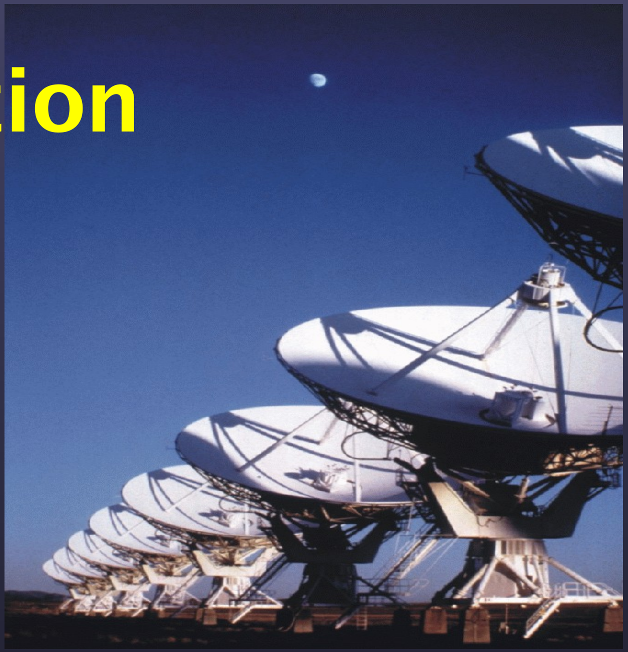
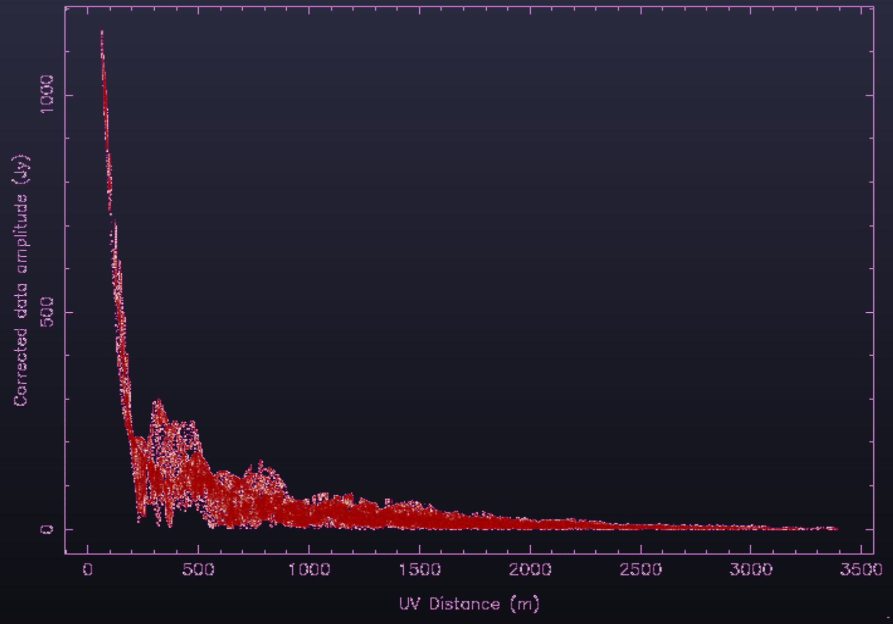
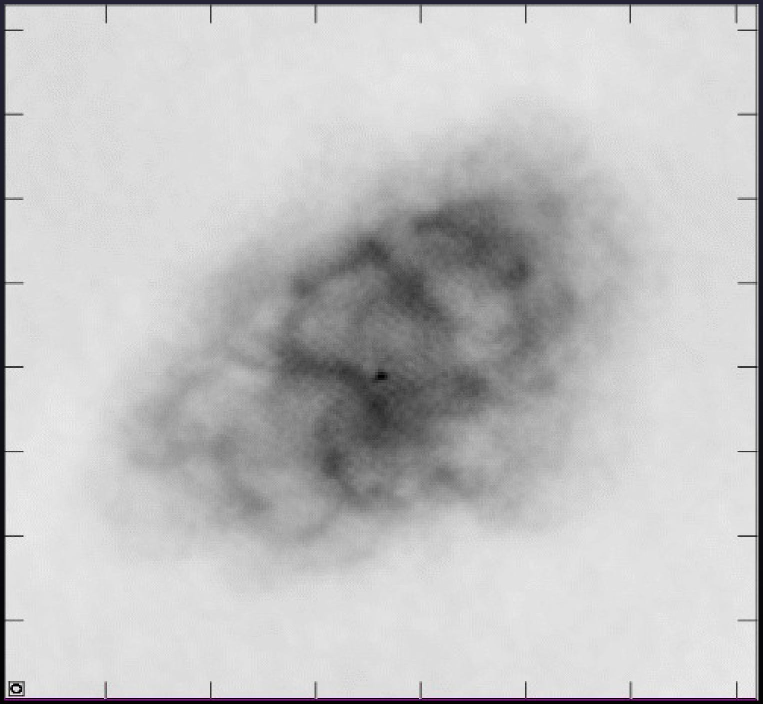
Imaging and deconvolution

S. Bhatnagar, NRAO







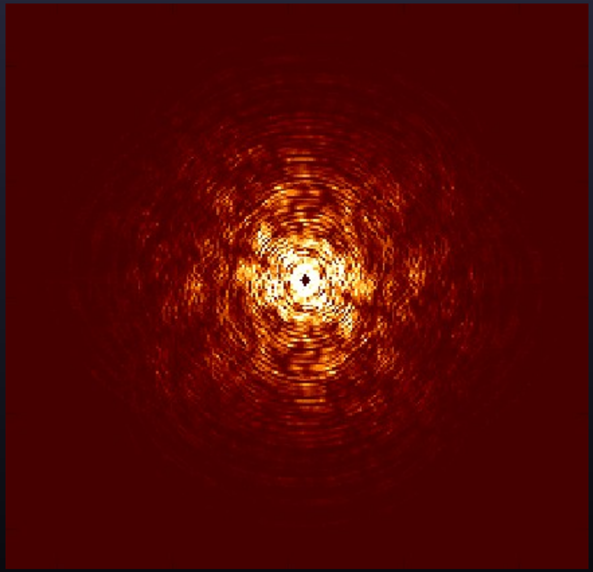
Tenth Synthesis Imaging Summer School

University of New Mexico, June 13- 20, 2006

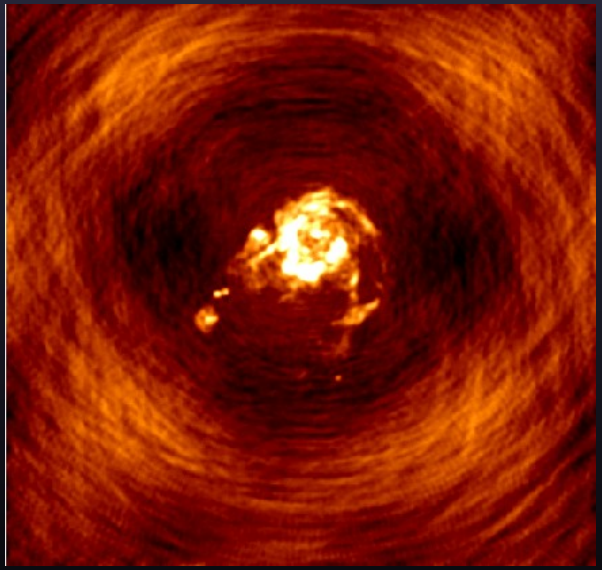
Plan for the lecture–I

- How do we go from the measurement of the coherence function (the Visibilities) to the images of the sky?
- First half of the lecture: Imaging

Measured Visibilities <--> Dirty Image



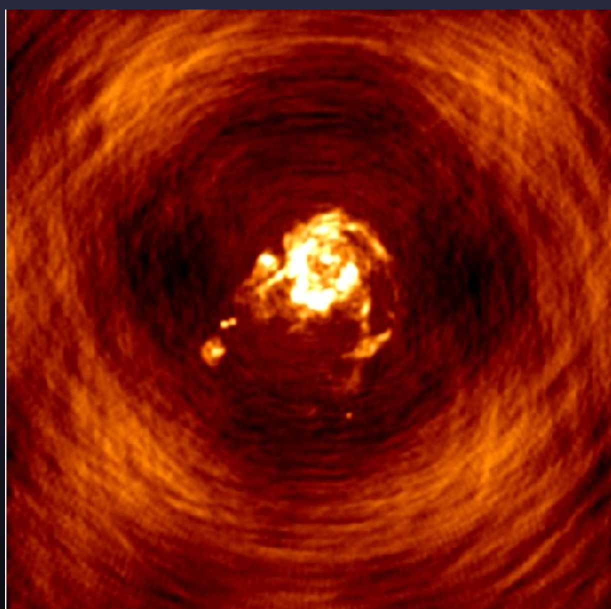
<-->



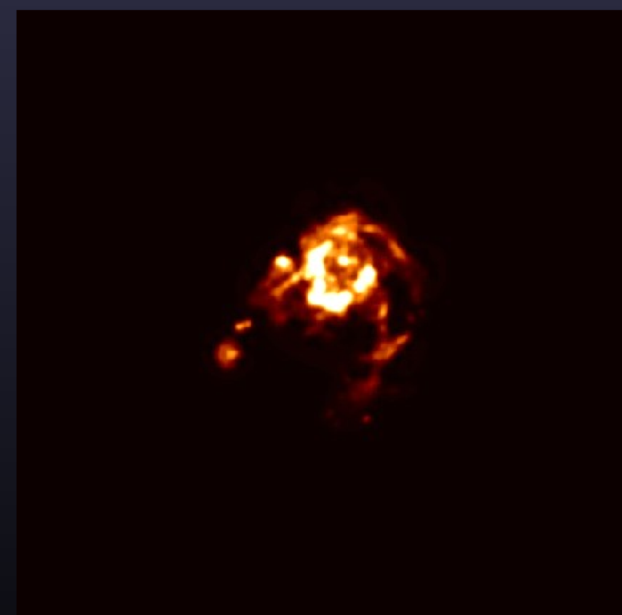
Plan for the lecture–II

- **Second half of the lecture: Deconvolution**

Dirty image <--> Model of the sky



<-->



Imaging

- **Interferometers are indirect imaging devices**

$$V^o(u,v,w)=\iint P(l,m)\,I(l,m)\,e^{2\pi\iota\left[ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right]}\frac{dl\,dm}{\sqrt{1-l^2-m^2}}$$

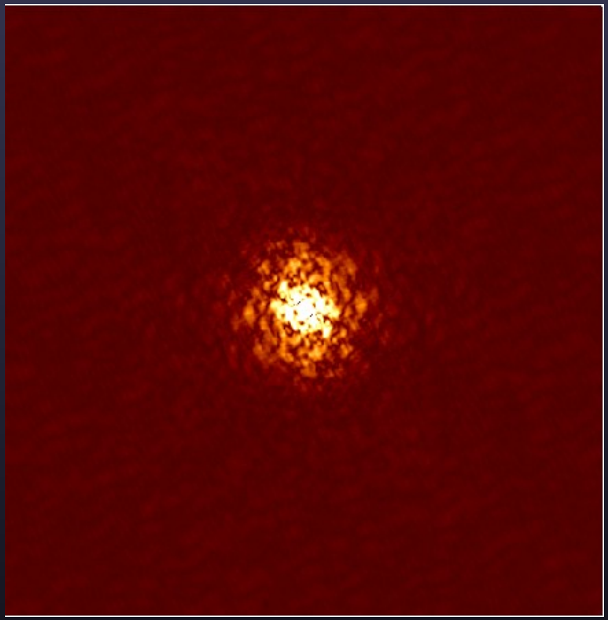
- **For small w (small max. baseline) or small field of view *(l^2 + m^2 << 1)* and *P(l,m)*~1, *I(l,m)* is 2D Fourier transform of *V(u,v)***

$$V^o(u,v)=\iint I(l,m)\,e^{2\pi\iota\left[ul+vm\right]}dl\,dm$$

$$V^o(u,v)\Leftrightarrow I(l,m)$$

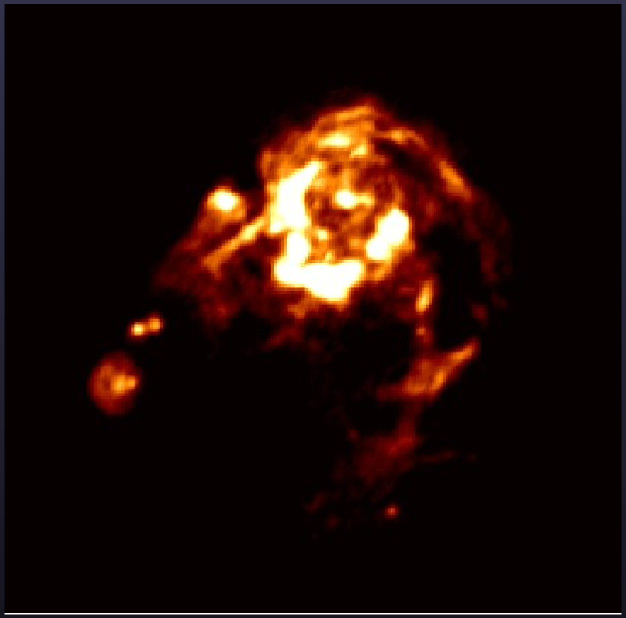
Imaging: Ideal 2D Fourier relationship

Ideal visibilities(*V*) — **True image(*I*)**



FT

<-->



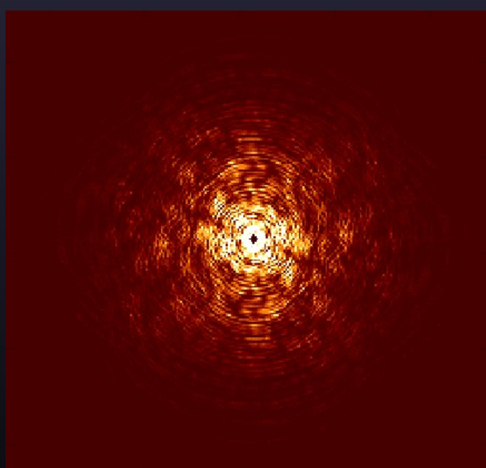
- **This is true ONLY if *V* is measured for all *(u,v)*!**

Imaging: UV-plane sampling

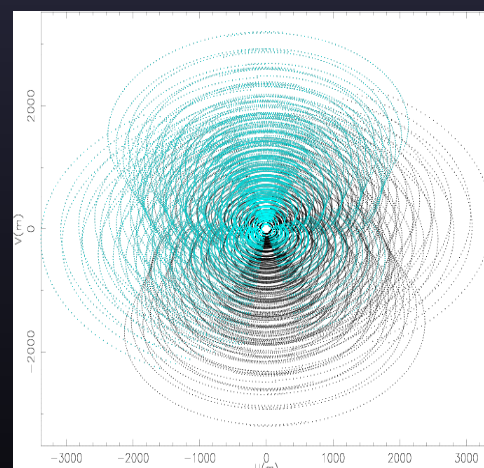
- **With limited number of antennas, the uv-plane is sampled at discrete points:**

$$S(u,v)=\sum_k \delta(u_k,v_k)$$

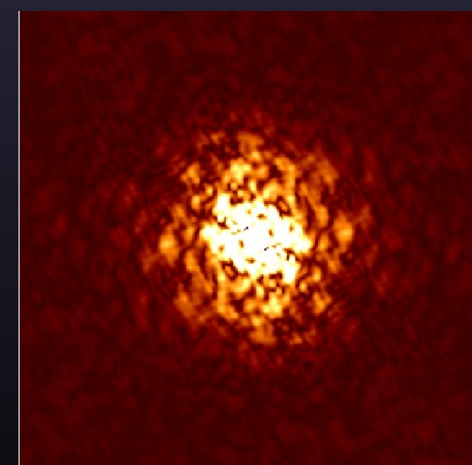
$$V^M(u,v)=S(u,v)\,V^o(u,v)$$



=



X



V^M S V^o

Convolution with the PSF

- **Effect of sampling the uv-plane:**

$$I^d(l,m) = FT^{-1}\left[V^o(u,v)\,S(u,v)\right]$$

- **Using the Convolution Theorem:**

$$I^d(l,m) = B(l,m) * I^o(l,m)$$

The Dirty Image (I^d) is the convolution of the True Image (I^o) and the Dirty Beam/Point Spread Function (B)

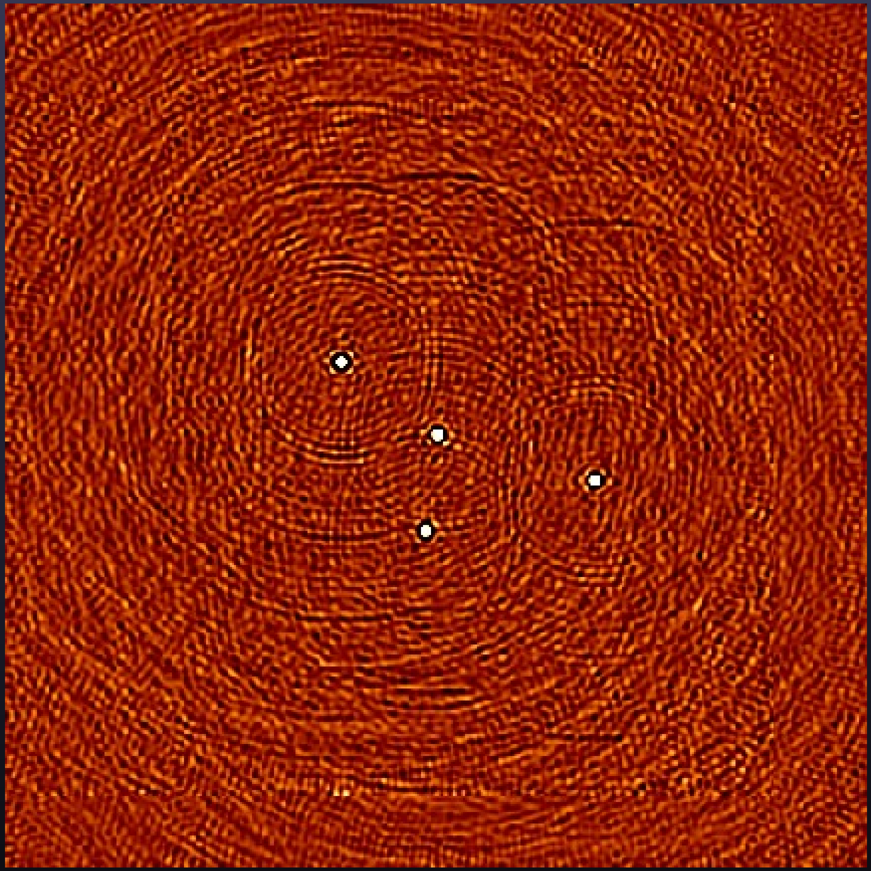
$$B = FT^{-1}(S)$$

- **In practice**

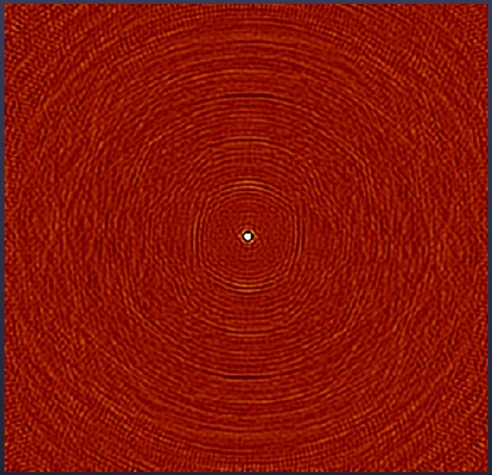
$$I^d = B * I^o + B * I^N \quad \text{where} \quad I^N = FT^{-1}(\text{Vis. Noise})$$

- **To recover I^o, we must deconvolve B from I^d. The algorithm must also separate $B*I^o$ from $B*I^N$.**

Convolution



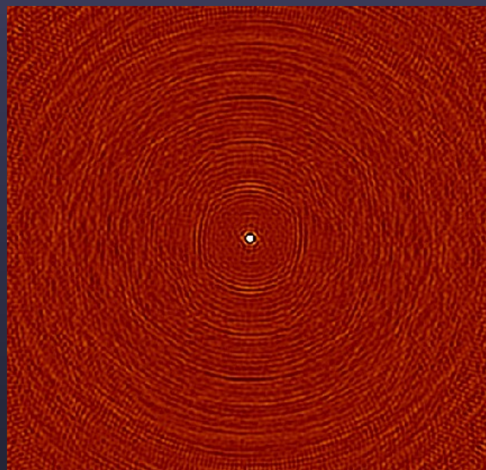
The PSF



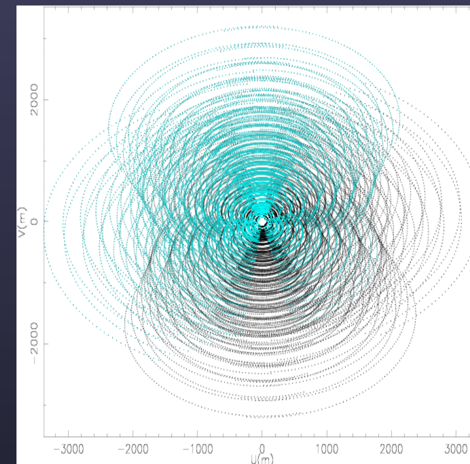
$$= I(x_0)B(x-x_o) + I(x_1)B(x-x_1) + \ldots$$

The Dirty Image

The PSF

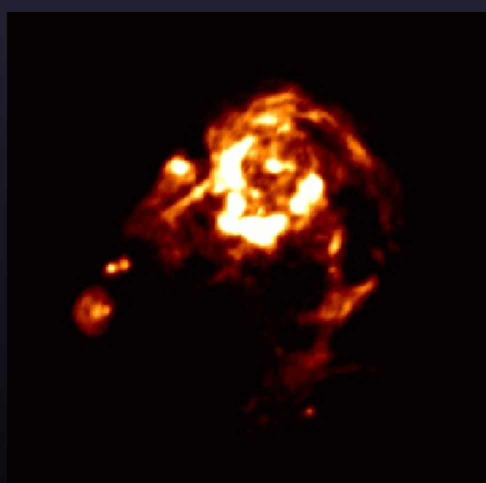


FT
<-->

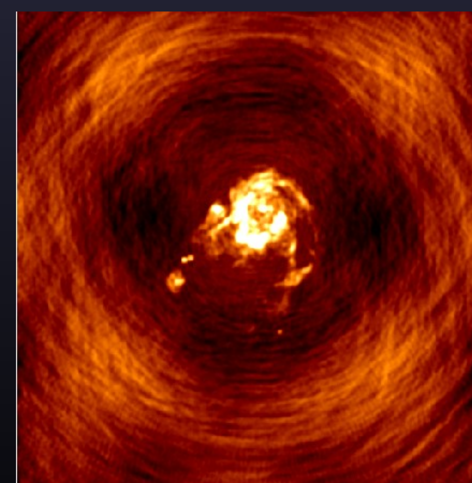


UV-coverage

*



-->



The
Dirty Image

Making of the Dirty Image

- **Fast Fourier Transform (FFT) is used for efficient Fourier transformation. It however requires regularly spaced grid of data.**
- **Measured visibilities are irregularly sampled (along uv-tracks).**
- **Convolutional gridding is used to effectively interpolate the visibilities everywhere and then re-sample them on a regular grid (the Gridding operation)**

$$V^S = V^M * C = (V^o S) * C \;\; => I^d \cdot FT^{-1}(C)$$

- ***C* is designed to have desirable properties in the image domain.**

Dirty Beam: Interesting properties

- **PSF is a weighted sum of cosines corresponding to the measured Fourier components:**

$$B(l,m) = \frac{\sum_k W_k \cos(u_{kl} + v_{km})}{\sum_k W_k}$$

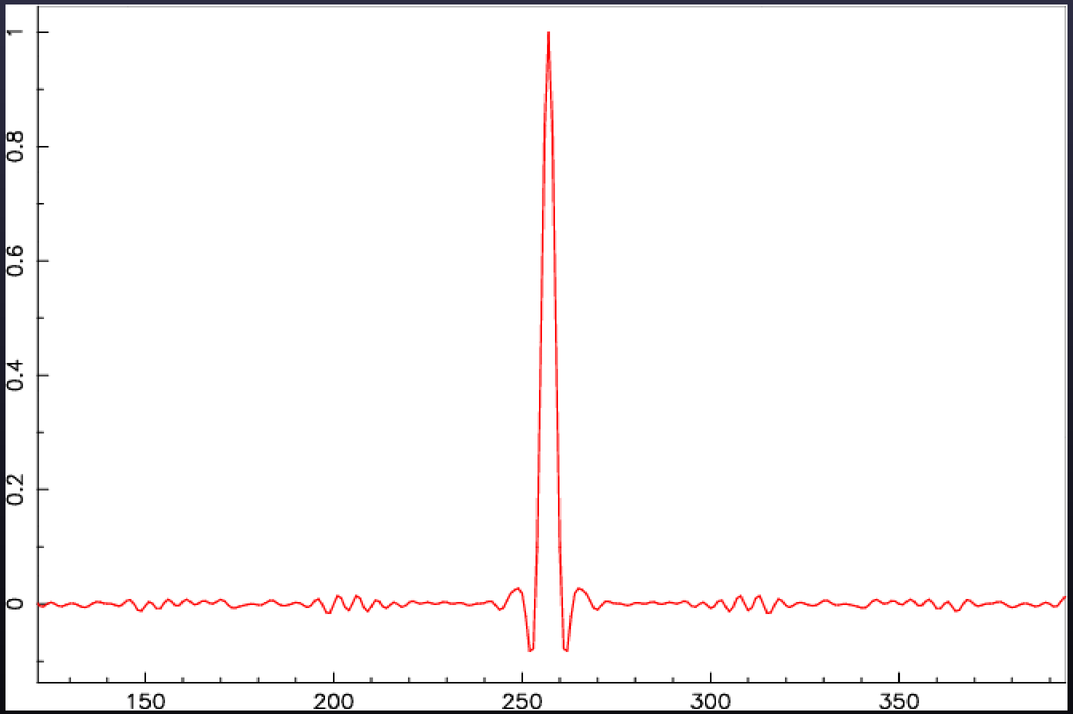
Visibility weights (w_i) are also gridded on a regular grid and FFT used to compute the Dirty Beam (or the PSF).

- **The peak of the PSF is normalized to 1.0**
- **The 'main lobe' has a size d$x \sim 1/u_{max}$ and d$y \sim 1/v_{max}$**

This is the 'diffraction limited' resolution (the Clean Beam) of the telescope.

Dirty Beam: Interesting properties

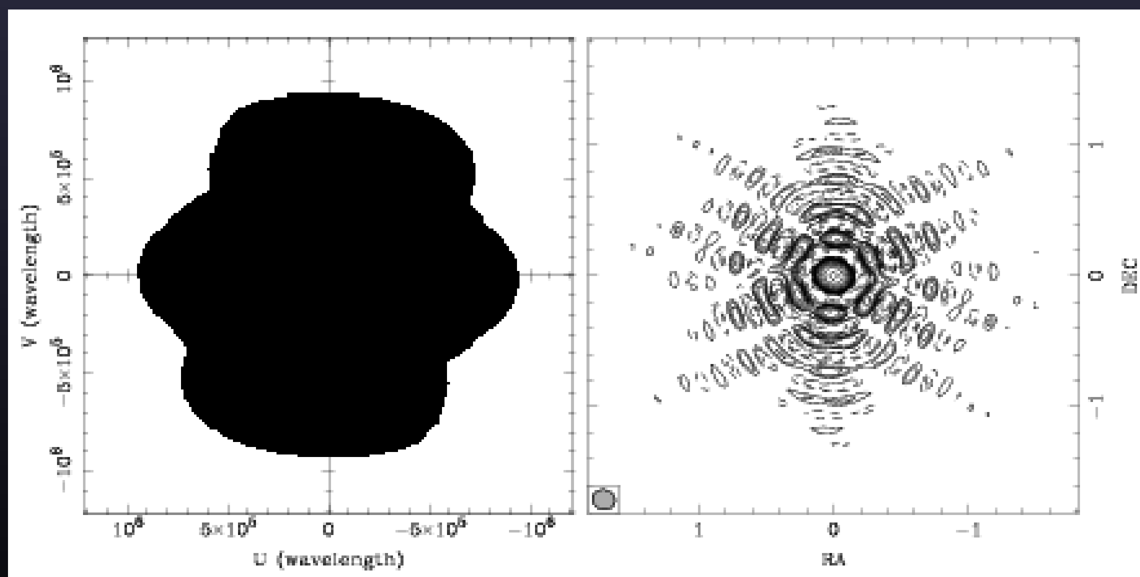
- **Side lobes extend indefinitely**
- **RMS ~ *1/N*   where *N* = No. of antennas**



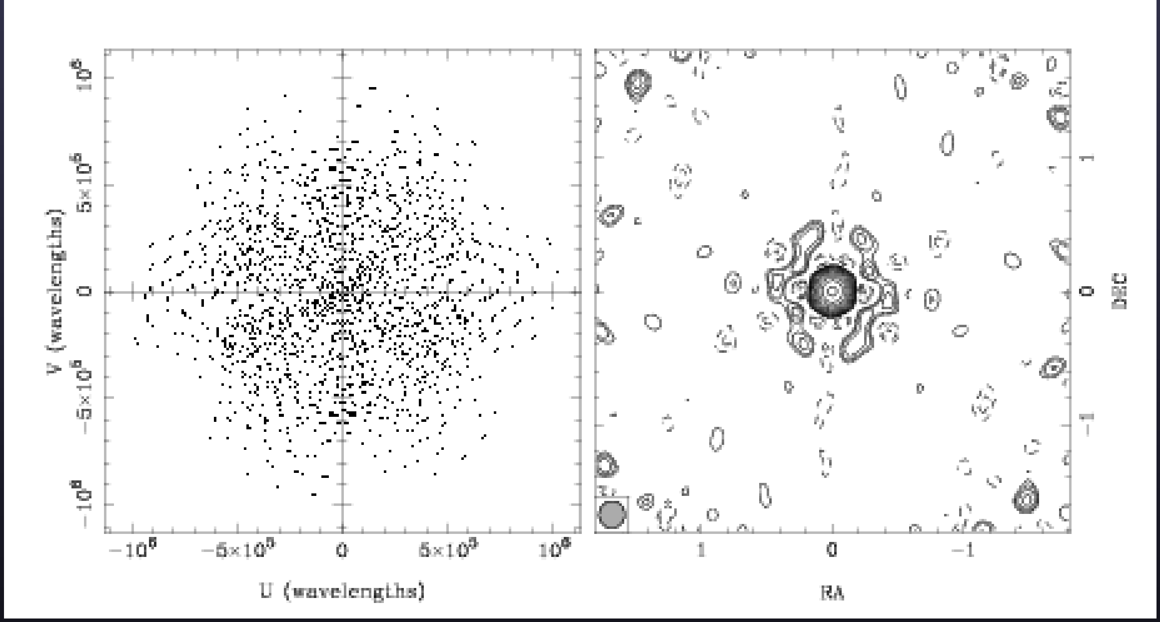
Close-in side lobes of the PSF

- **Close-in side lobs of the PSF are controlled by the uv-coverage envelope.**

  E.g., if the envelop is a circle, the side lobes near the main lobe must be similar to the FT of a circle: Bessel function/Radius



Close-in side lobes: VLA uv-coverage



PSF forming: Weighting...

- Weighting function (W_k) can be chosen to modify the side lobes

$$B(l,m)=\frac{\sum_k W_k \cos(u_{kl}+v_{km})}{\sum_k W_k}$$

- **Natural Weighting**

$W_k=1/\sigma_k^2$ where σ_k^2 is the RMS noise

- Best RMS across the image.
- Large scales (smaller baselines) have higher weights.
- Effective resolution less than the inverse of the longest baseline.

...Weighting...

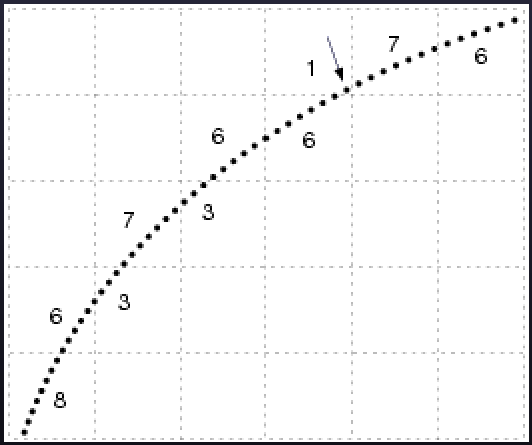
- **Uniform weighting**

  $W_k = 1/\rho(u_k, v_k)$ **where $\rho(u_k, v_k)$ is the density of uv-points in the k^{th} cell.**

- **Short baselines (large scale features in the image) are weighted down.**

- **Relatively better resolution**

– **Increases the RMS noise.**



- **Super uniform weighting:**

  Consider density over larger region.

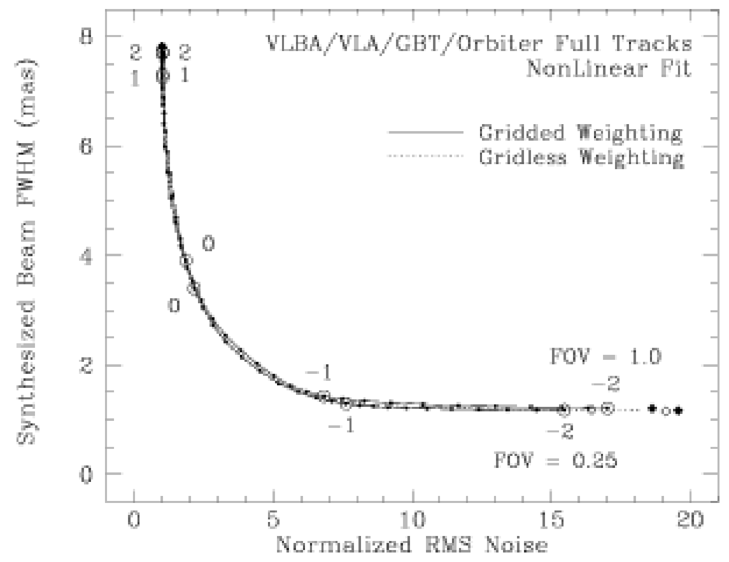
  Minimize side lobes locally.

...Weighting

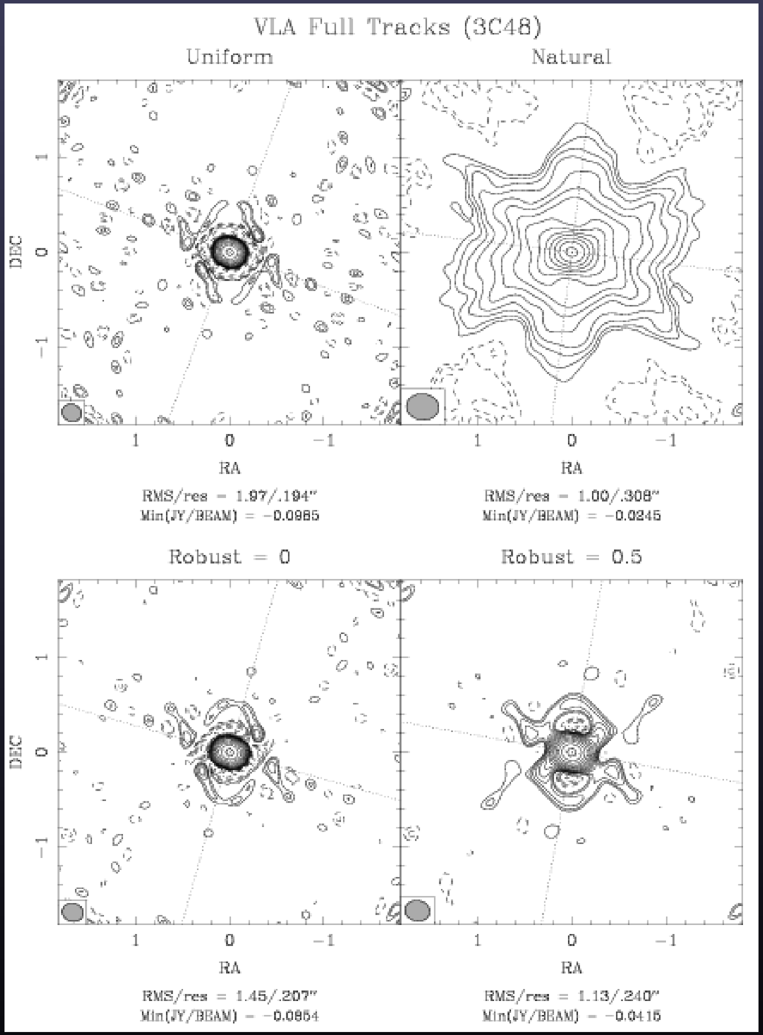
- **Robust/Briggs weighting:**

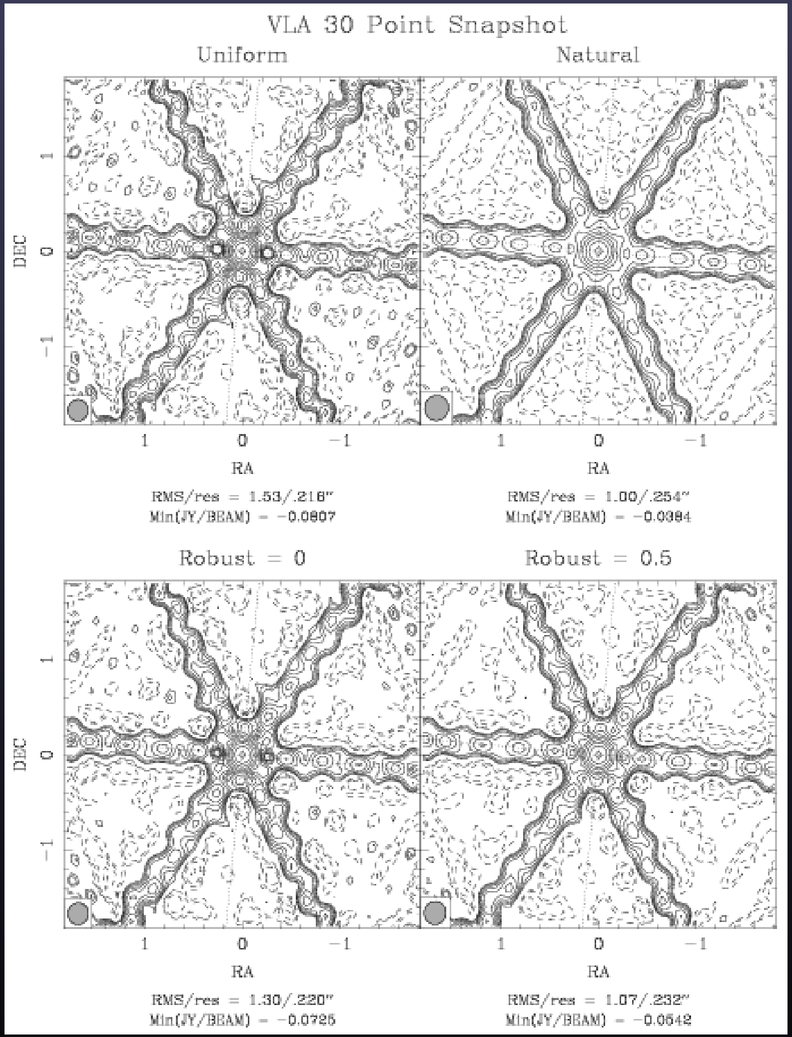
$$W_k = 1/[S\,\rho(u_k, v_k) + \sigma_k^2]$$

- **Parameterized filter – allows continuous variation between optimal resolution (uniform weighting) and optimal noise (natural weighting).**



Examples of weighting





PSF Forming: Tapering

- **The PSF can be further controlled by applying a tapering function on the weights (e.g. such that the weights smoothly go to zero beyond the maximum baseline).**

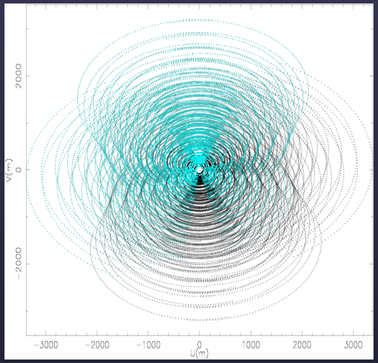
$$W'_k = T(u_k, v_k)\, W_k(u_k, v_k)$$

- **Bottom line on weighting/tapering:**

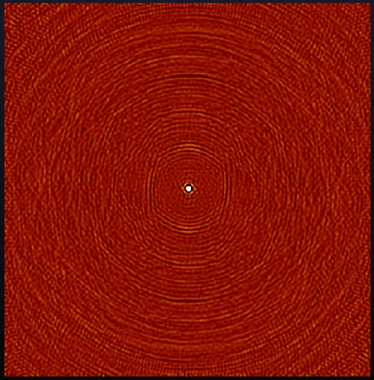
These help a bit, but imaging quality is limited by the deconvolution process!

The missing information

- As seen earlier, not all parts of the uv-plane are sampled – the 'invisible distribution'



1. "Central hole" below u_{min} and v_{min}:
   - Image plane effect: Total integrated power is not measured.
   - Upper limit on the largest scale in the image plane.
2. No measurements beyond u_{max} and v_{max}:
   - Size of the main lobe of the PSF is finite (finite resolution).
3. Holes in the uv-plane:
   - Contribute to the side lobes of the PSF.



More on missing information

- Missing 'central hole' means that the total flux, integrated over the entire image is zero.

$$V(u=0, v=0) = \iint I^d(l, m)\, dl\, dm = 0$$

- Total flux for scales corresponding to the Fourier components between u_{max} and u_{min} can be measured.
  - In the presence of extended emission, the observations must be designed keeping in mind:
    - the required resolution ==> maximum baseline
    - the largest scale to be reliably reconstructed ==> minimum baseline

Recovering the missing information

- **For information beyond the max. baseline, one requires extrapolation. That's unphysical (unconstrained).**
- **Information corresponding to the "central hole": possible, but difficult (need extra information).**
- **Information corresponding to the uv-holes: requires interpolation. The measurements provide constraints – hence possible. But non-linear methods necessary.**

If *Z* is the unmeasured distribution, then $B*Z=0$. If I^M is a solution to $I^d=B*I^M$, then so is $I^M + \alpha Z$ for any value of α.

Deconvolution = interpolation in the visibility plane.

Prior knowledge about the sky

- **What can we assume about the sky emission:**
  1. ***Sky does not look like cosine waves***
  2. ***Sky brightness is positive (but there are exceptions)***
  3. ***Sky is a collection of point sources (weak assertion)***
  4. ***Sky could be smooth***
  5. ***Sky is mostly blank (sometimes justifies "boxed" deconvolution)***
- **Non-linear deconvolution algorithms search for a model image I^M such that the residual visibilities $V^R=V^o-V^M$ are minimized, subject to the constraints given by the (assumed) prior knowledge.**

Small digression: Vector notation

- **Let**

  A **= Measurement matrix to go from the image domain to the visibility domain (the measurement domain).**

  I **= Vector of the image pixel values**

  V **= Vector of visibilities**

  B **= Operator (matrix) for convolution with the PSF**

  N **= The noise vector**

- **Then,**

$$I^d = BI^o + BI^N \quad \text{where } BI^N = A^T AN$$

$$V^M = AI^M \quad \text{and} \quad V^o = AI^o + N$$

$$V^R = V^o - AI^M$$

Some observations

- A is rectangular (not square) and is a collection of sines and cosines corresponding to only the measured Fourier components.
  - A is singular ==> A^{-1} does not exist
  - $I^M = A^{-1} V^M$ not possible ==> non-linear methods needed
- N is independent gaussian random process. Noise in the image domain = BI^N
  - Pixel-to-pixel noise in the image is not independent

...more observations

- For successful recovery of I^o given I^d, prior knowledge must fundamentally separate *BI^o and BI^N.*
- χ^2 is the optimal estimator. Deconvolution then is equivalent to:

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

Deconvolution is equivalent to function minimization

- Algorithms differ in the parameterization of P_k, the type of constraints and the way the constraints are applied.

Deconvolution algorithms

- **Scale-less algorithms:**

$$P_k = A_k \delta(x - x_k)$$

Popular ones: Clean, MEM and their variants

- **Scale-sensitive algorithms (new turf!):**

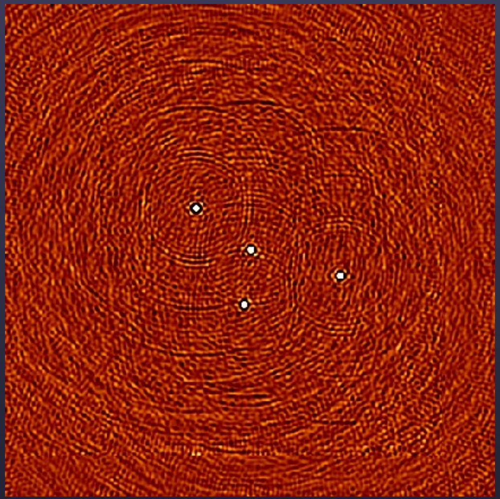
$$P_k = A_k f\left(Position, Scale\right)$$

Existing ones: Multi-scale Clean, Asp-Clean

- **Image plane corrections (in use)**

Existing ones: w-projection, pb-projection

The classic Clean algorithm (Hogbom, 1974)



- **Prior knowledge about the sky:**
  - **is composed of point sources**
  - **is mostly blank**

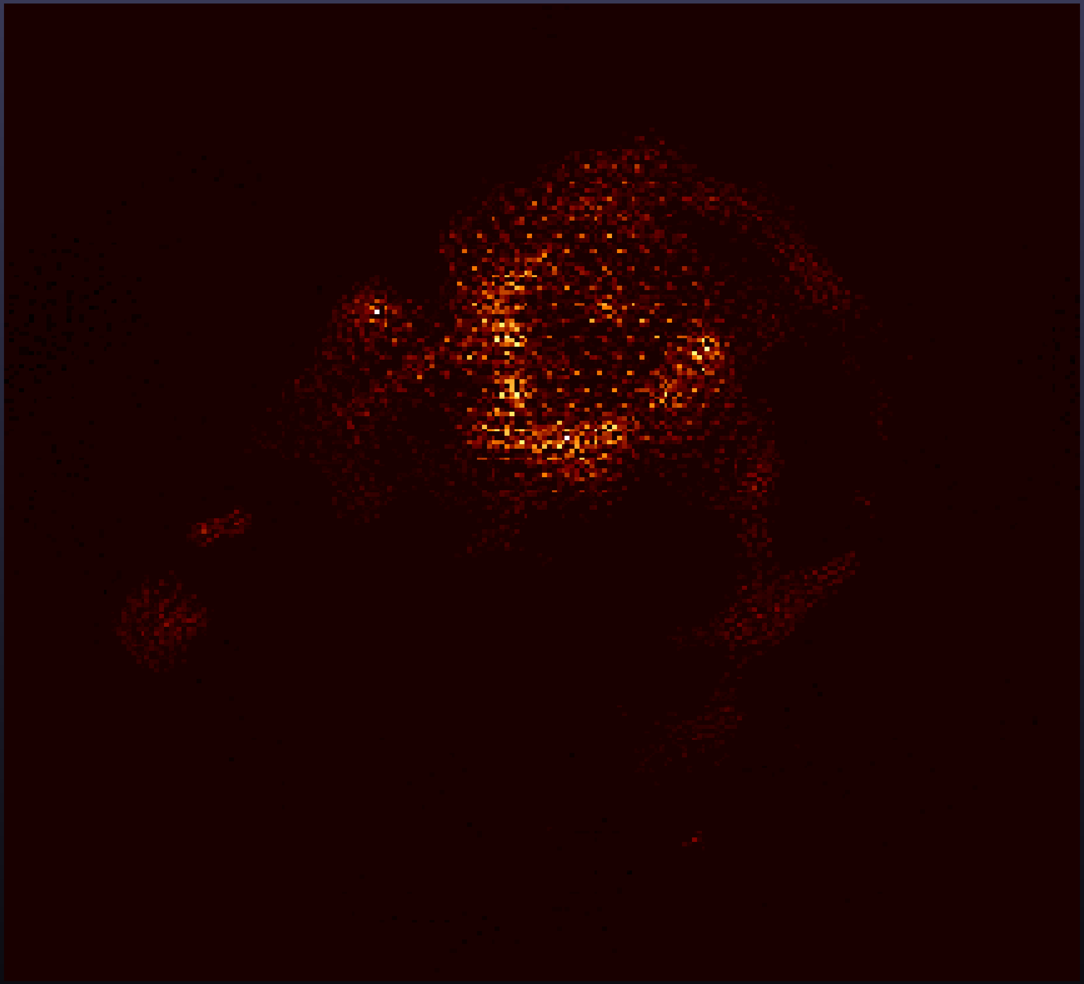
- **Algorithm:**
  1. **Search for the peak in the dirty image.**
  2. **Add a fraction *g* (loop gain) of the peak value to I^M.**
  3. **Subtract a scaled version of the PSF from the position of the peak $I^R_{i+1} = I^R_i - g\, B\, max(I^R_i)$**
  4. **If residuals are not "noise like", goto 1.**
  5. **Smooth I^M by an estimate of the main lobe (the "clean beam") of the PSF and add the residuals to make the "restored image"**

Details of Clean

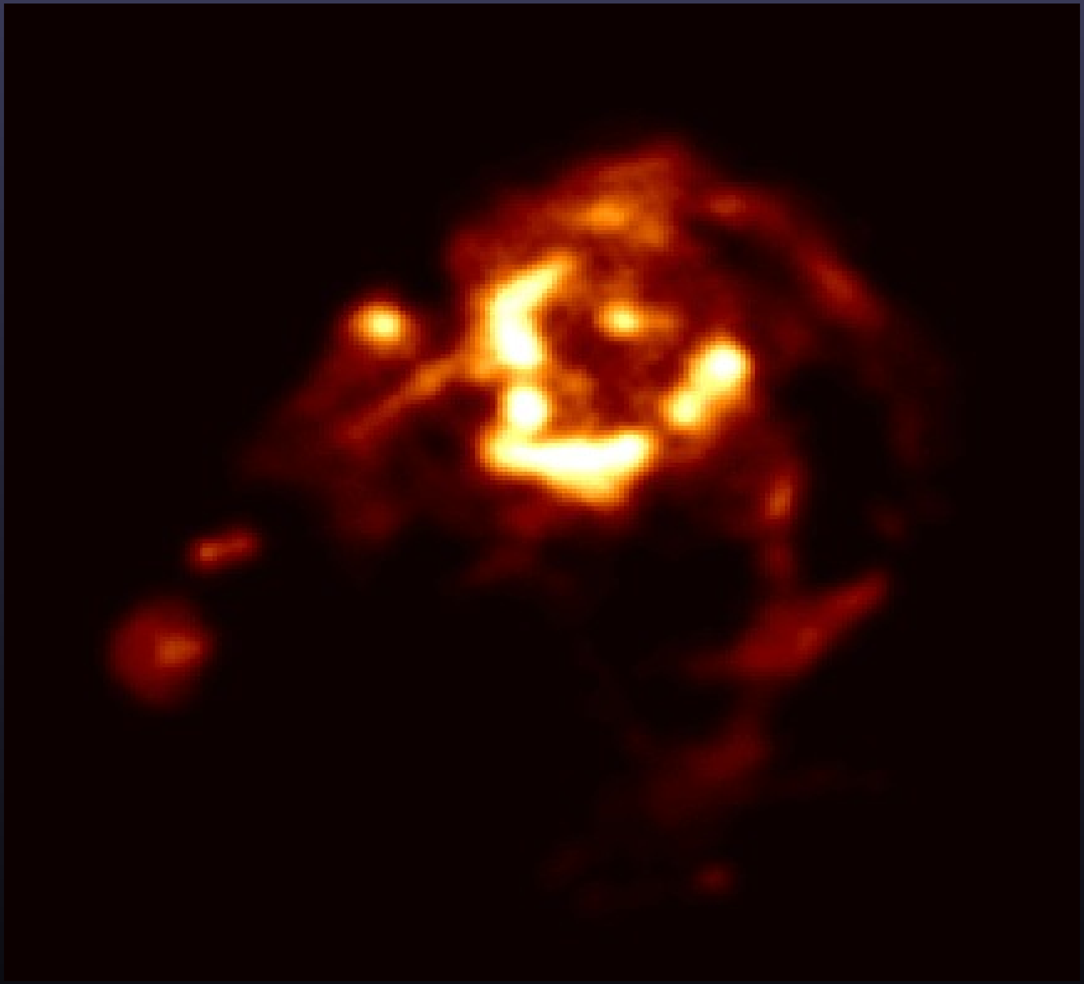
- It is a steepest descent minimization.
- Model image is a collection of delta functions – a scale insensitive algorithm.
- A least square fit of sinusoids to the visibilities which is proved to converge (Schwarz 1978).
- Stabilized by keeping a small loop gain (usually g=0.1–0.2).
- Stopping criteria: either the max. iterations or max. residuals some multiple of the expected peak noise.
- Search space constrained by user defined windows.

– Ignores coupling between pixels (extended emission) – assumes an orthogonal search space.

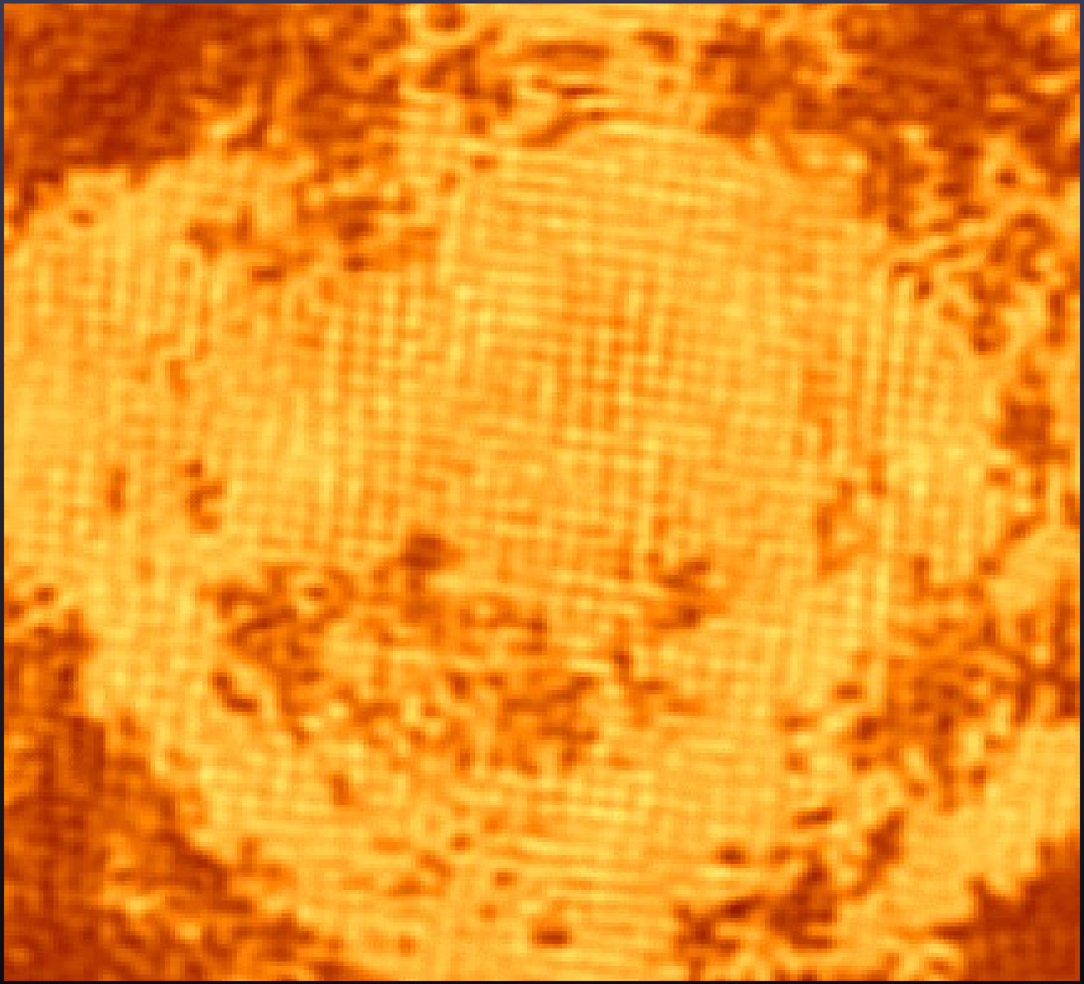
Clean: Model



Clean: Restored

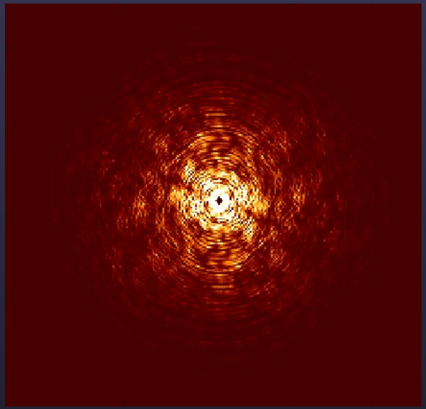


Clean: Residual

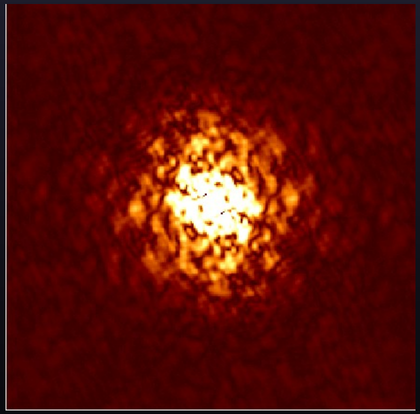


Clean: Model visibilities

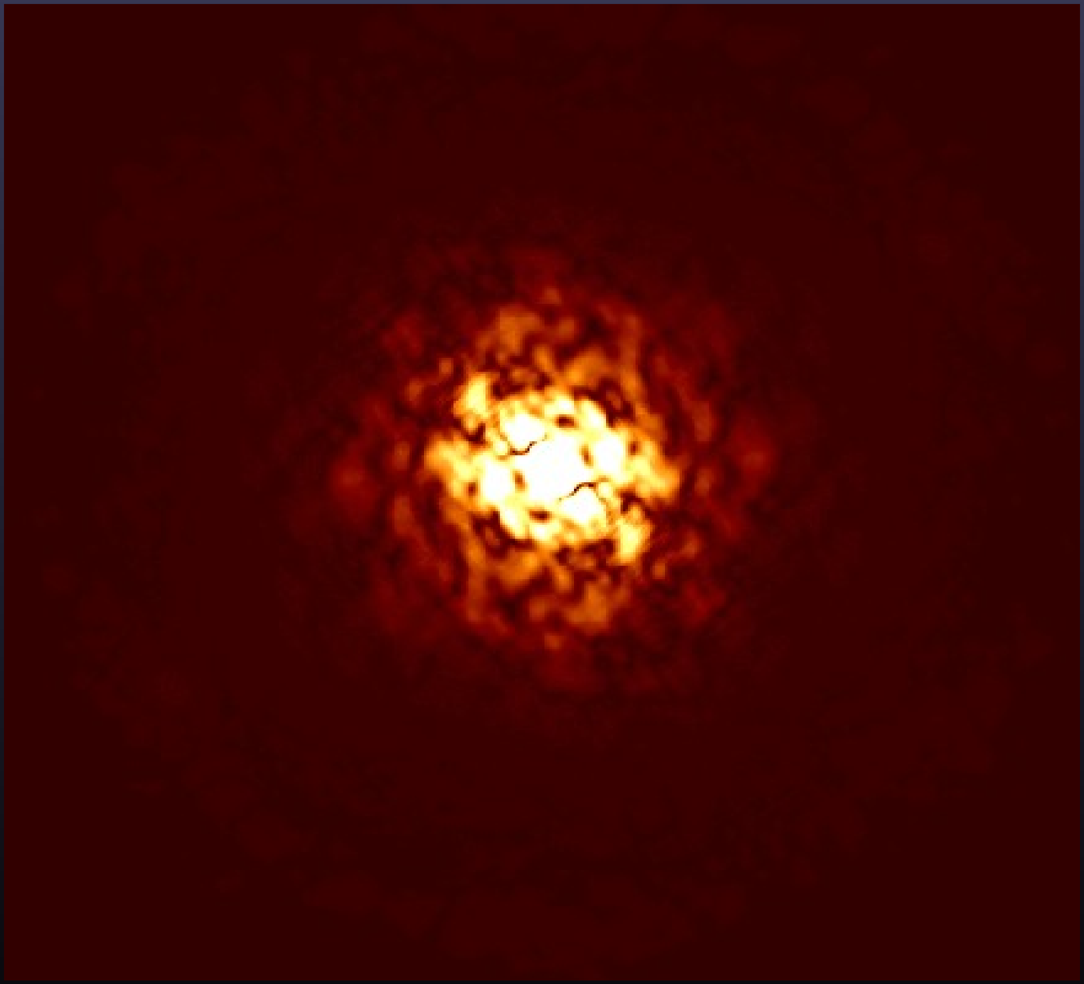
Sampled Vis



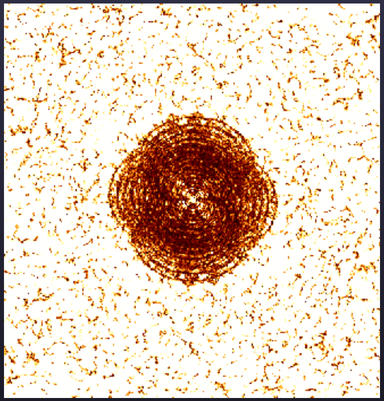
True Vis.



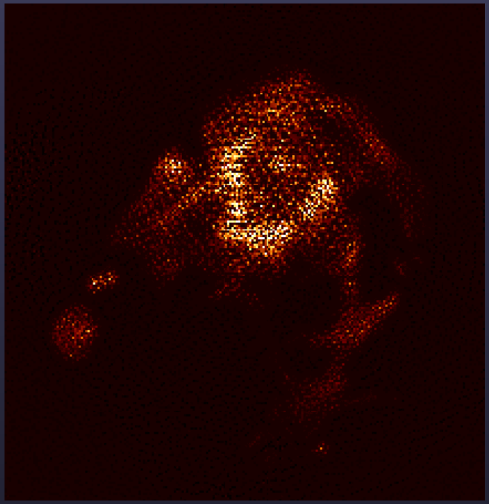
Model Vis.

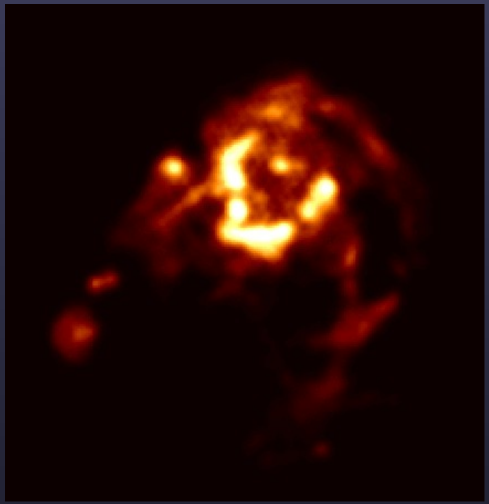


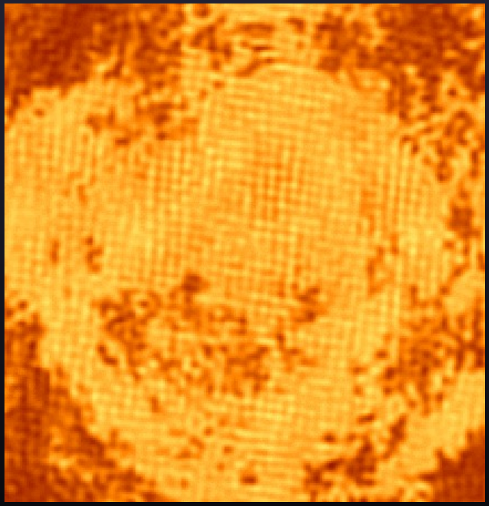
Residual Vis.

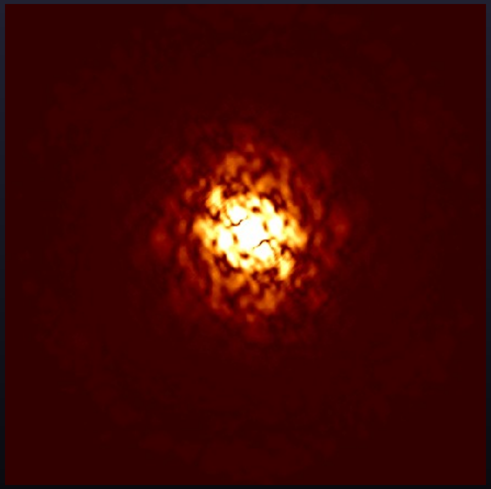


Clean: Example









Variants of the Classic Clean

- **Clark Clean – uses FFT to speed up**

  Minor cycle(inexpensive) : Clean the brightest points using an approximate PSF to gain speed

  Major cycle(expensive): Use FFT convolution to accurately remove the point sources found in the minor cycle

- **Cotton-Schwab Clean: A variant of Clark Clean**

  Subtract the point sources from the visibilities directly.

  Sometimes faster and always more accurate then Clark Clean.

  Easy to adapt for multiple fields.

Deconvolution algorithms: MEM

- **MEM is a constrained minimization algorithm.**
- **Fast non-linear optimization algorithm due to Cornwell&Evans(1983).**
- **Solve the convolution equation, with the constrain of smoothness via the 'entropy'**

$$H(I) = -\sum_k I_k - \log\left(I_k / m_k\right)$$

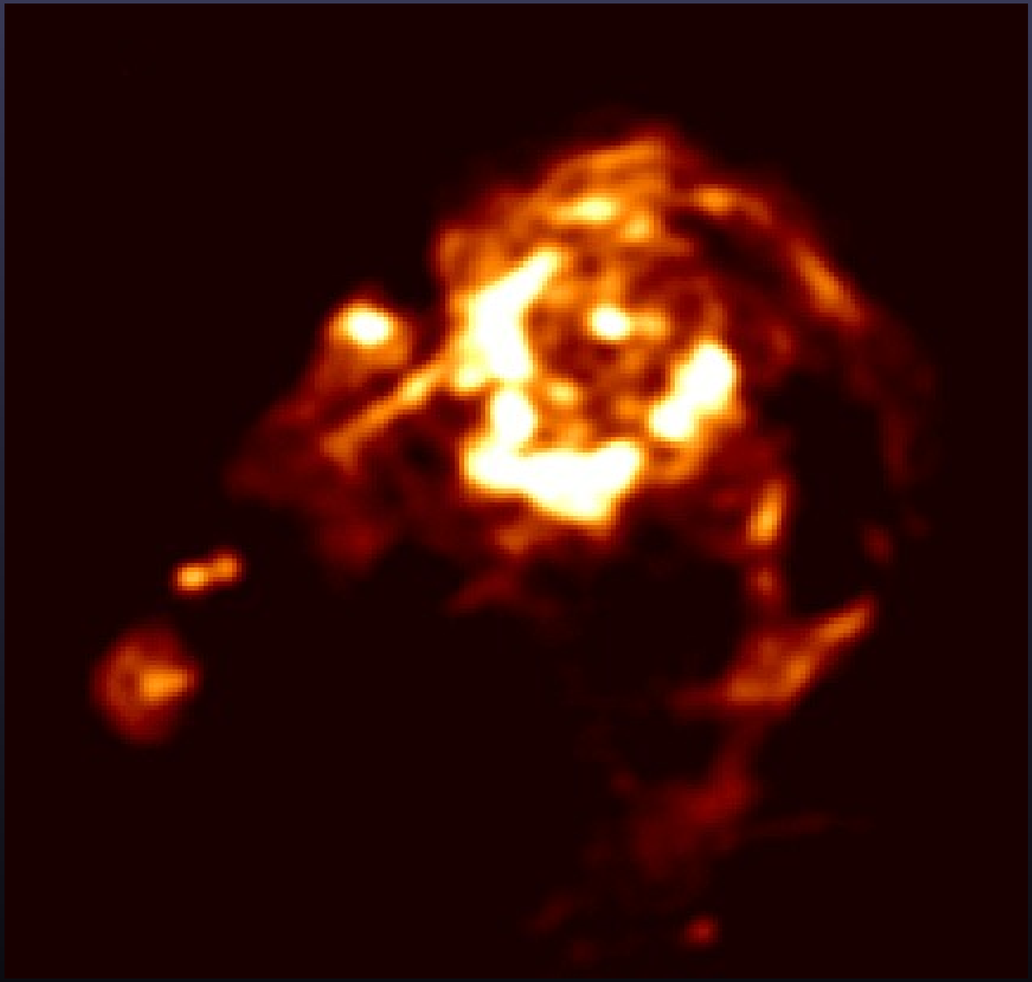
m_k is the prior image - usually a flat default image.

- **Default image is a very useful in incorporating model images from other algorithms etc.**
- **Naturally useful when final image is some combination of images (like mosaic images).**

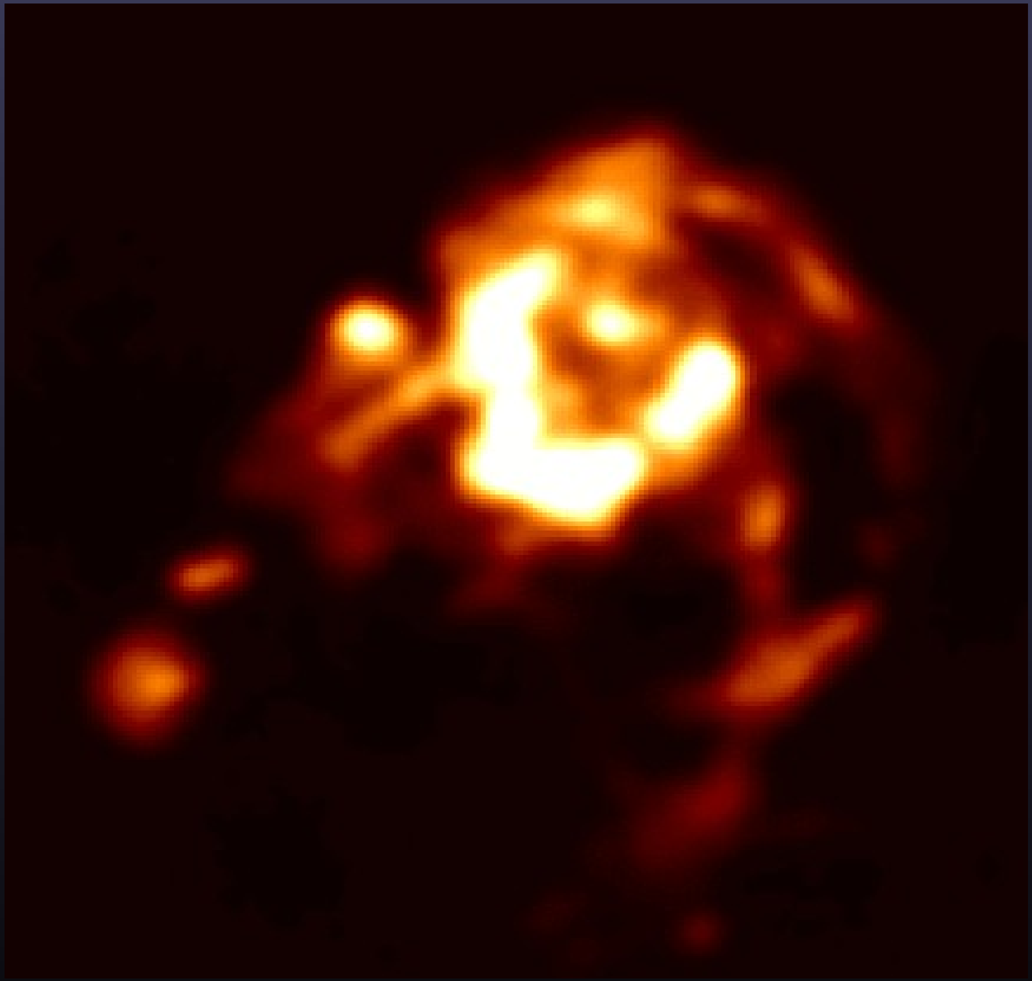
MEM: Some points

- Works better than Clean for extended emission.
- Every pixel is treated as a potential degree of freedom – a scale insensitive algorithm.
- Point sources are a problem, particularly along with large scale background emission – but can be removed with, say, Clean before hand.
- Easier to analyze and understand.

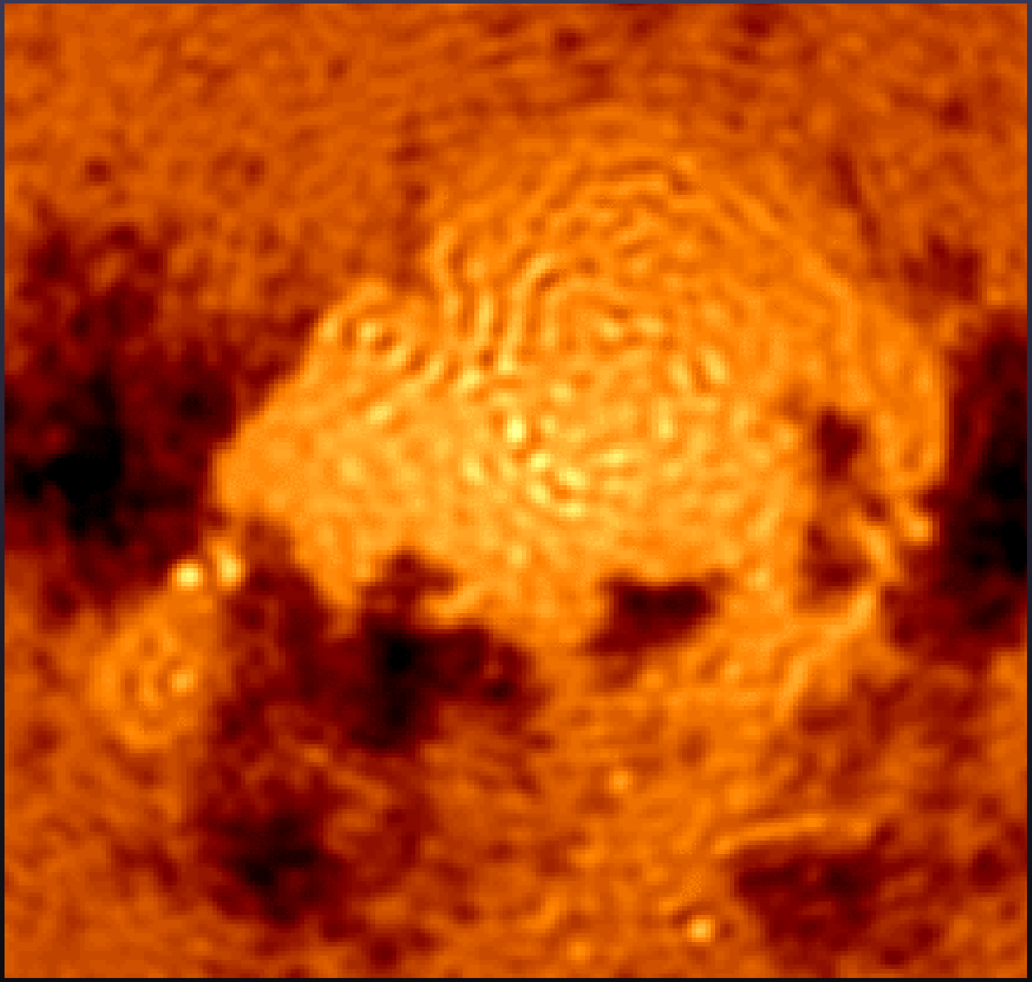
MEM: Model



MEM: Restored

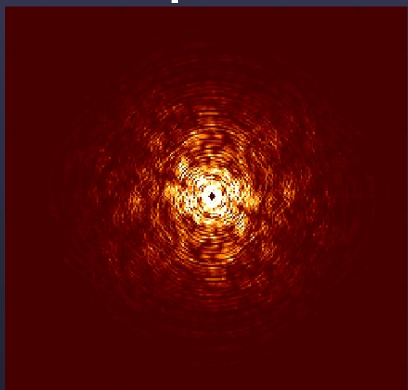


MEM: Residual

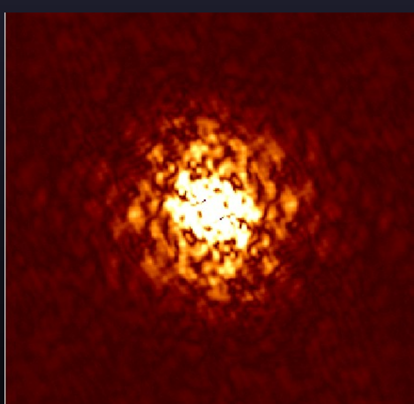


MEM: Model visibilities

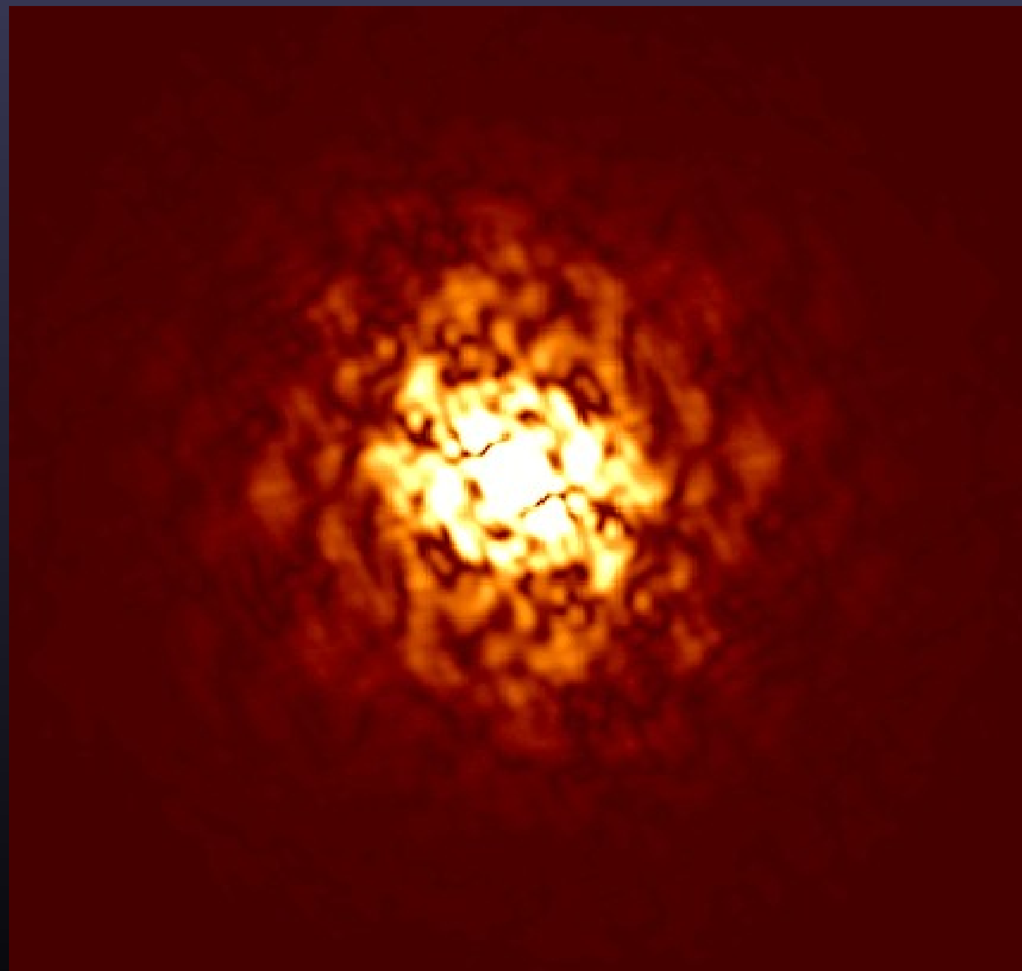
Sampled Vis



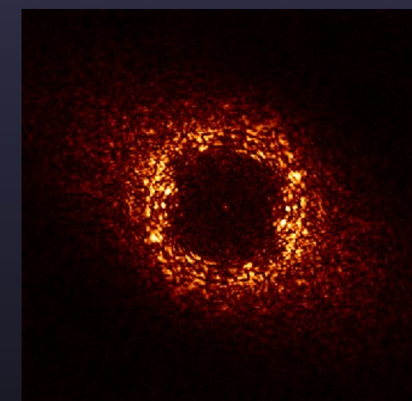
True Vis.



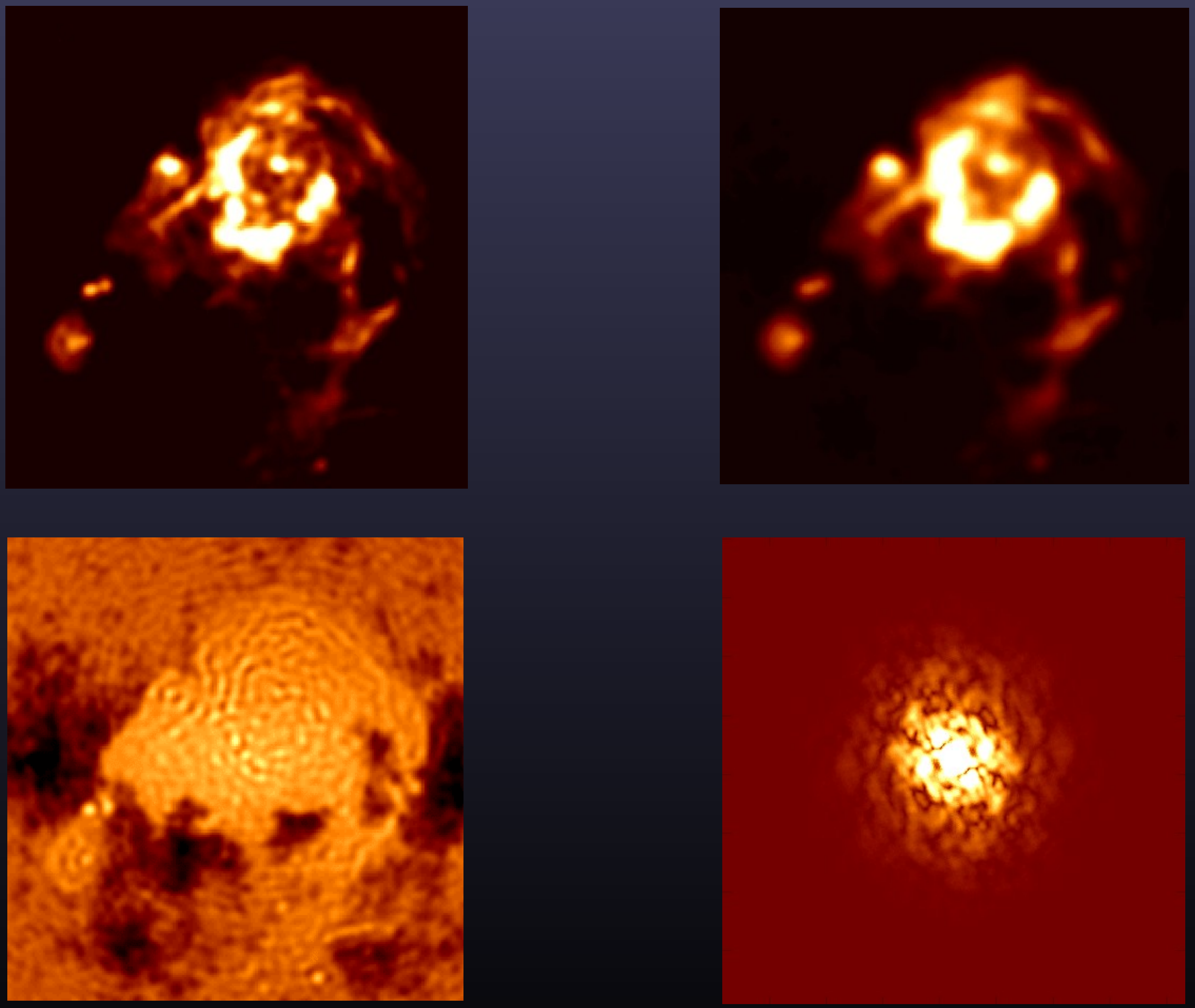
Model Vis.



Residual Vis.



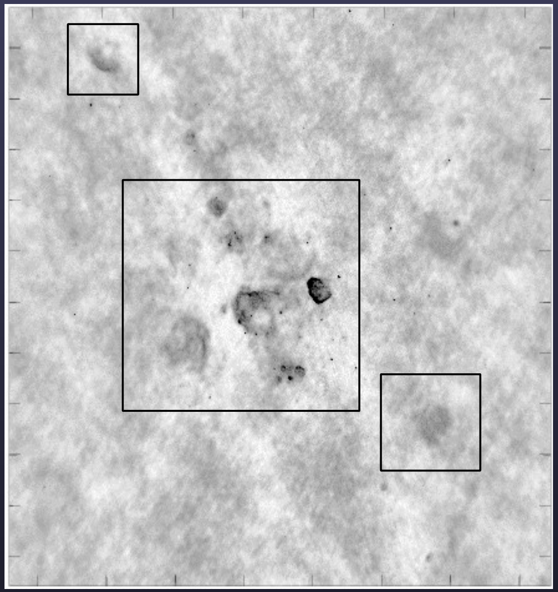
MEM: Example



Role of boxes

- Limit the search for components to only parts of the image.

  A way to regularize the deconvolution process.



- Useful when small no. of visibilities (e.g. VLBI/snapshots).
- Do not over-Clean within the boxes (over-fitting).
- Deeper Clean with no/loose boxes and lower loop gain can achieve similar (more objective) results.
- Stop when Cleaning within the boxes has no global effect (insignificant coupling of pixels due to the PSF).

Fundamental problem with scale-less decomposition

- **Each pixel is not an independent degree of freedom (DOF).**

    E.g., a gaussian shaped source covering 100 pixels can be represented by 5 parameters.

- **Clean/MEM treats each pixel within a clean-box as an independent degree of freedom.**
- **Scale fundamentally separates noise and signal.**
    - ***Largest*** **coherent scale in** BI^N ***~ the size of the resolution element.***
    - ***Physically plausible*** I^M ***is composed of scales >= the resolution element (smallest scale is of the size of the resolution element).***

- ***Scale-sensitive reconstruction therefore leaves more noise-like (uncorrelated) residuals.***

Scale Sensitive Deconvolution: MS-Clean

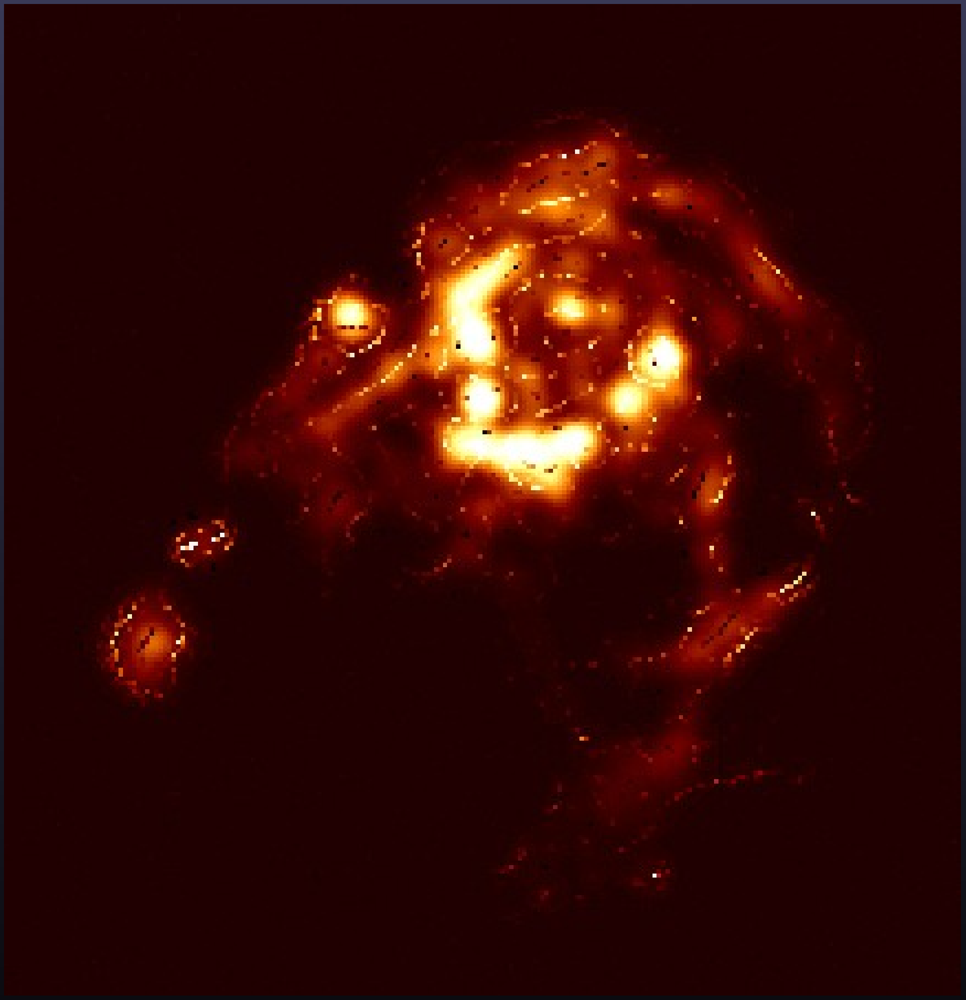
- Inspired by the Clean algorithm (Cornwell & Holdaway).
- Decompose the image into a pre-computed set of symmetric "blobs" at a few scales (e.g. Gaussians).

- Algorithm
  1. Make residual images smoothed to a few scales.
  2. Find the peak among these residual images.
  3. Subtract from all residual images a blob of scale corresponding to the scale of the residual image which had the peak.
  4. Add the blob to the model image.
  5. If more peaks in the residual images, goto 1.
  6. Smooth the model image by the "clean beam" and add the residuals.

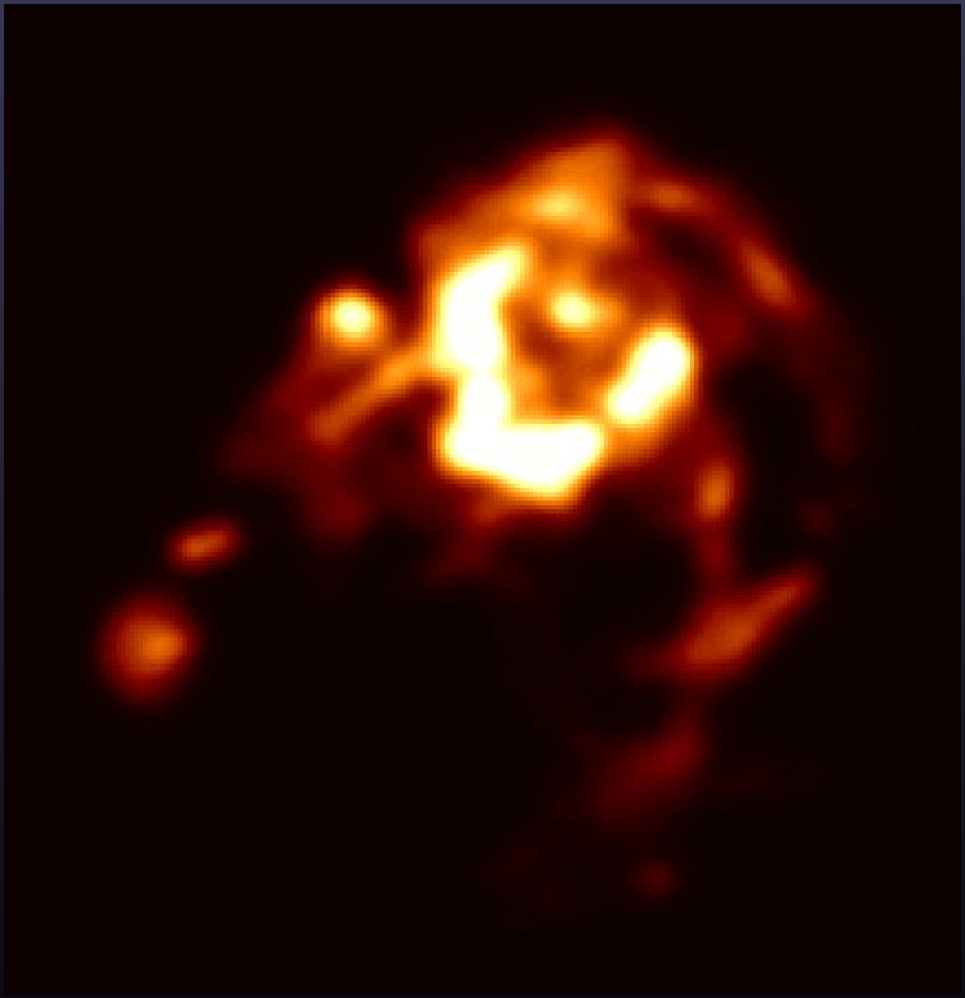
MS-Clean details

- ✔ **Deals with compact as well as extended emission better (need to include a blob of zero scale).**
- ✔ **Retains the scale-shift-n-subtract nature of Clean – easy to implement.**
- ✔ **Reasonably fast (for what it does!)**
- ✗ **Breaks up non-symmetric structures (as in Clean – but the errors are at larger scales than in Clean).**
- ✗ **Ignores coupling between blobs.**
  - **Assumes an orthogonal space and steepest descent minimization.**

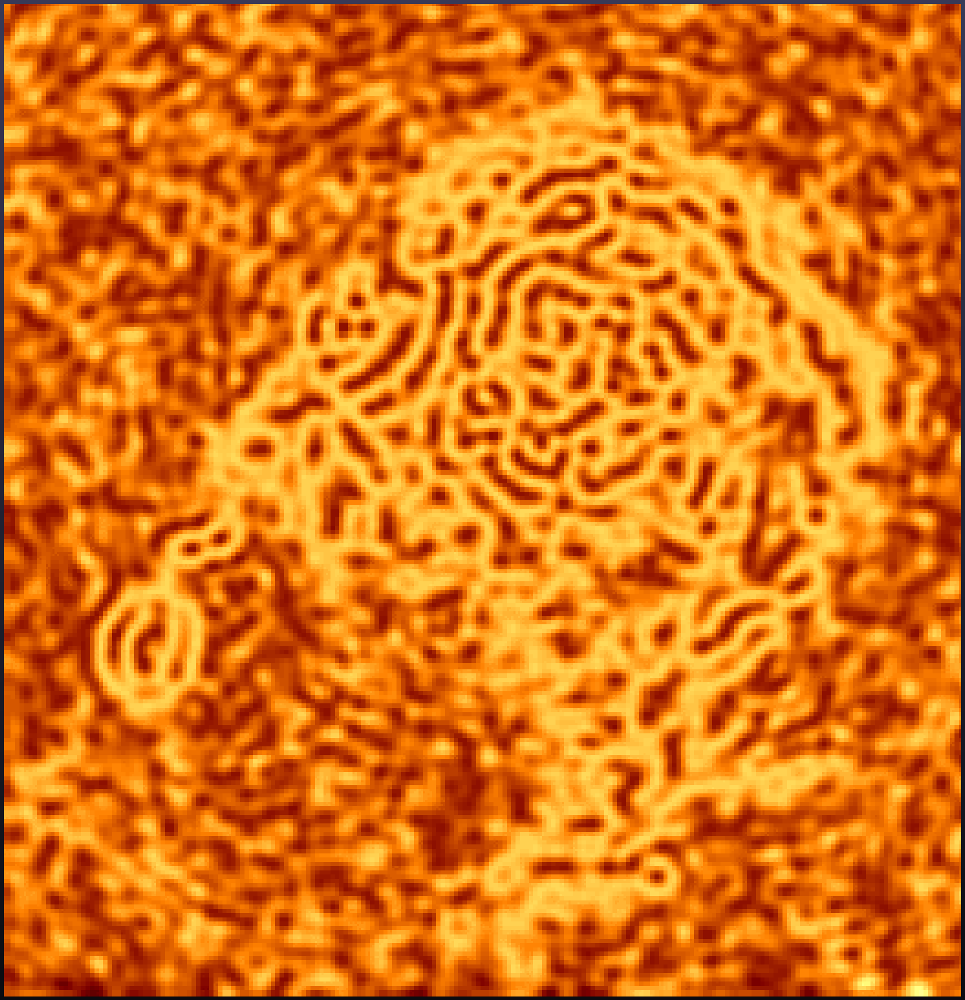
MS-Clean: Model



MS-Clean: Restored

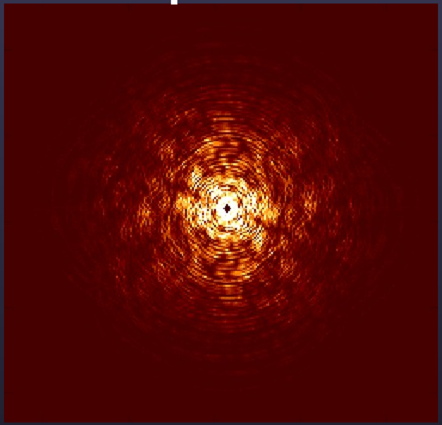


MS-Clean: Residual

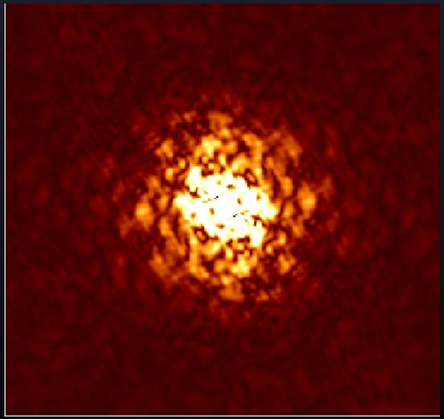


MS-Clean: Model visibilities

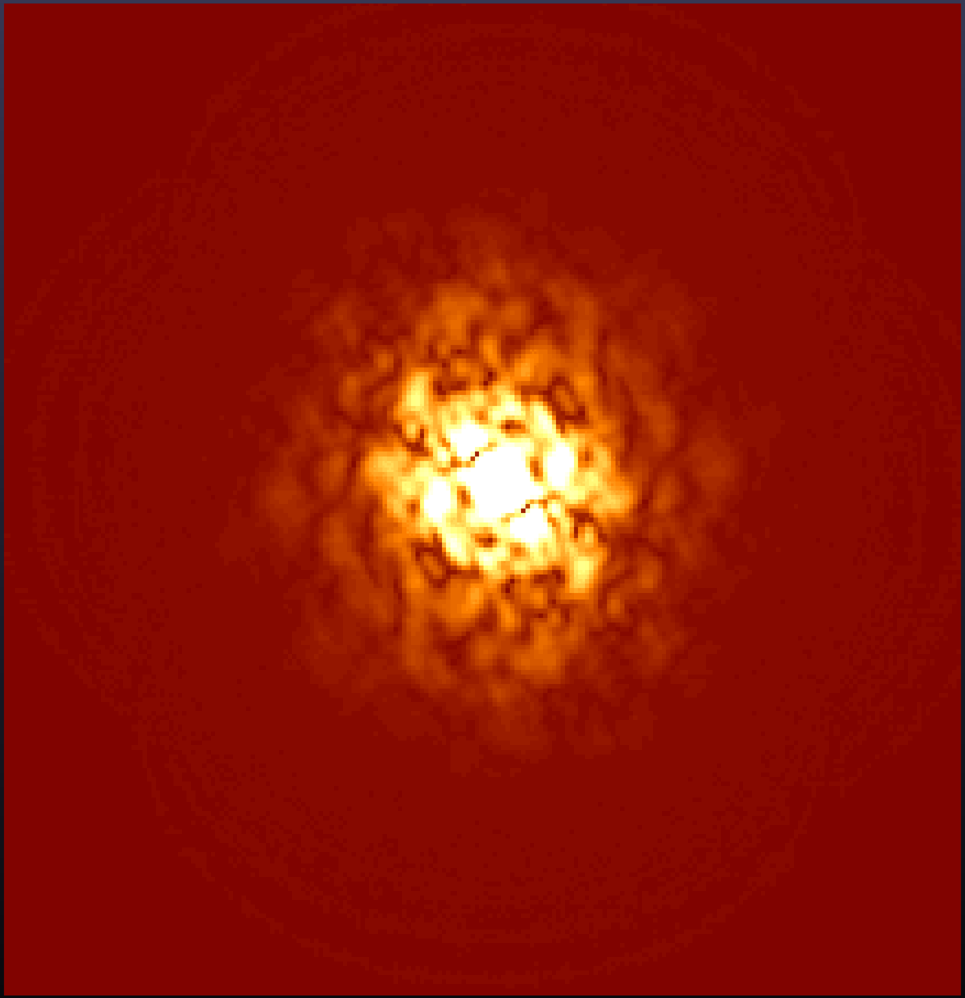
Sampled Vis



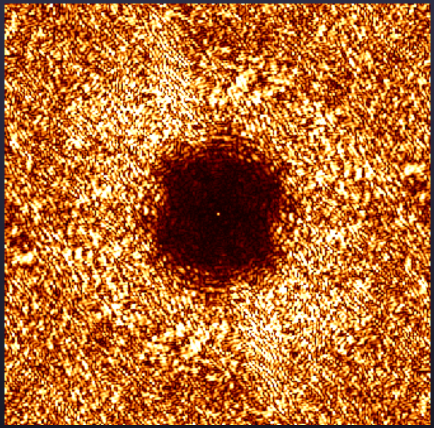
True Vis.



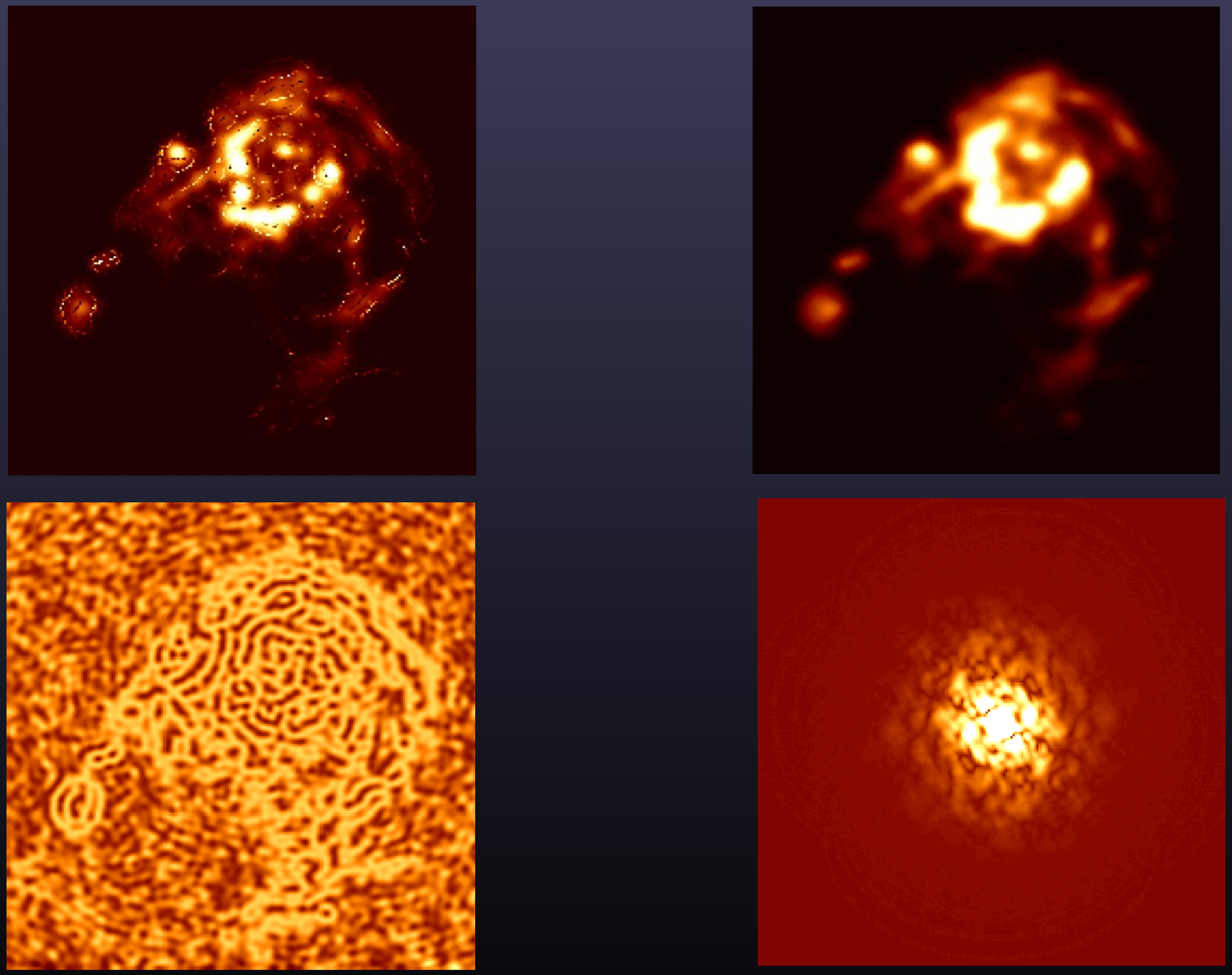
Model Vis.



Residual Vis.



MS-Clean: Example



Multi-resolution vs. Multi-scale Clean

- **Subtle difference between AIPS and AIPS++ implementations of scale sensitive**

  AIPS++: Each *iteration of the minor cycle* removes the optimal scale (one which reduces the residuals globally). Effectively, this achieves a "simultaneous" deconvolution at various scales [Multi-Scale Clean]

  AIPS: A decision, based on a user defined parameter, is made at the *start of each minor cycle* about the optimal scale to deconvolve [Multi-resolution Clean].

- **MS-Clean naturally detects and removes the scale with maximum power**
- **Removal of the optimal scale in MR-Clean strongly depends on the value of the user defined parameter.**

Scale Sensitive Deconvolution: Asp-Clean

- **Inspired by Pixon reconstruction (Puetter & Pina, 1994).**
- **Decompose the image into a set of Adaptive Scale Pixel (Asp) model (Bhatnagar & Cornwell, 2004).**
- **Algorithm:**
  1. **Find the peak at a few scales, and use the scale with the highest peak as an initial guess for the optimal dominant scale.**
  2. **Make a set of active-aspen containing Aspen found in earlier iterations and which are likely to have a significant impact on convergence.**
  3. **Find the best fit set of active-Aspen (expensive step).**
  4. **If termination criterion not met, goto 1.**
  5. **Smooth with the clean-beam. Add residuals if it has systematics.**

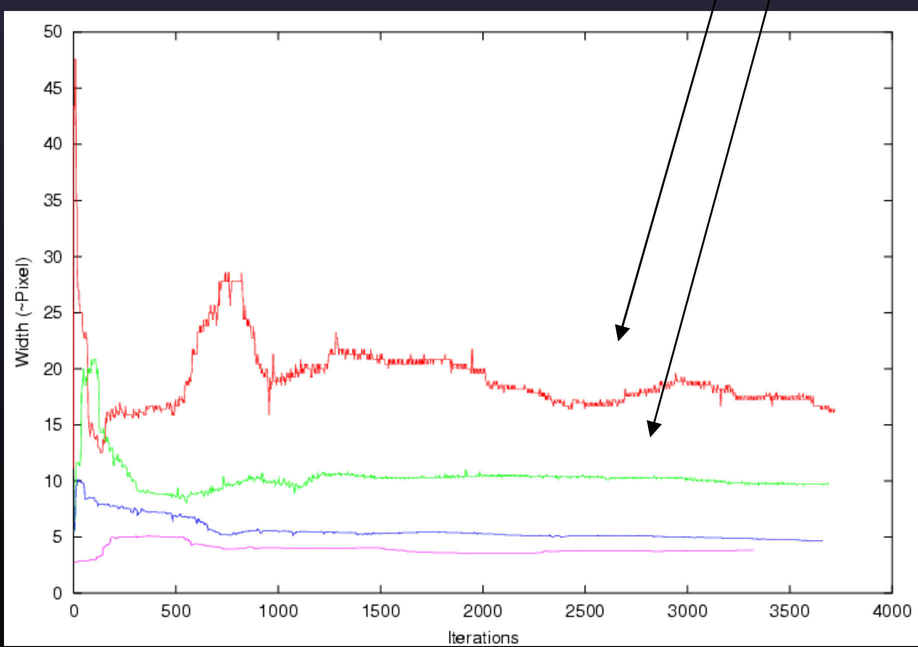
Asp-Clean details

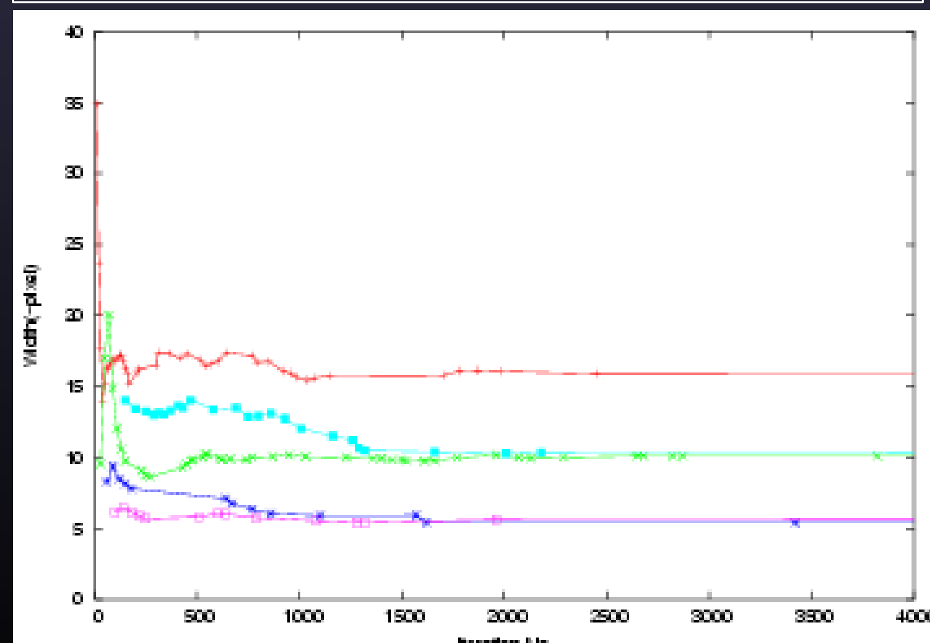
- ✔ Deals with non-symmetric structures better.
- ✔ Incorporates the fact that scale changes across the image. Residuals are more noise-like.
- ✔ Incorporates the fact that search space is potentially non-orthogonal (inherent coupling between Aspen).
- ✔ Aspen found in earlier iterations are not frozen.
- ✔ Scales well with computing power.
- ✗ Slower in execution speed.

Asp-Clean details: acceleration

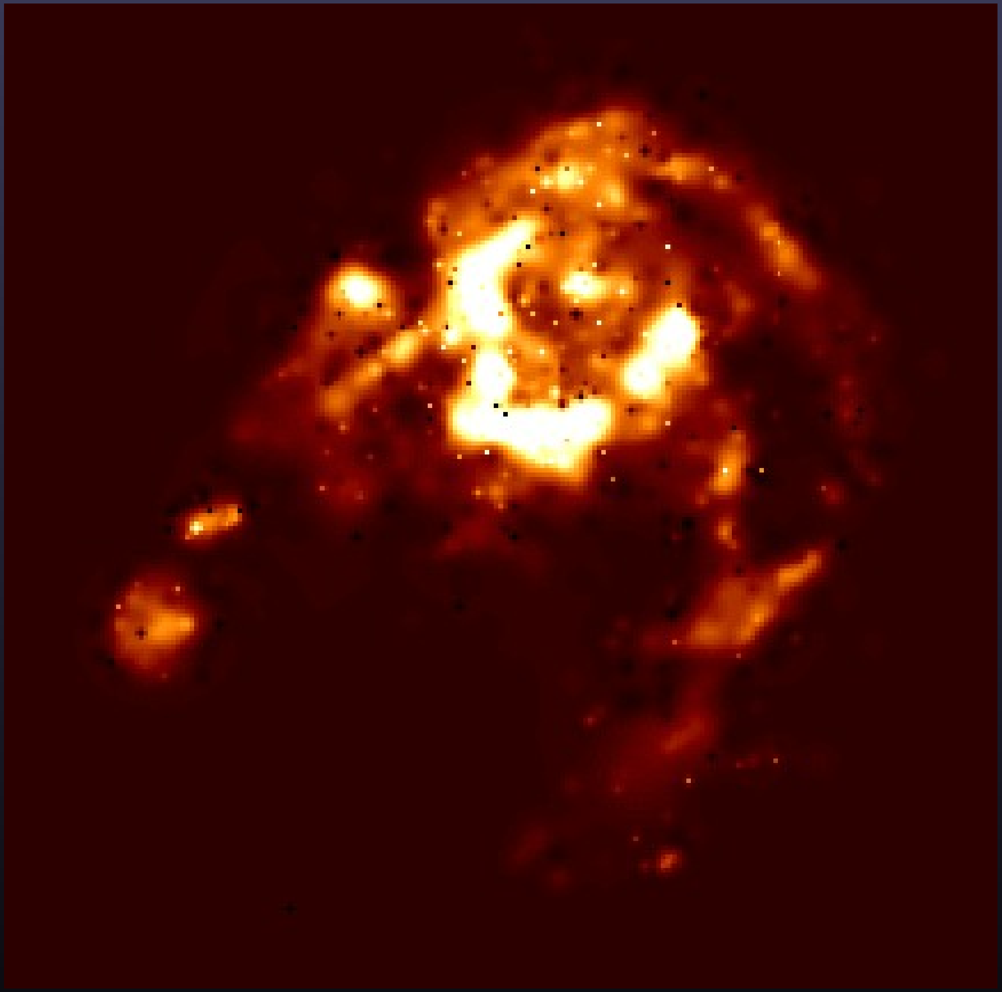
Fig 1: All Aspen are kept in the problem for all iterations. Scales all Asp scales evolve as a function of iterations.
Not all Aspen evolve significantly for at all iterations.

Fig 2: The active-set is determined by thresholding the first derivative. Only those Aspen, shown by symbols, are kept in the problem which are likely to evolve significantly at each iteration.

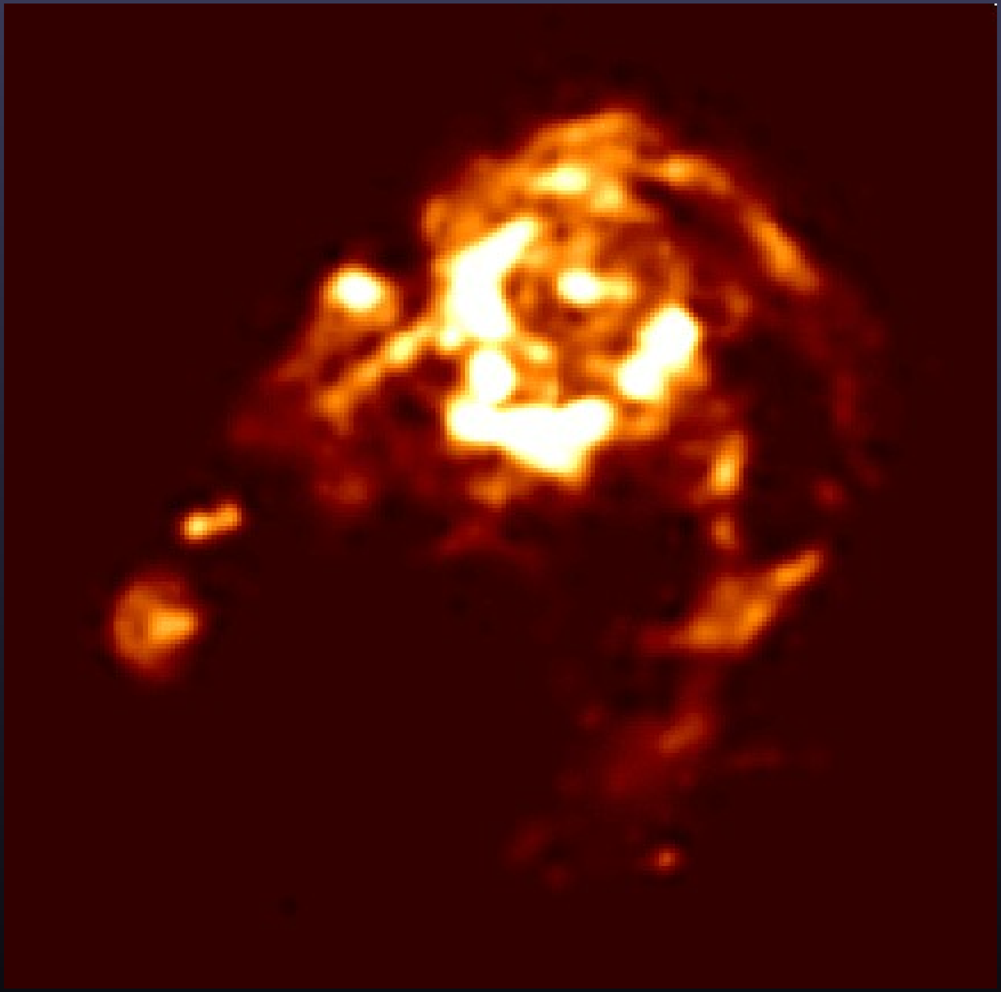




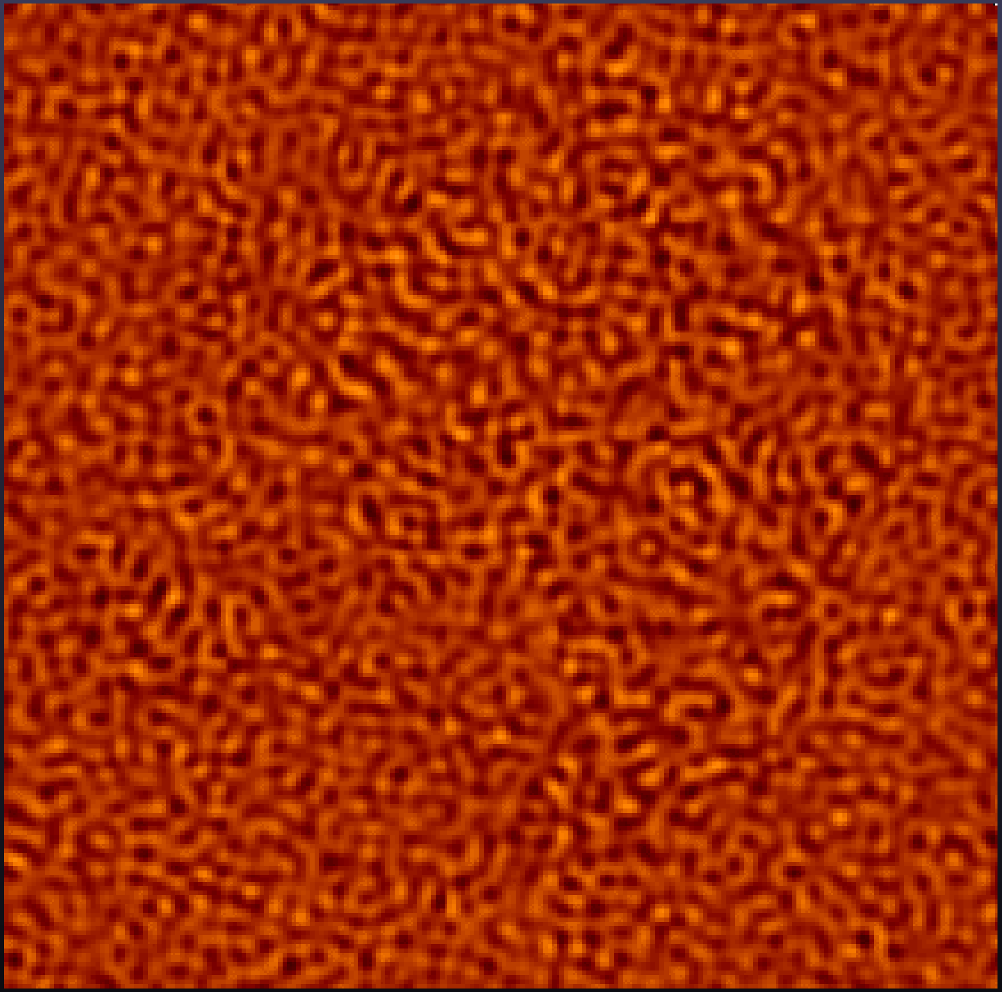
Asp-Clean: Model



Asp-Clean: Restored

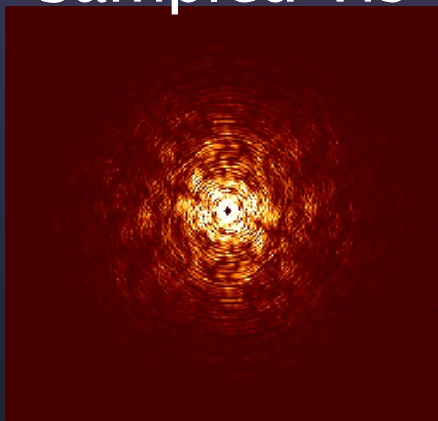


Asp-Clean: Residual

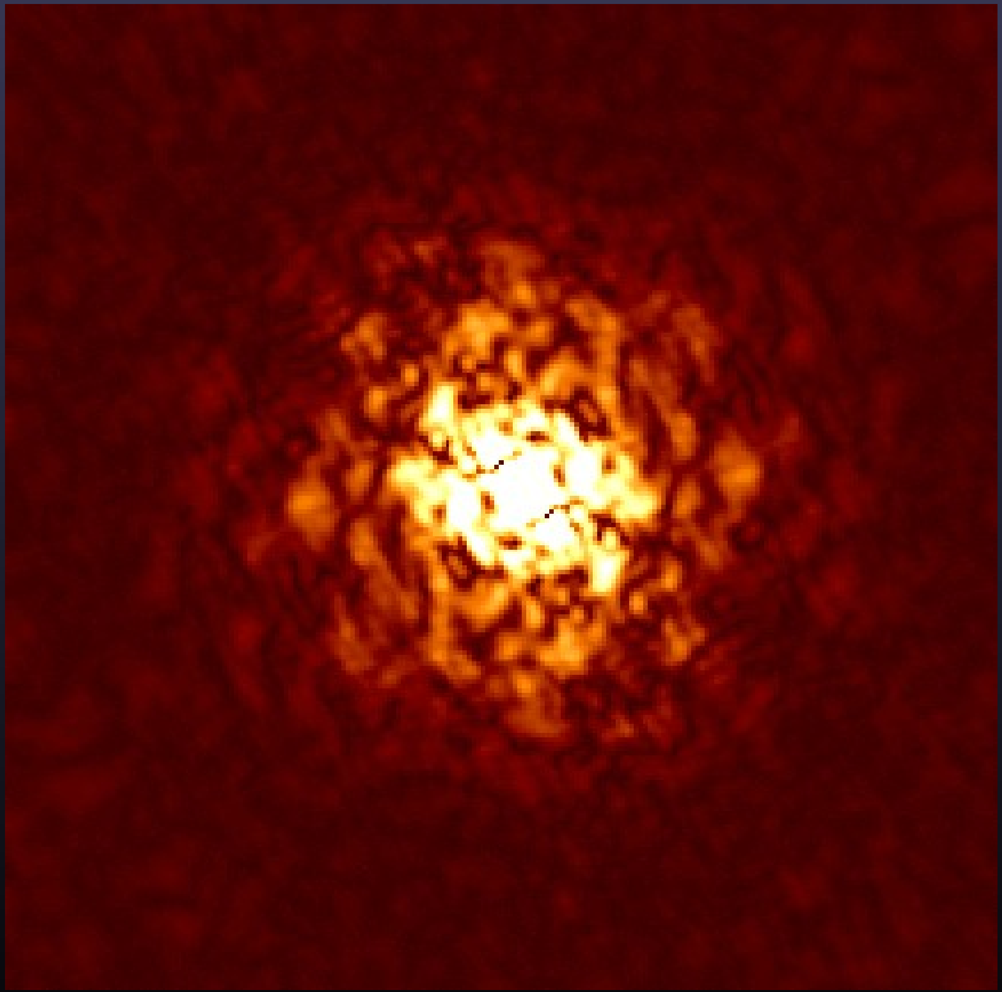


Asp-Clean: Model visibilities

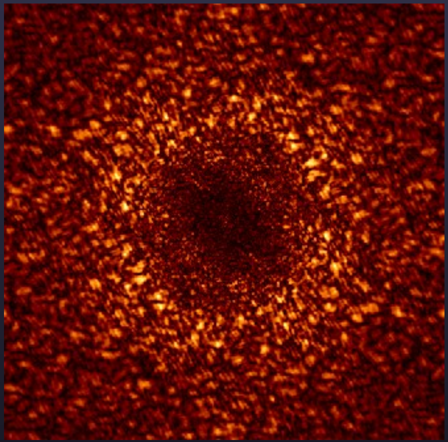
Sampled Vis



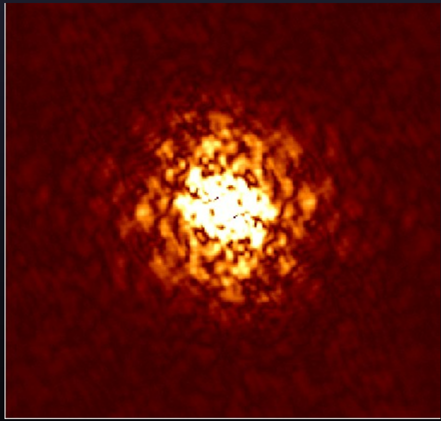
Model Vis.



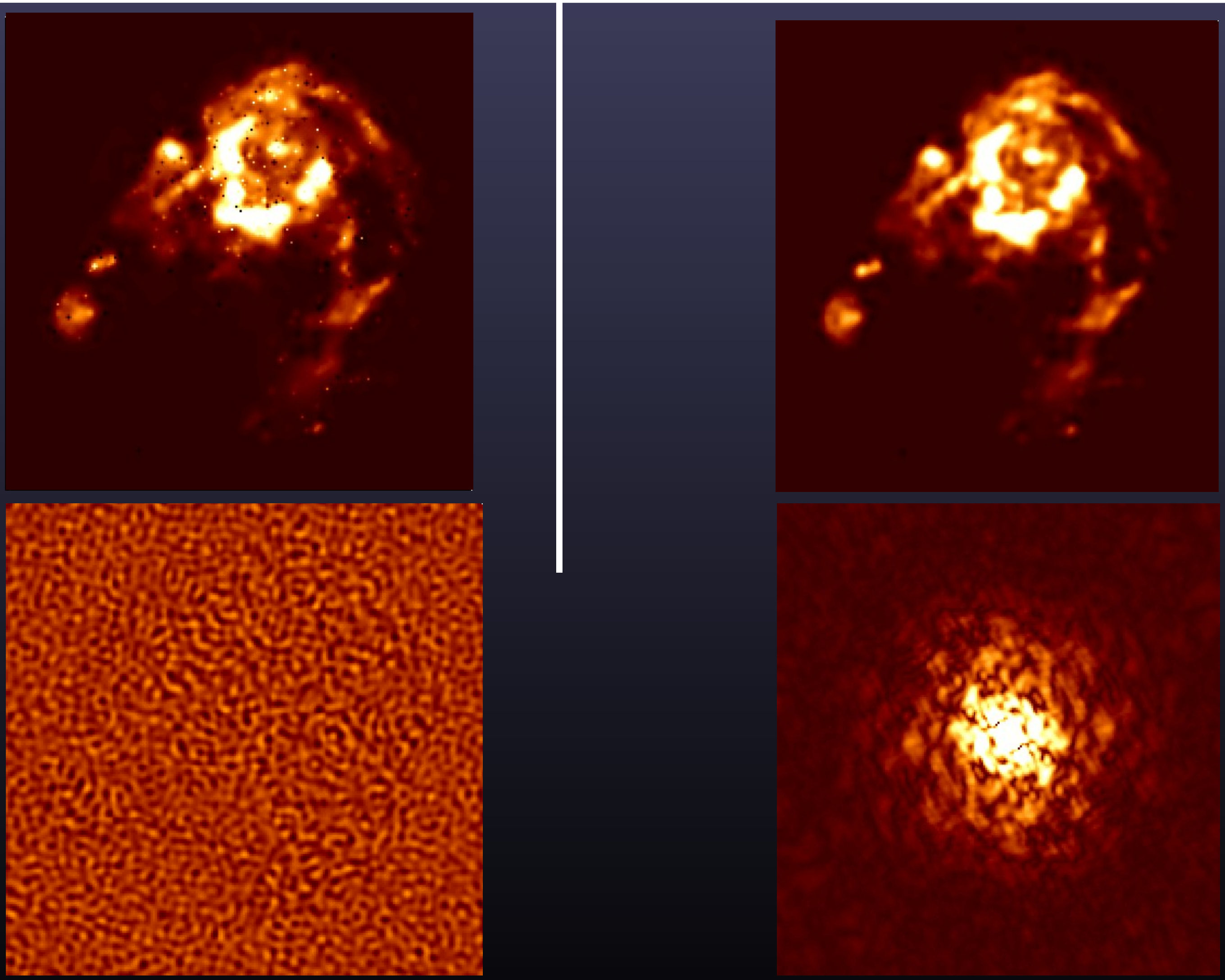
Residual Vis.



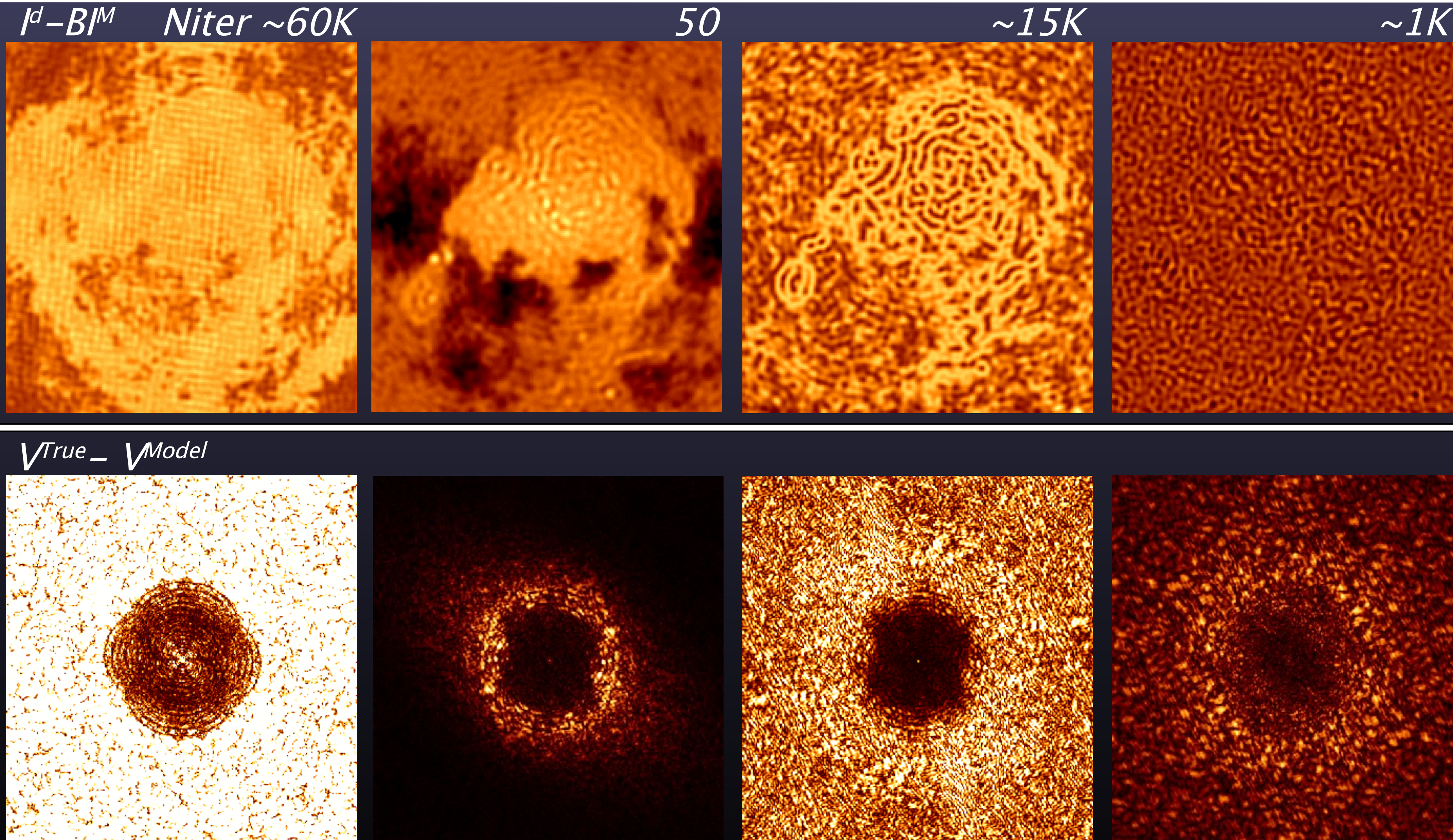
True Vis.



Asp-Clean: Example



Clean, MEM, MS-Clean, Asp-Clean



High Dynamic Range imaging

- **Full Beam Imaging**

$$V_{ij}^{obs} = J_{ij} \int J_{ij}^{Sky}(\vec{S}) I(\vec{S}) e^{2\pi \iota \vec{U}.\vec{S}} d\vec{S}$$

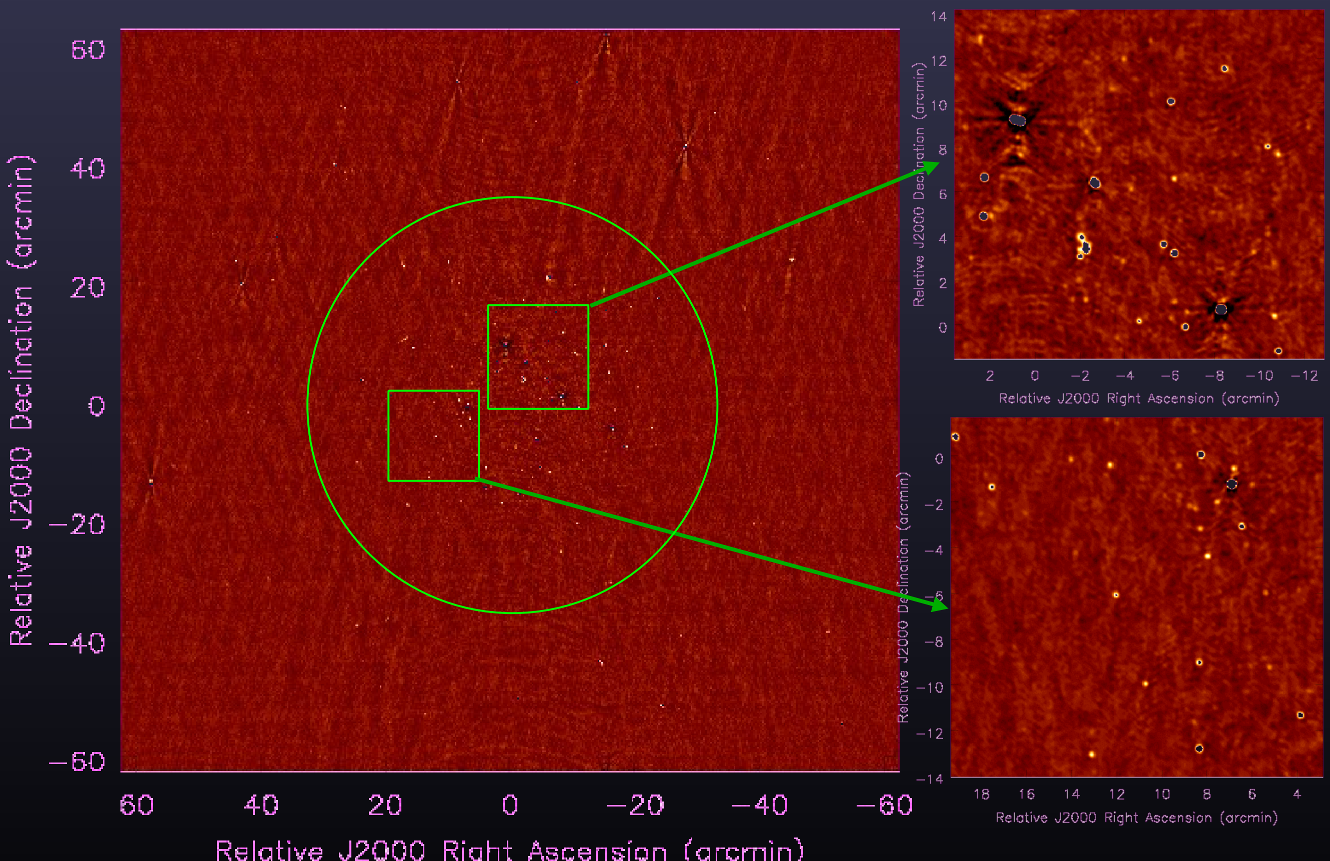
  - $J_{ij} \equiv \text{Direction independant gains}$
  - $J_{ij}^{Sky} \equiv$ **Direct dependent gains(e.g. Antenna Power Patterns *P(l,m)*)**
    - Rotation of PB on the sky + pointing errors gives time varying direction dependent gains (limits Mosaicking DR).
  - **Antenna polarization response is azimuthally asymmetric**
    - VLA polarization squint, full beam full Stokes imaging correction required even for moderate dynamic ranges
  - **Sky and antenna response frequency dependent**
    - **EVLA:** $\nu_{max} : \nu_{min} = 2:1$ **PB changes by a factor of 2!**

Mathematical Framework

- $V^{obs} = E A I$ where $E = A^T J^{Sky} A$
  - If E is unitary (or approximately so), use $A^T E^T$ as the inverse operator
  - Use $A^T E^T$ for update direction (Minor Cycle)
  - Use EA for residuals computation (Major Cycle)
- The modified transforms correct for image place errors
  - The W-Projection algorithm: Correction for non co-planarity
  - The PB-Projection algorithm: Correction for PB effects (pointing errors, poln., PB rotation)

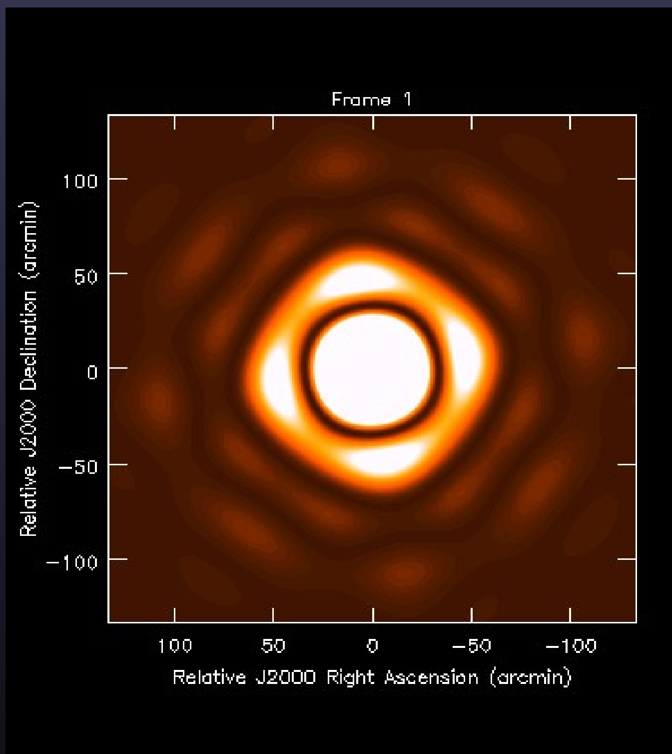
Wide field imaging

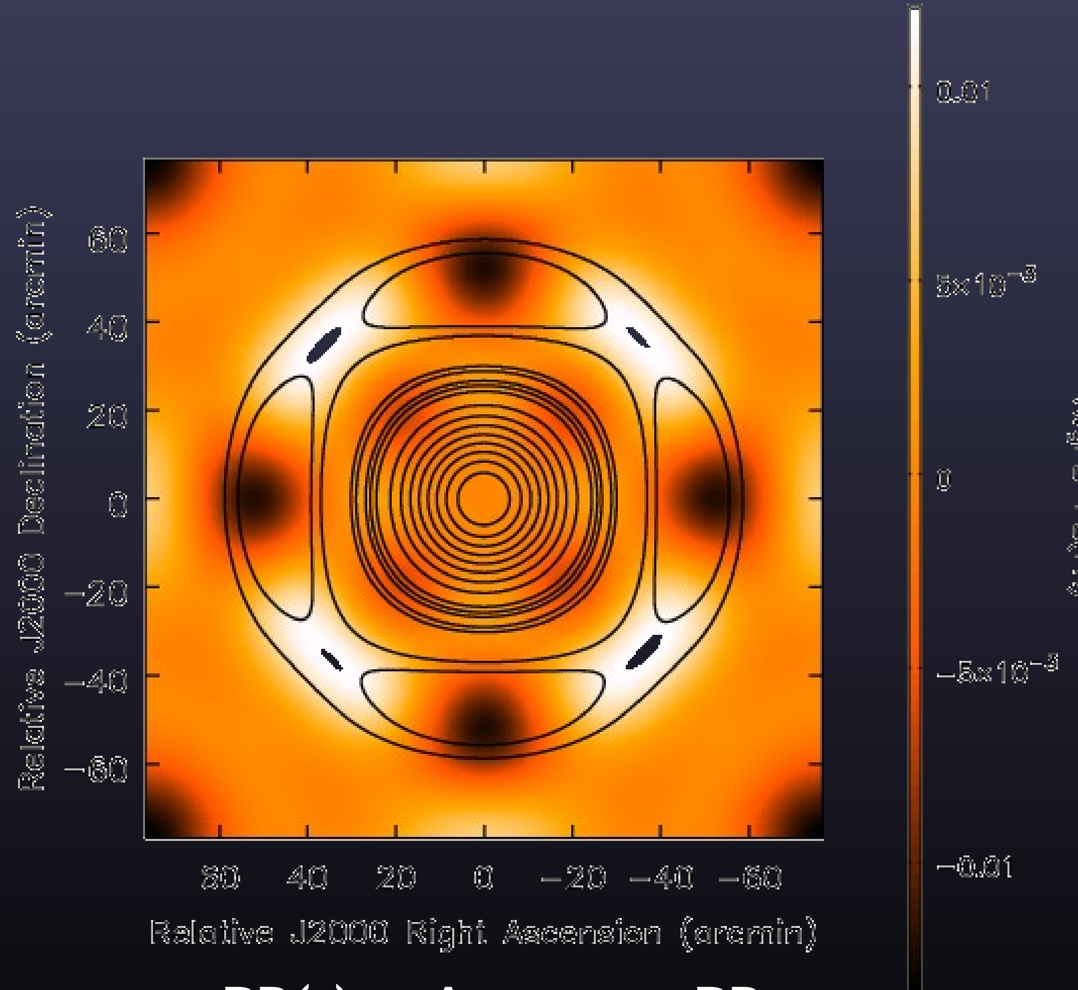
- W-projection: $\text{w-term} = e^{2\pi \iota w\left(\sqrt{1-l^2-m^2}-1\right)}$



- Errors not due to w-term
- Limited by pointing errors/PB-rotation?
- Errors in the first sidelobe due to PB-rotation

Primary Beams



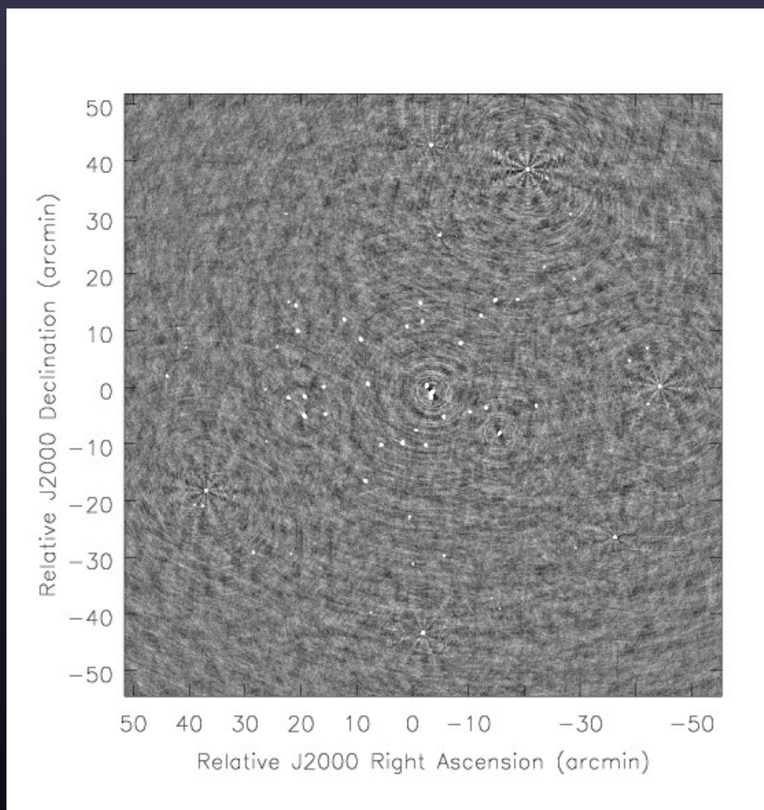


Stokes-I PB(PA)

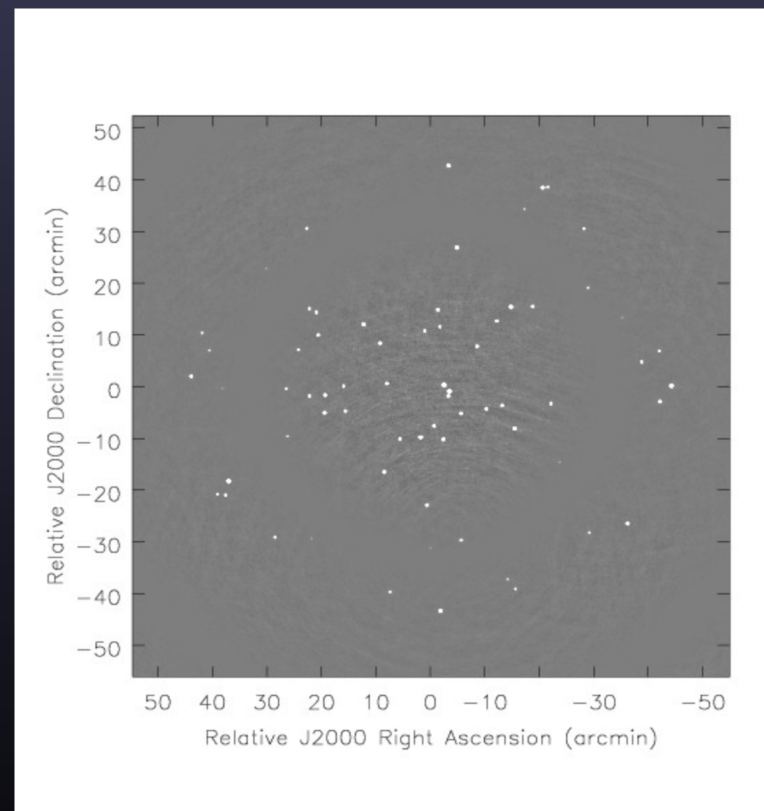
PB(t) - Average PB

Simulations: Stokes-I

- Stokes-I imaging with and without PB effects (Polarization squint, Pointing offsets, PB rotation)



RMS ~15μJy/beam



RMS ~1μJy/beam

Full-beam full-Stokes imaging

- Full Stokes imaging requires full Sky Muller matrix

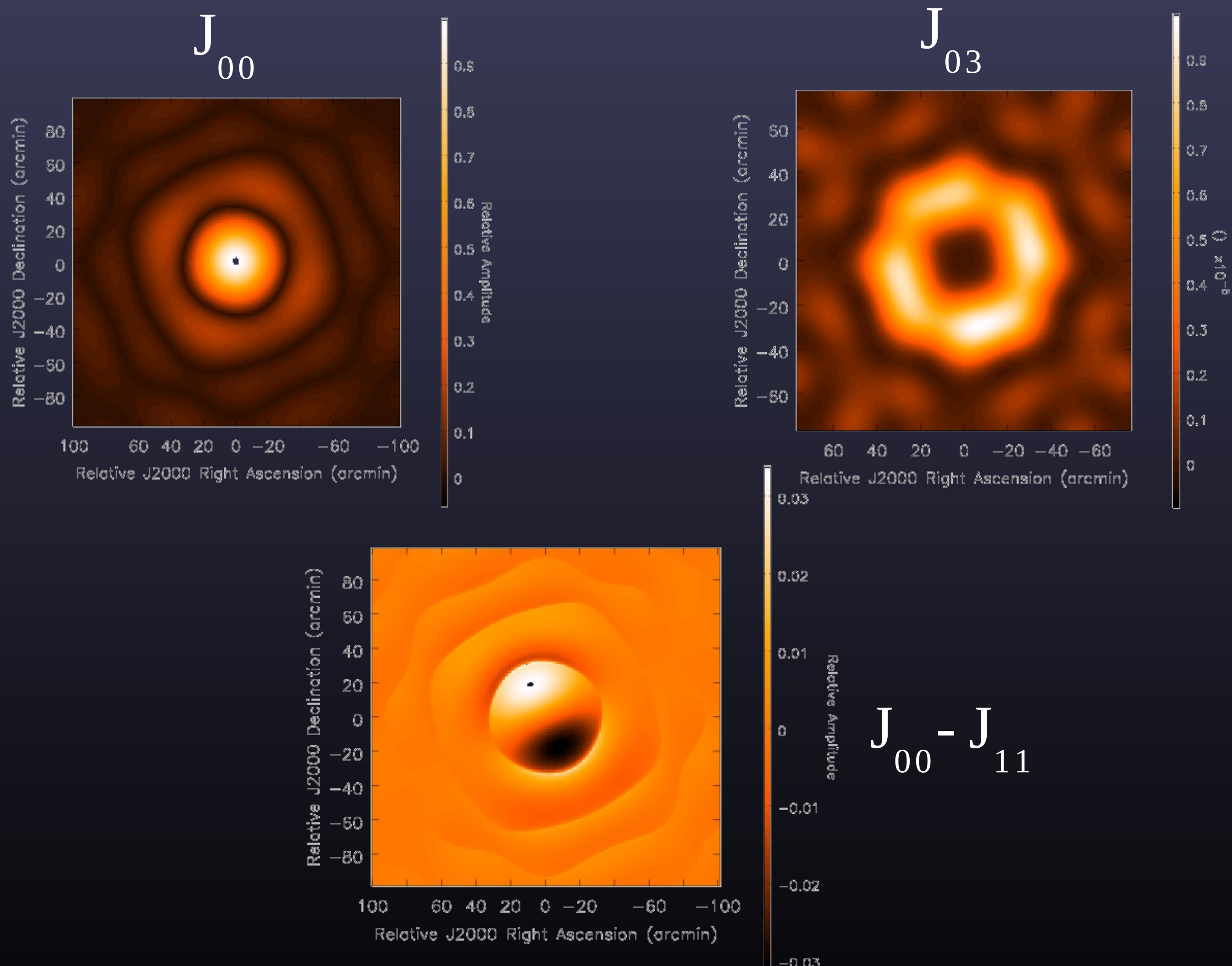
$$J_i(S) = \begin{bmatrix} J_i^{p} & J_i^{pq} \\ J_i^{qp} & J_i^{q} \end{bmatrix}$$

$$\begin{bmatrix} OUT_{pp} \\ OUT_{pq} \\ OUT_{qp} \\ OUT_{qq} \end{bmatrix} = \begin{bmatrix} J_i^{p} J_j^{p^*} & J_i^{p} J_j^{pq^*} & J_i^{pq} J_j^{p^*} & J_i^{pq} J_j^{pq^*} \\ J_i^{p} J_j^{qp^*} & J_i^{p} J_j^{q^*} & J_i^{pq} J_j^{qp^*} & J_i^{pq} J_j^{q^*} \\ J_i^{qp} J_j^{p^*} & J_i^{qp} J_j^{pq^*} & J_i^{q} J_j^{p^*} & J_i^{q} J_j^{pq^*} \\ J_i^{qp} J_j^{qp^*} & J_i^{qp} J_j^{q^*} & J_i^{q} J_j^{qp^*} & J_i^{q} J_j^{q^*} \end{bmatrix} \begin{bmatrix} IN_{pp} \\ IN_{pq} \\ IN_{qp} \\ IN_{qq} \end{bmatrix}$$

$$J_{ij}(S) = J_i(S) \otimes J_j^*(S)$$

- $J_i^p(\vec{S}), J_i^q(\vec{S}) \equiv$ Antenna voltage pattern in orthogonal polarization
- $J_i^{pq}(\vec{S}), J_i^{qp}(\vec{S}) \equiv$ Leakage of polarization due to reflection from curved surface
- $J_{ij}(\vec{S})$ is not identity or even diagonal for DR$>10^4$

Structure of the Sky-Muller Matrix



References

1. **Synthesis Imaging in Radio Astronomy II; Imaging and deconvolution.**
2. **High Fidelity Imaging of Moderately Resolves Sources; Briggs, D. S., PhD Thesis, New Mexico Tech., 1995**
3. **Understanding radio polarimetry. I,II,III; Hamaker, J., P., Bregman, J.D., & Sault, R.J. A&A, 117,137, 1996**
4. **Scale sensitive deconvolution of astronomical images: I-Adaptive Scale Pixel (Asp) model; Bhatnagar,S. & Corwnell,T.J., A&A, 426, 747-754, 2004 (astro-ph/0407225).**
5. **W Projection: A New Algorithm for Non-Coplanar Baselines; Cornwell,T.J., Golap,K., & Bhatnagar,S., 2003, EVLA Memo 67**
6. **Solving for the antenna based pointing errors; Bhatnagar,S., Cornwell,T.J., & Golap,K., 2004, EVLA Memo 84**
7. **Corrections of errors due to antenna power patterns during imaging; Bhatnagar,S., Cornwell,T.J., & Golap,K., 2006, EVLA Memo 100**