as pointing error (optical pointing error).
- Squint would be symmetric about the origin in El-Az plane in the absence of antenna servo pointing errors.
- Pointing errors for various antennas detected in the range 1-7 arcmin.
- Pointing errors confirmed independently via the EVLA online system.

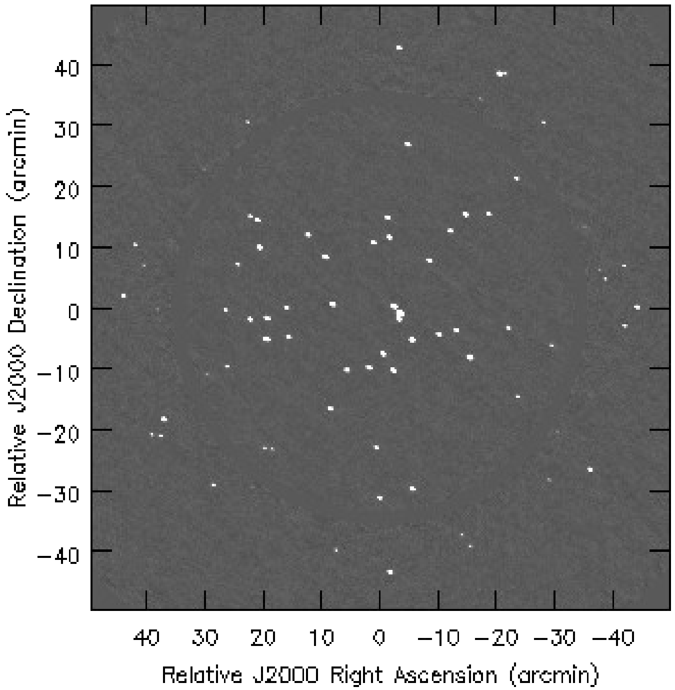
[Bhatnagar et al, 2004]

A-Projection algorithm

Before Correction



After Correction

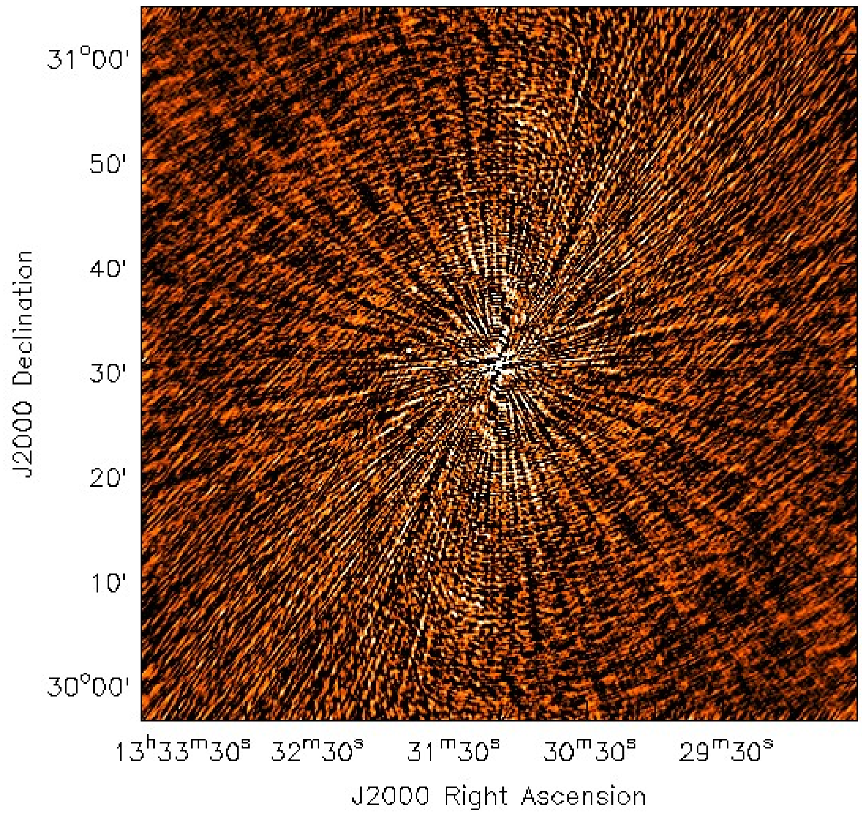


$$\text{Minimize}: V^{O}_{ij} - E_{ij} * [F I^{M}] \quad w.r.t. \quad I^{M}$$

Goal: Full-field, full-polarization imaging at full-sensitivity

A-Projection: Bhatnagar et al., A&A,487, 2008

MS-MFS + A-Projection



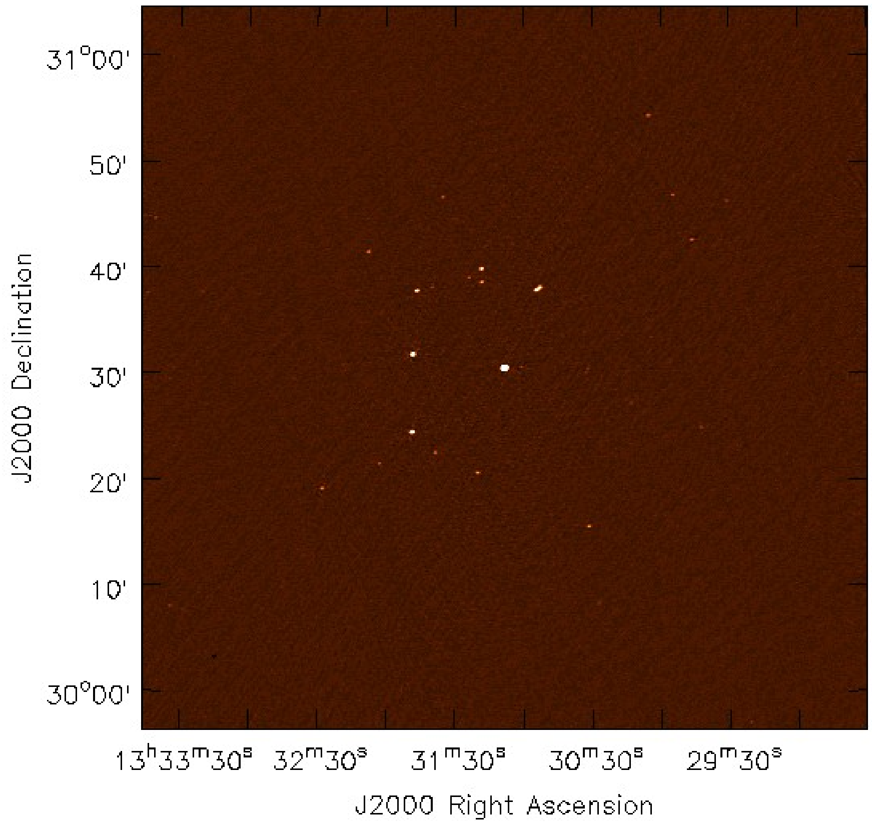


Image: Rua et al.
MS-MFS: Rau&Cornwell, A&A, 2011

DD Effects in RA and Blind Deconv.

- The effect of DD terms in the image domain is to make the "effective PSF" vary across the FoV
  - Effective PSF = PSF° x F(l,m)
  - PSF° in RA can be shown to be shift invariant
- RA algorithms utilize this fact to solve for parameterized F(l,m)
- Projection algorithms use physical parametrization
  - PB, Antenna Pointing errors, PB shape, etc.
- Partitioning algorithms similar to Blind Deconvolution
  - Solved for the PSF at multiple locations
  - However, indirectly inferring a model for F(l,m) is still required to make noise-like residual image

Discussion Items

- Two-level iterative image reconstruction in RA appears to be similar to some of the descriptions (CLASH?)
  - RA terminology: Major Cycle and Minor Cycle

- The A-Matrix (Measurement Matrix) is fixed and strictly defined by the physics of the observations
  - S_{ij}: Array configuration
  - Transforms include EM propagation (W-Term), antenna far-field pattern, polarization and frequency dependence...
  - Wide-band AW-Projection, MT-MFS

- Modern image reconstruction decomposes sky emission in sparse basis
  - Scale sensitive deconvolution

- We use approximate operators to compute update directions. Data prediction is accurate (leads to convergence).
  - Construct pseudo-inverse of the A-Matrix

Discussion Items

- Noise in the image is not independent
  - Correlated at the scale of the PSF (convolved with the PSF)
- PSF sidelobes also couple distant pixels in the image
  - Hessian is not diagonal (or even diagonally dominant)
- "One step threshold methods": Lucy-Richardson Algorithm
  - Known to be insufficient for decades in RA
- A-priori information used:
  - Sky does not look like the PSF
  - Limited support of the emission, positivity
- Many of the techniques/issues/algorithms described appear rather similar.
  - Re-invention? Just language differences? Or real differences not understood?