Imaging algorithms, AIPS++, and the GMRT

S. Bhatnagar
AOC, NRAO

22March, 2004

Imaging algorithms: General structure

- Interferometric observations can be described as:

  $\boldsymbol{V} = \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{N})$ where $\boldsymbol{A}$ = FT operator, $\boldsymbol{I}$=Image vector, and $\boldsymbol{N}$ the noise vector.

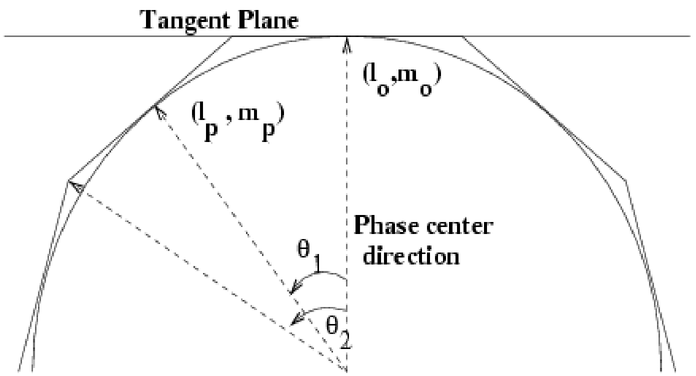
- Iterative algorithms involve two major operations:
  - Reconcile with the data: $\boldsymbol{A}^T(\boldsymbol{V} - \boldsymbol{A}\boldsymbol{I}^M)$ (major cycle)
  - Update model image using $\boldsymbol{B}'(\boldsymbol{V} - \boldsymbol{A}\boldsymbol{I}^M)$ *where* $\boldsymbol{B}'$ *is an approx. of* $\boldsymbol{A}^T$ (minor cycle)

- For scale-less deconvolution, major cycle costs dominate. This is addressed in the first half of the talk (w-projection).

- Quality of image reconstruction, image fidelity and dynamic range depends on the image modeling in the minor cycle. This is addressed in the second half of the talk (Asp Clean).

Wide field imaging: Image plane faceting

- Major cycle in wide-field imaging uses multiple facet-images to compute the residual visibilities.
- Image plane faceting



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

- Re-phase and image on multiple tangent planes such that $l^2+m^2\approx 0$ within the facet
- Re-project facet images onto the plane tangent to the phase center $\begin{pmatrix} l \\ m \end{pmatrix} \approx \boldsymbol{R}_2 \begin{pmatrix} l' \\ m' \end{pmatrix} + \begin{pmatrix} l_p \\ m_p \end{pmatrix}$
- Multiple images and facet edge effects are a problem

Wide field image: UV-faceting

- Linear image co-ordinate transformation (***C***) has an equivalent UV co-ordinate transformation (*Sault et al. (1996)* ; rotation and stretching in the image plane):

$$\boldsymbol{I}(\boldsymbol{C}\,l) \rightarrow |det(\boldsymbol{C})|^{-1}\,\boldsymbol{V}(\boldsymbol{C}^{-1^T}u)$$

- Projection error: $\epsilon = \sin(\theta_1)(1-\cos(\theta_2)) \approx \frac{1}{2}\theta_1\theta_2^2$ same as in image plane faceting!

  - ✓ Single image
  - ✓ No edge effects
  - ✓ Global deconvolution possible
  - ✓ Region definition as in the usual case

Interferometry: Recap

- $V(u,v,w) = G(u,v,w) * V(u,v,w=0)$

  $where\, \overline{G}(l,m,w) = e^{2\pi\iota\left[w\left(\sqrt{1-l^2-m^2}\right)\right]}$

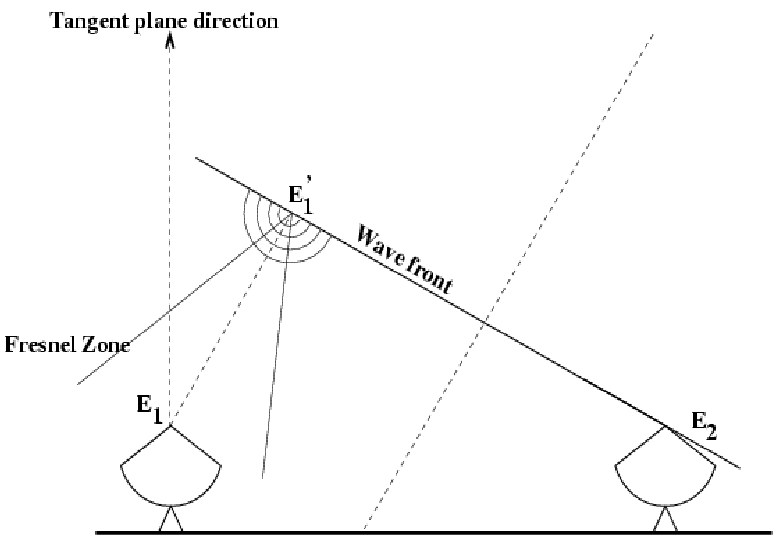
- $\rho_{12} = \langle \boldsymbol{E}_1(u,v,w=0)\, \boldsymbol{E}_2^{\,*}(0,0,0) \rangle$

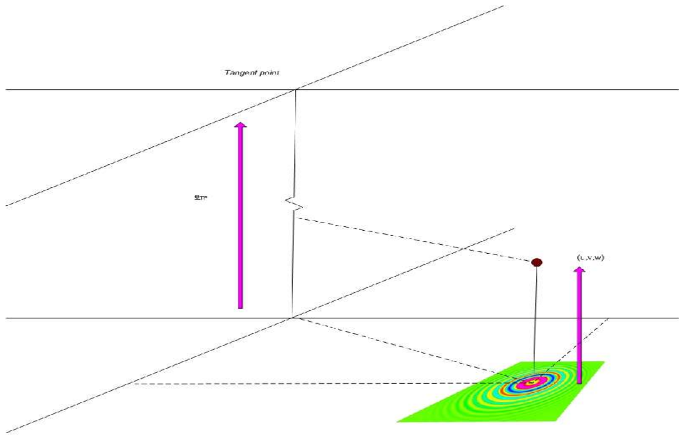
  ***E**$_1$**=E'**$_1$(u,v,w)* propagated using Fresnel diffraction theory. The above convolution equation is reproduced with $\frac{r_F}{\lambda} \approx \sqrt{w}$

- A $w \neq 0$ interferometer is not a device to measure a single Fourier component.

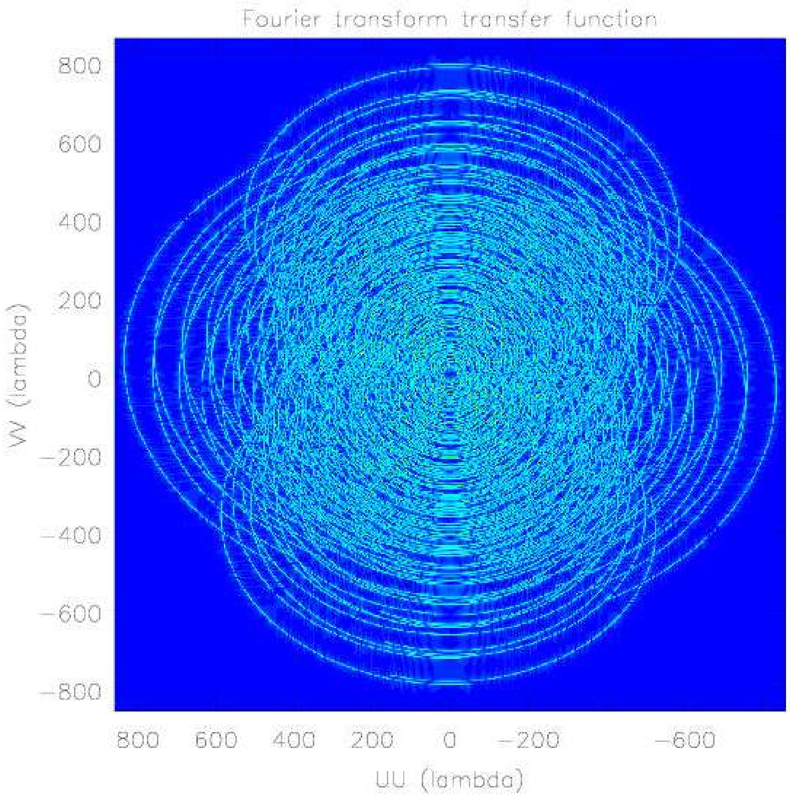
- Thickness of the uv-tracks changes along the tracks as *w* changes.

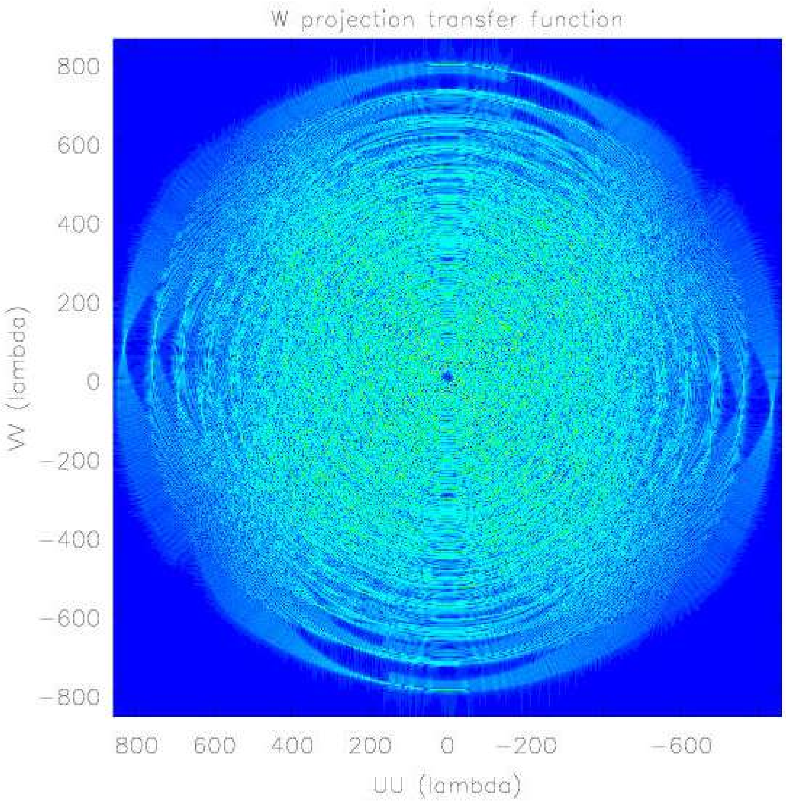
- The concept of redundant baselines is much more restricted than usually thought.





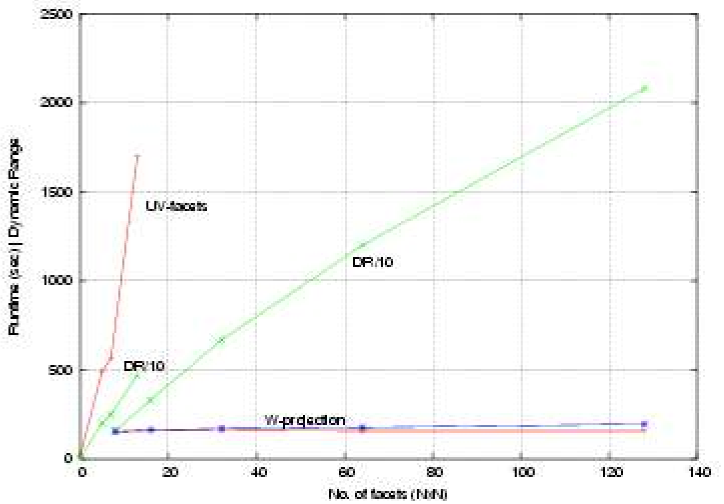
UV-coverage: Recap





W-projection

- Visibility computation (de-gridding):
  - Multiply model image by taper *T*
  - Do a 2D Fourier transform
  - Evaluate the convolution to de-grid
- Dirty image computation (gridding):
  - Evaluate the convolution for $(u,v,w=0)$ plane
  - Do a 2D inverse Fourier transform
  - Divide the image by *T*



Scaling laws:

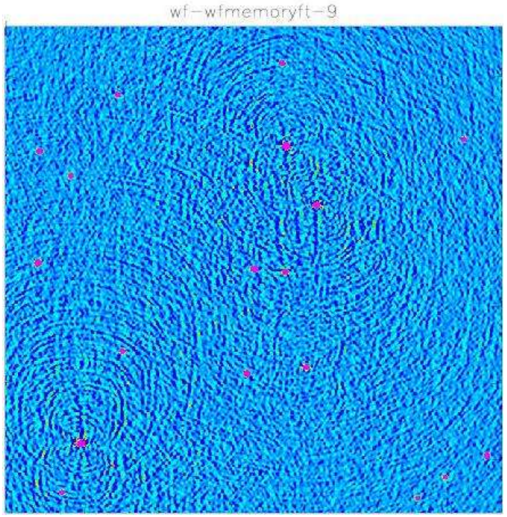
W-projection: $(N^2_{wproj} + N^2_{GCF}) N_{vis}$

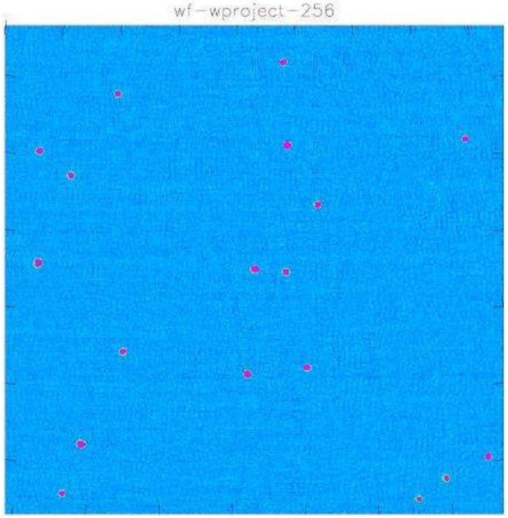
UV-facet: $N^2_{facets} N^2_{GCF} N_{vis}$

Ratio: $\approx N^2_{GCF}$ *for large no. of facets*

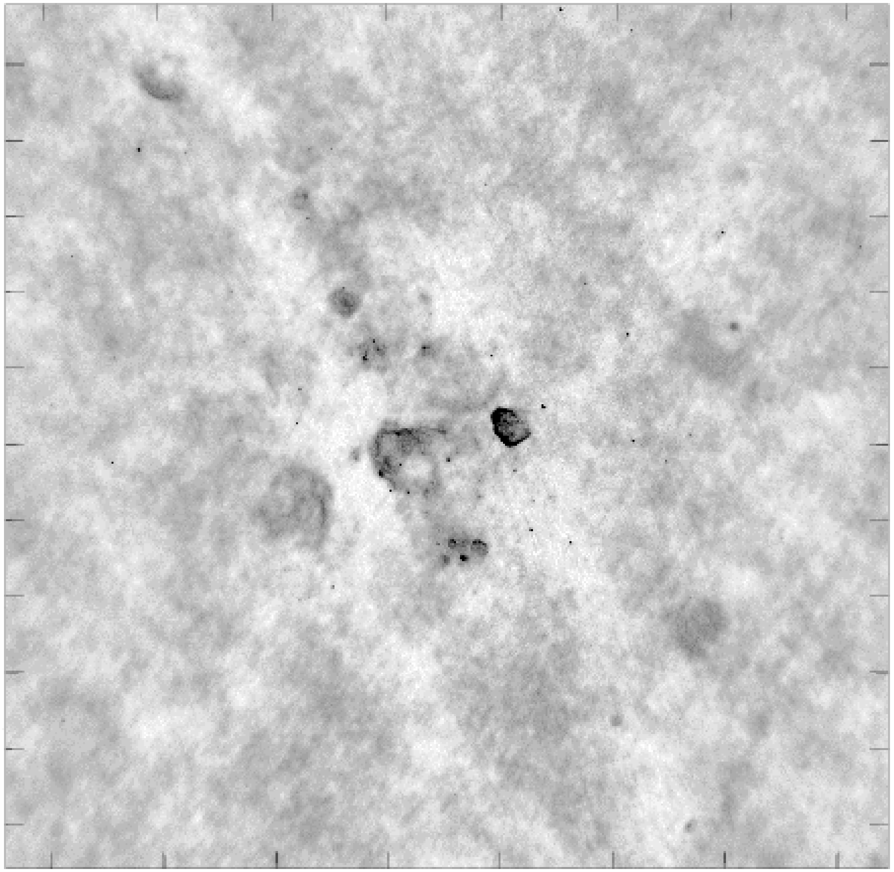
- Pre-compute $T(u,v)G(u,v,w)$ with uniform sampling in $\sqrt{w}$ such that aliasing effects are less than the required dynamic range.

W-projection: Simulation/Example





Galactic Plane at P-band – VLA B,C,D (Brogan et al.)



Scaleless deconvolution

- Measurement equation describing an interferometer can be written as:

$$V = A\,\boldsymbol{I}^{o} + AN \qquad where\ \boldsymbol{I}^{o} \equiv True\ image\ vector\,,\ \boldsymbol{N} \equiv Noise\ vector$$

- Deconvolution is a search for a model image ($\boldsymbol{I}^M$) which solves:

$$A^T V = A^T A\,\boldsymbol{I}^M + A^T A\,N \qquad \vee \qquad \boldsymbol{I}^D = \boldsymbol{B}\boldsymbol{I}^M + \boldsymbol{I}^N$$

- Clean (and its variants) is steepest descent minimization of the objective function $\chi^2 = [V - A\boldsymbol{I}^M]^T W [V - A\boldsymbol{I}^M]$ using the update relation:

$$\boldsymbol{I}_k^R = \boldsymbol{I}_{k-1}^R - g\,[\boldsymbol{max}\,\boldsymbol{I}_{k-1}^R]$$

- MEM (and its variants) is constraint minimization of the objection function, with prior (assumed) knowledge encoded in the entropy *H*:

$$f(\boldsymbol{I}^M, \lambda) = H - \lambda\,\chi^2$$

Multiscale methods

- However both use $I^M = \sum_k F_k \delta(x - x_k)$ which has no scale information (leads to a diagonal approximation of the Hessian).
- Scale fundamentally separates signal from noise without which large scale low level emission cannot be recovered (***BI^M*** and ***I^N*** are comparable)
- **Pixon method** *(Putter&Pina (1994); Putter&Yahil(1999))*

  Explicitly assumes finite support PSF (filled aperture telescopes) and independent image-plane noise (direct imaging devices). Not useful for interferometric imaging.

- **MS-Clean** *(Cornwell&Holdaway (in prep.))*

  Decomposes the image into a set of a few symmetric kernels.

  ✓ Reasonably fast and retains the shift-scale-n-add nature of Clean

  ✗ Non-symmetric features poorly reconstructed + slow convergence

  ✗ Effectively uses diagonal approximation of the Hessian (ignores coupling)

  ✗ Scales poorly

Scale sensitive (Asp) decomposition

- Scale sensitive parametrization using **Adaptive Scale Pixel (Asp) model**

$$I^M = \sum_k F_k P(\vec{p}_k)$$

- Compute the approximate Asp model in the minor cycle. Use analytical form to compute the gradient: $\Delta_k = [\boldsymbol{I}^R] \left[\frac{\partial \boldsymbol{P}}{\partial p_k} \right]$ (second term provides a finite support).

- Use w-projection for full reconciliation with the data.

  - ✔ Uses minimum DOFs compared to other algorithms.
  - ✔ Continuous range of scales.
  - ✔ Good reconstruction of all shapes. Residuals noise-like.
  - ✔ Explicitly retains coupling between Aspen.
  - ✗ Comparatively slower (3 times slower than MS-Clean).
  - ✗ Slows down with iterations (dimensionality of the search space increases).

Asp decomposition: Examples

S. Bhatnagar and T.J. Cornwell: Scale sensitive deconvolution of interferometric images 5

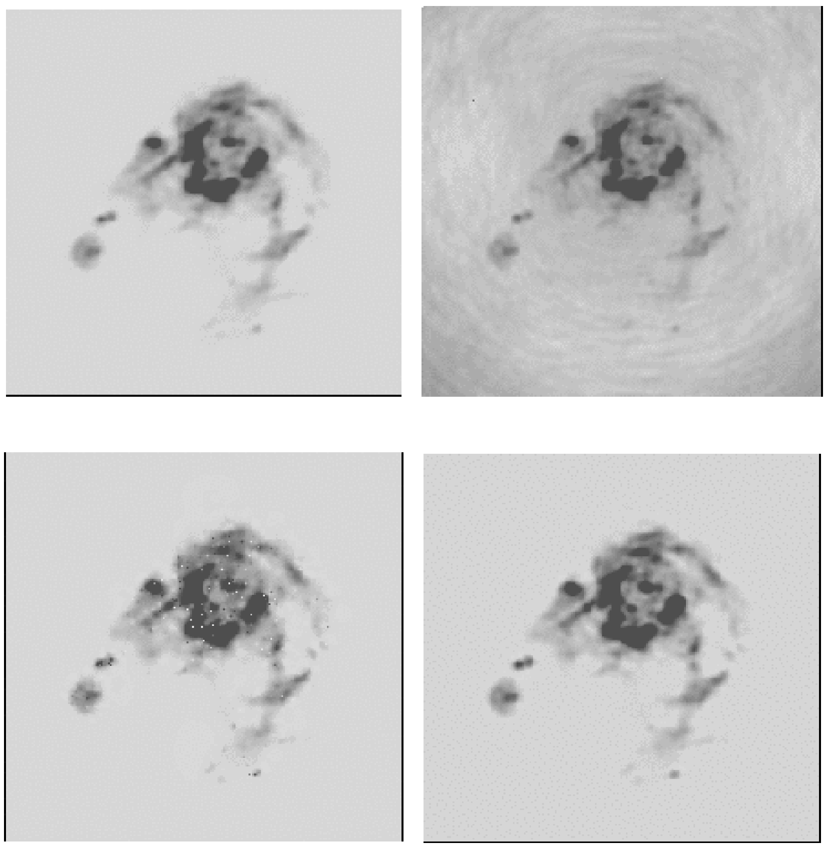
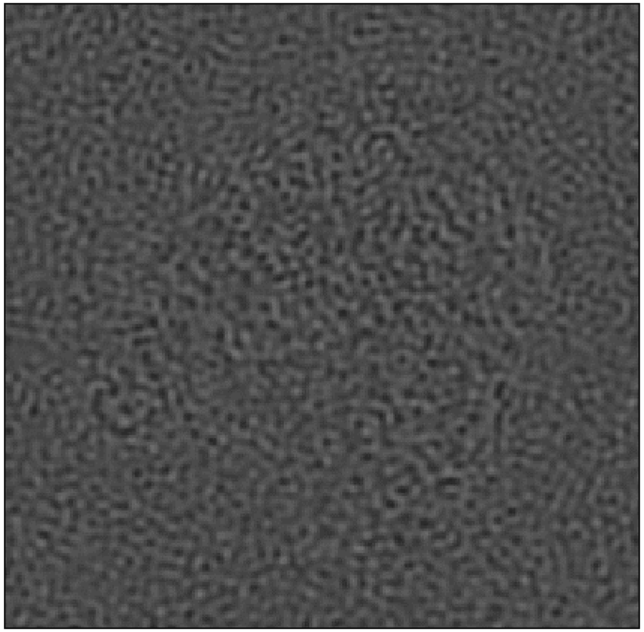
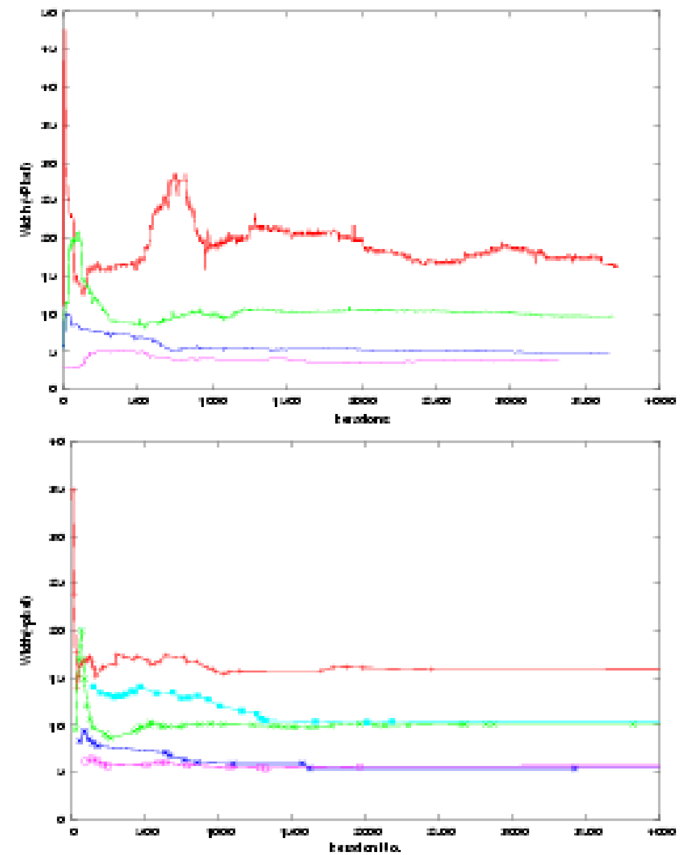


Fig. 2. Figure showing an example of Asp reconstruction of a typical astronomical image. Top left panel shows the HI image made with the VLA, used as the "true image" (I^o) for the simulation. The image contains ~ 10 000 pixels with significant emission. This image was used to simulate visibilities corresponding to a VLA observation. The corresponding dirty image (I^d), shown in the top right panel, was then deconvolved using the Asp-Clean algorithm. A 800-Asp component reconstructed model image (I^M) is shown in the lower left panel. The lower right panel shows the restored Asp-model image ($CI^M + I^R$, where C is the smoothing operator corresponding to the resolution element).

The residual image



Asp decomposition: Acceleration



- Not all Aspen are active at all iterations. Significant loss of efficiency due to in-active Aspen.
- Adaptively contain the dimensionality of the search space by retaining only the active-set.
- Active set computed at the beginning of each iteration by thresholding $L_k = |\nabla_k \chi^2|$
- Threshold $L_o \propto \sum_i \boldsymbol{I}_i^R$
- To Do:
  - Fast computation of the covariance matrix for thresholding.
  - Use more exotic Asp forms (further reduce the final DOFs).
  - MCMC? Its variations?

Asp decomposition: Acceleration

- Gaussians allow control on the scale, and orientation. There is no handle on the shape or support/roll-off.
- Leads to larger number of components and higher runtime.
- Higher Order Gaussians (HOGs)?