Experiences With EVLA Data

CALIM 2010, Dwingeloo, Aug. 24th. 2010

S. Bhatnagar

NRAO

High sensitivity imaging

- Sensitivity $\propto \frac{N_{ant} A_{ant} \sqrt{N_t \tau N_{chan} \Delta \nu}}{T_{sys}}$
- Higher sensitivity is achieved using larger collecting area ($\propto \mathbf{N_{ant}}$), wider band-widths ($\mathbf{N_{channels}}$) and longer integrations in time ($\mathbf{N_t}$)
  - Data volume $\propto \mathbf{N^2_{ant}\ N_{channels}\ N_t}$

- Implications for high dynamic range imaging
  - Wider field imaging required → finer sampling in time and frequency
  - $\mathbf{N_{channels}}$ = 1-10GHz/MHz and $\mathbf{N_t}$ = 10hr/(1-10sec)
  - Wider range of angles on the sky (==> Direction Dependence)
  - Smaller scale variations over larger parameters space to be accounted for
  - Algorithm efficiency remains a critical parameter
  - 10-100x increase in the number of samples to achieve the required sensitivities

The EVLA

- 27-antenna array: 25m diameter, Az-El mount
- Continuous frequency coverage from 1-50 GHz in 8 bands
- Bandwidth ratio (n_{high} : n_{low}) = 2:1

- Instantaneous bandwidth: 8 GHz (currently 2x1GHz)
- WIDAR Correlator: Upto 16K channels

- Data rate: Currently ~12 Mbytes/s (512 channels)
  - 350 GB/per typical observation (6hr)

- Sensitivity: Thermal noise ~1 microJy/beam in ~1 hr.
  → Dynamic range of 10^6 for a 1 Jy point source

Synthesis Imaging

- Measurement Equation

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

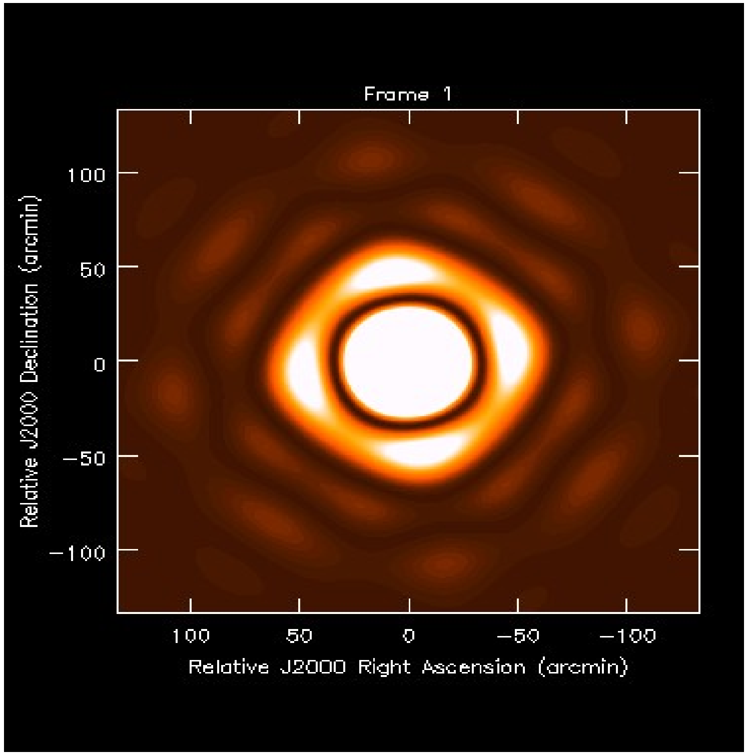
$M_{ij}(\nu,t) = J_i(\nu,t) \otimes J_j^*(\nu,t)$:Direction <u>independent</u> gains

$M_{ij}^{S}(s,\nu,t) = J_i(s,\nu,t) \otimes J_j^*(s,\nu,t)$:Direction <u>dependent</u> (DD) gains

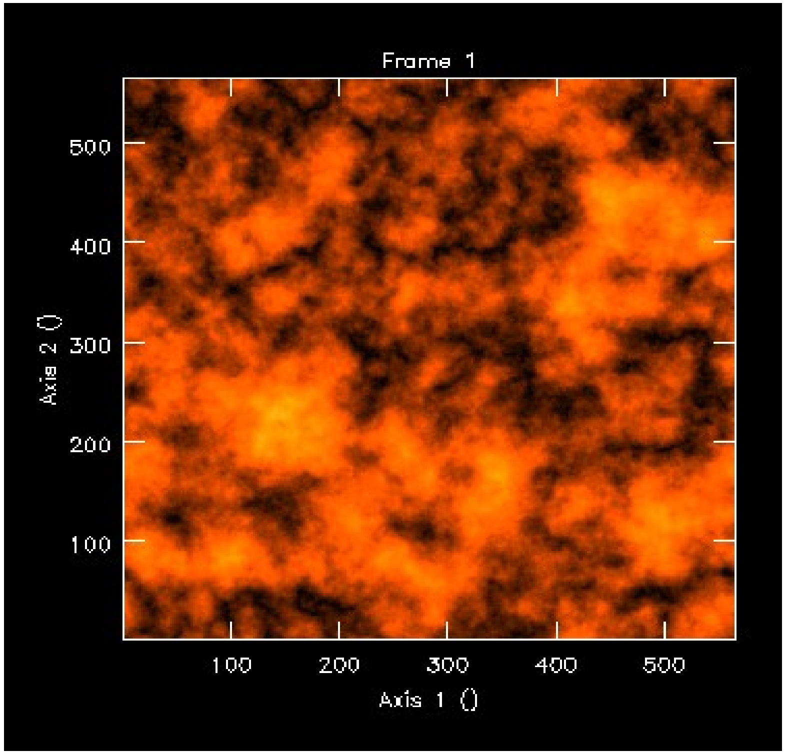
- Requirements: Full beam, full band, full Stokes imaging
  - Wide-band, narrow field: Ignore $M_{ij}^{S}(s,\nu,t)$
  - Narrow-band, wide-field: Ignore frequency dependence of ***I***
  - Wide-band, wide-field: A-Projection + MS-MFS
  - High dynamic range: All the above + DD solvers
    - Time dependent pointing errors, PB shape change, etc.

- ***Combined RHS determines the "time constant" over which averaging will help***

Examples of DD effects

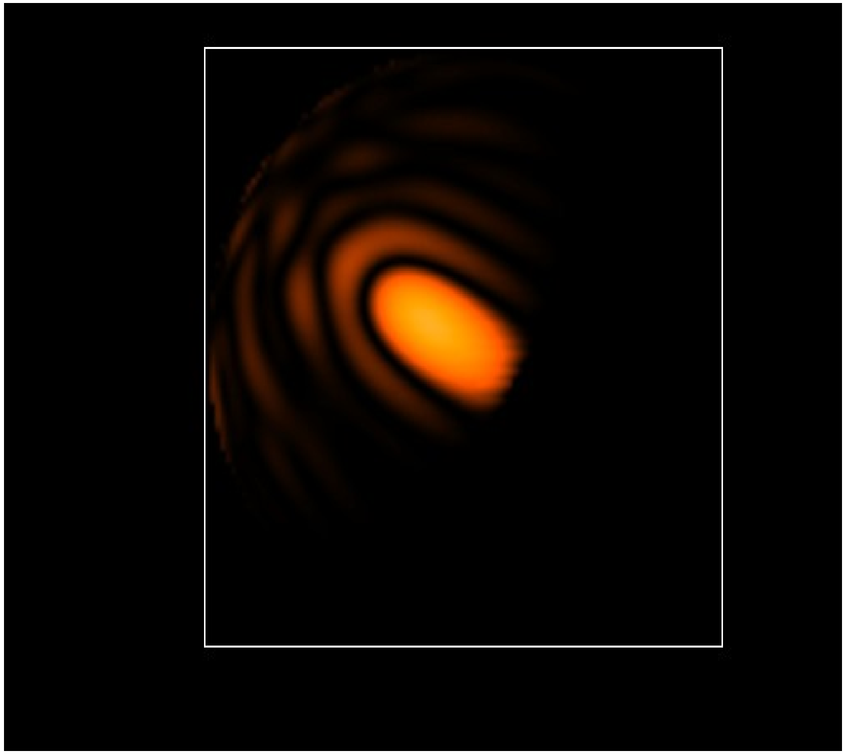


Time and DD Primary Beam: EVLA

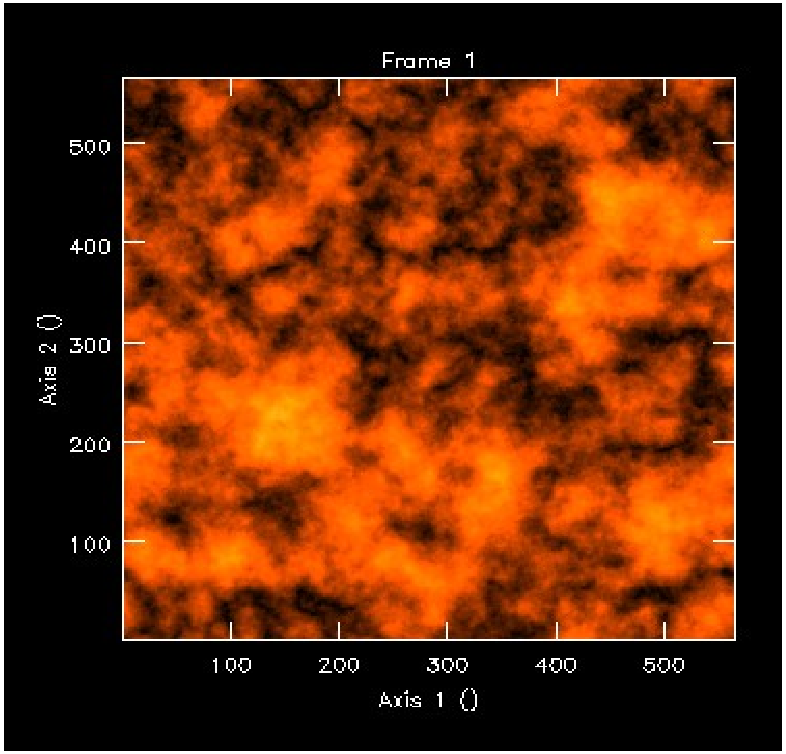


Ionospheric Phase Screen

Examples of DD effects

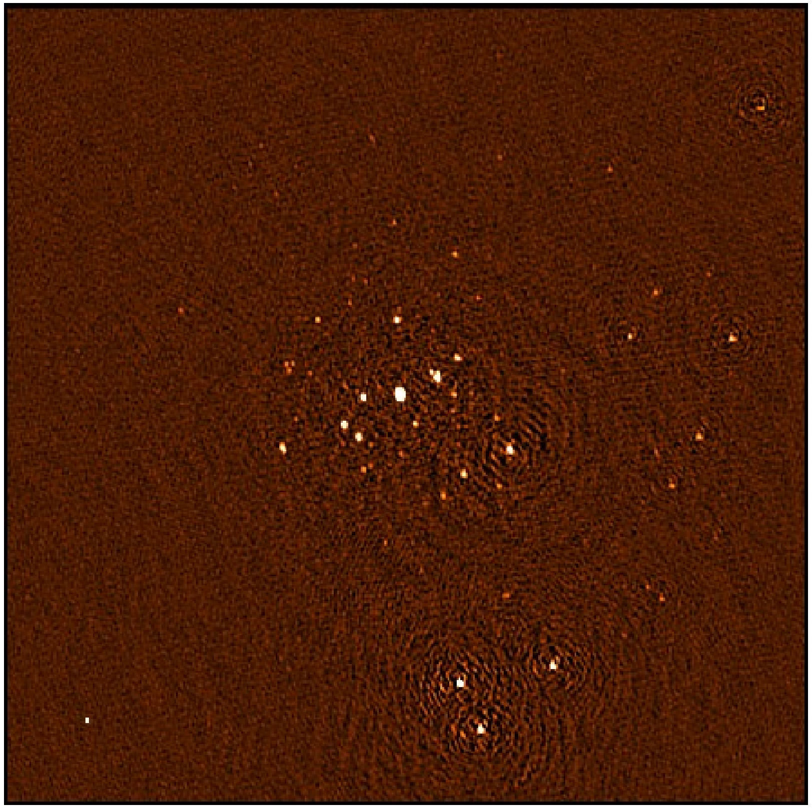


Time and DD Primary Beam: LWA

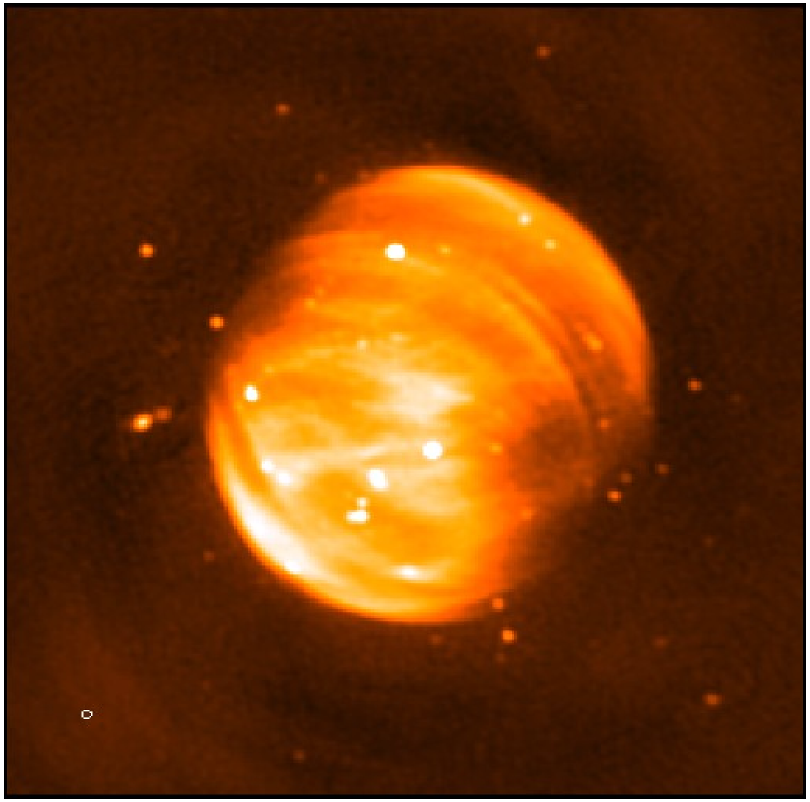


Ionospheric Phase Screen

Range of imaging challenges



Field with compact sources filling the FoV



Compact + extended emission filling the FoV

Used mostly auto-flagging + some manual flagging

Dominant DD Effects

- Time varying PB effects
  - All frequencies: Rotation with Parallactic Angle for El-Az mount antennas (GMRT, EVLA, ALMA)
  - All telescopes: Pointing errors, structural deformation
    - Projection effects (Aperture Array elements)
  - Frequency and polarization dependence (most telescopes)
  - Heterogeneous antenna arrays

- Algorithm development approach taken
  - Algorithm R&D (SNR per DoF, error propagation, computing requirements,...)
  - Proof-of-concept tests with realistic simulations
  - Apply to real data to test computing and numerical performance

Parametrized Measurement Equation

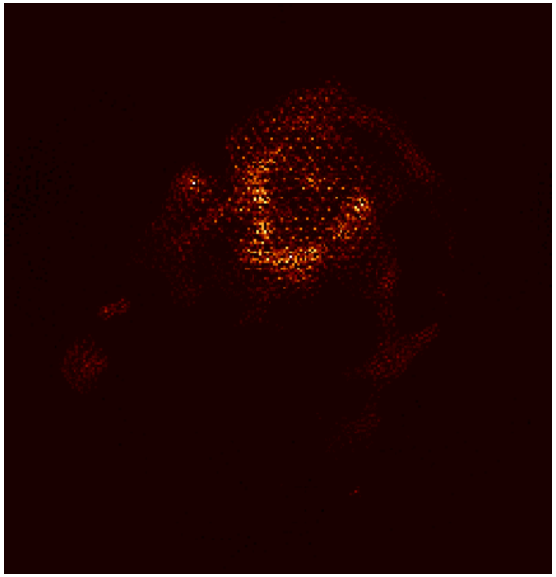
- Need more sophisticated parametrization of the ME
  - Better parametrization of the J_i , J_i^S *and the* Sky *(I^M)*
  - Solver for the (unknown) parameters
  - Forward and reverse transform that account for the DD terms
  - Efficient run-time implementation
- Useful parametrization:
  - Which models the effects well and with minimum DoF
  - For which efficient solvers can be implemented
  - Which optimally utilizes the available SNR
- Noise on the solved parameters:
  - Calculate Covariance matrix for $V_{ij} - V_{ij}^M(p_k, I^M) = \sigma_{ij}$

$$\sigma(p) = \left[\frac{2 k_b T_{sys}}{\eta_a A \sqrt{N_{ant} \nu_{corr} \tau_{corr}} \sqrt{N_{SolSamp}}} \right] \frac{1}{S}$$

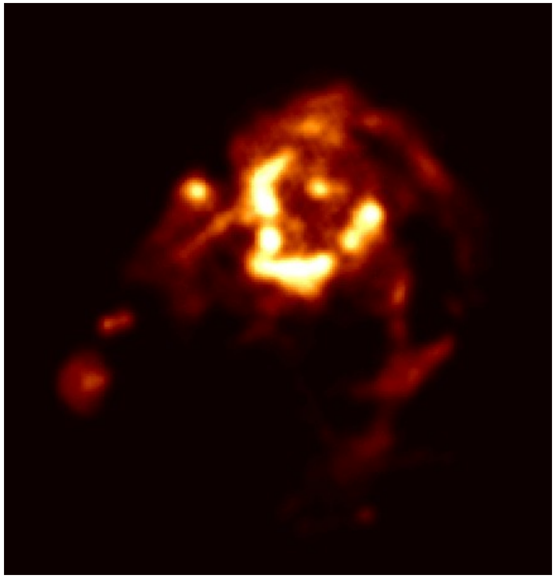
$$\text{where } S = \int \frac{\partial E_i(\mathbf{s}, p)}{\partial \mathbf{s}} E_j^*(\mathbf{s}, p) I^M(\mathbf{s}) e^{2\pi\iota \mathbf{s}.\mathbf{b}_{ij}} d\mathbf{s}$$

Deconvolution: Parametrization of the emission

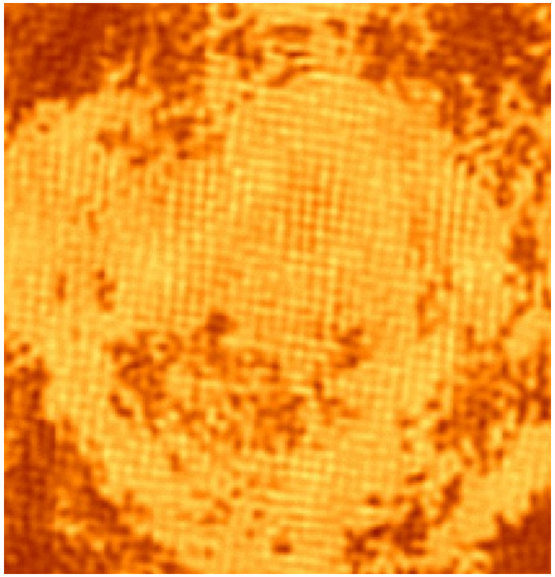
- Scale-less deconvolution algorithms:
  - $I^M = \sum_k A_k \delta(x - x_k)$:Treat each pixel as an independent DoF
  - CLEAN (and its variants), MEM (and its variants)



Component Model



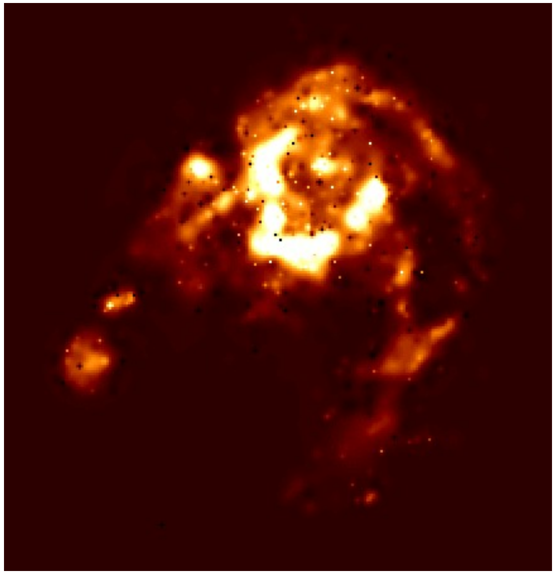
Restored Model



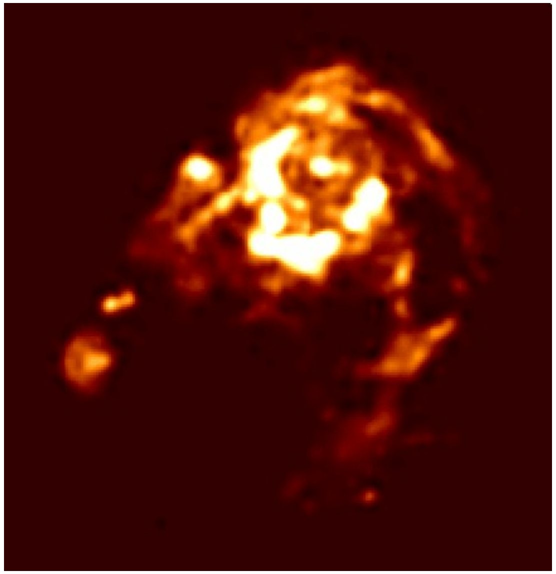
Residuals

Deconvolution: Parametrization of the emission

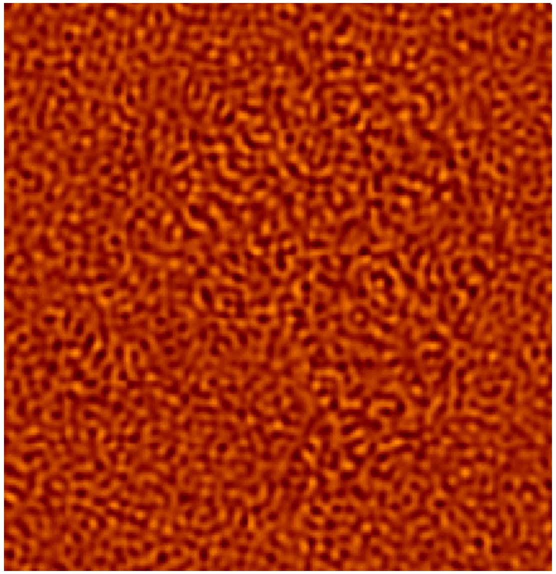
- Scale-sensitive deconvolution algorithms:
  - $I^M = \sum_k A_k f(Scale, Position)$:Decompose the image in a scale-sensitive basis
  - Asp-Clean (A&A, 747, 2004 (astro-ph/0407225), MS-Clean (IEEE JSPSP, Vol. 2, No. 5, 2008)



Component Model

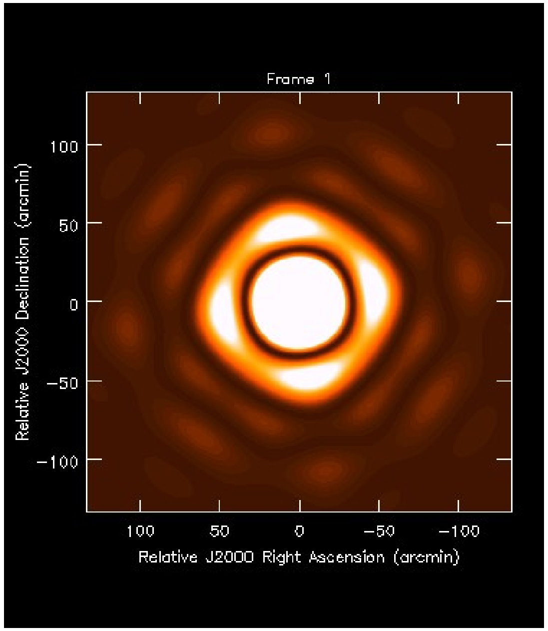


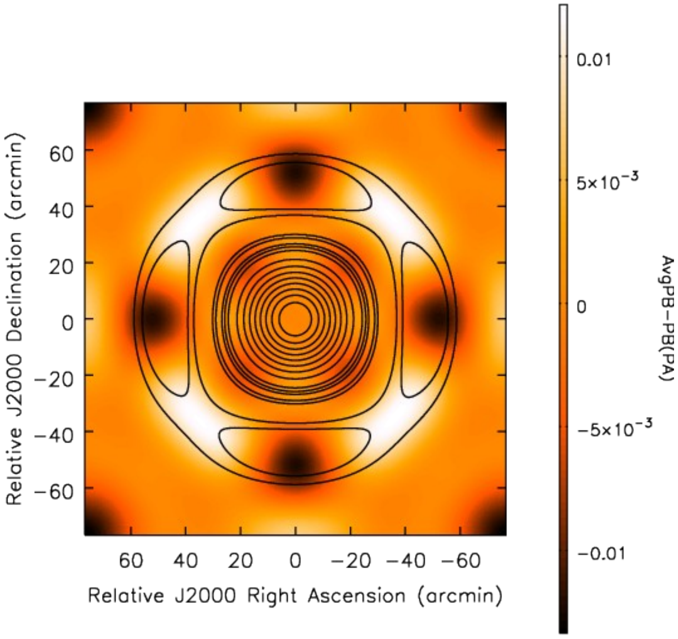
Restored Model



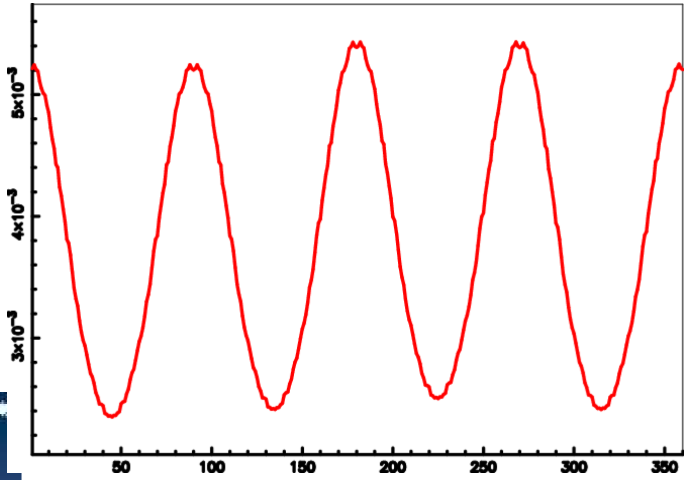
Residuals

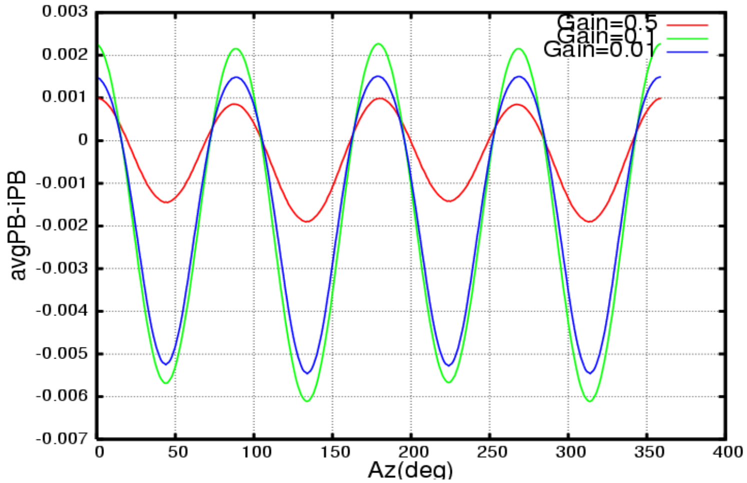
Time varying DD gains due to PB





$\Delta\, PB$



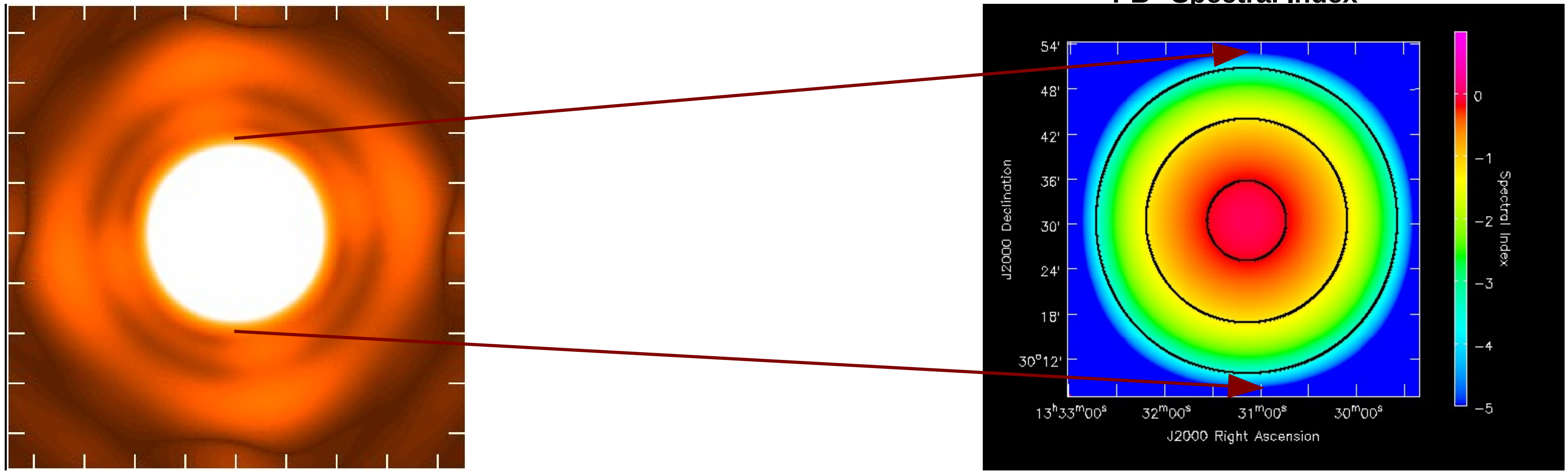


$$\Delta\, I = PSF * (I\, \Delta\, PB)$$

Wide-band PB effects

- Frequency dependence of the PB is a first order effect for wide-band observations

PB "Spectral Index"



  - Is it $M(s,\nu)$ or is it $I(s,\nu)$?
  - Fundamental separation: Include PB as part of the measurement process (include its effect as part of forward and reverse transforms)

$$V(b_{ij}) \quad = \quad \int M_{ij}^{S} I(s)\, e^{2\pi\iota\left(b_{ij}.s\right)}\, ds$$

Full beam imaging

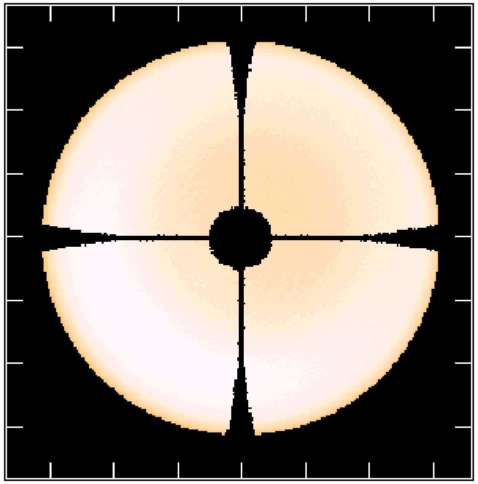
- Limits due to the rotation of asymmetric PB
  - Max. temporal gain variations @ ~10% point
  - DR limit: few X 10^4:1

- Limits due to antenna pointing errors
  - In-beam error signal max. @ 50% point
  - DR limit: few X 10^4:1
  - Limits for mosaicking would be worse
    - Significant flux at half-power point
    - Significant flux in the side-lobes for most pointing

The A-Projection algorithm

$$V^o(u,v,w)=V^M(u,v)*G(u,v;Time,Poln.)$$

- Modified forward and reverse transforms:
  - No assumption about sky properties
  - Spatial, time, frequency and polarization dependence naturally accounted for
  - Done at approximately FFT speed



Model for EVLA aperture illumination (real part)

One element of the Sky-Jones (Jones Matrix per pixel)

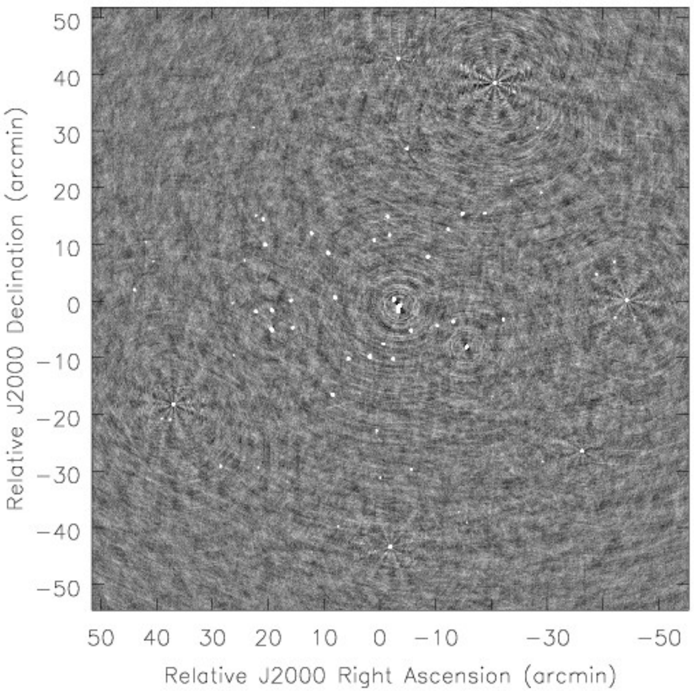
- Combining with W-Projection or image plane part of the various deconvolution algorithms is straight forward (algorithm complexity is lower)

- Efficient solvers to solve for more precise parametrized models (Pointing SelfCal and its extensions)

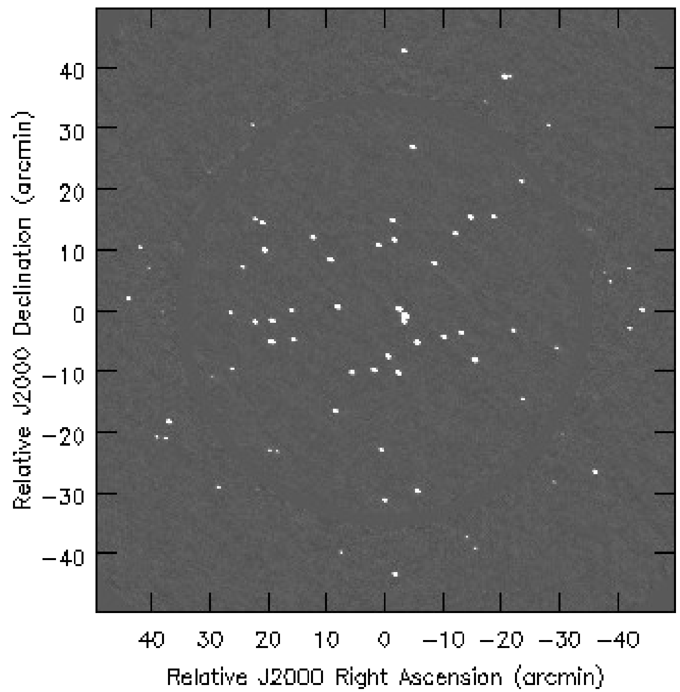
A-Projection algorithm, A&A 2008

A-Projection algorithm: Simulations

Before Correction



After Correction

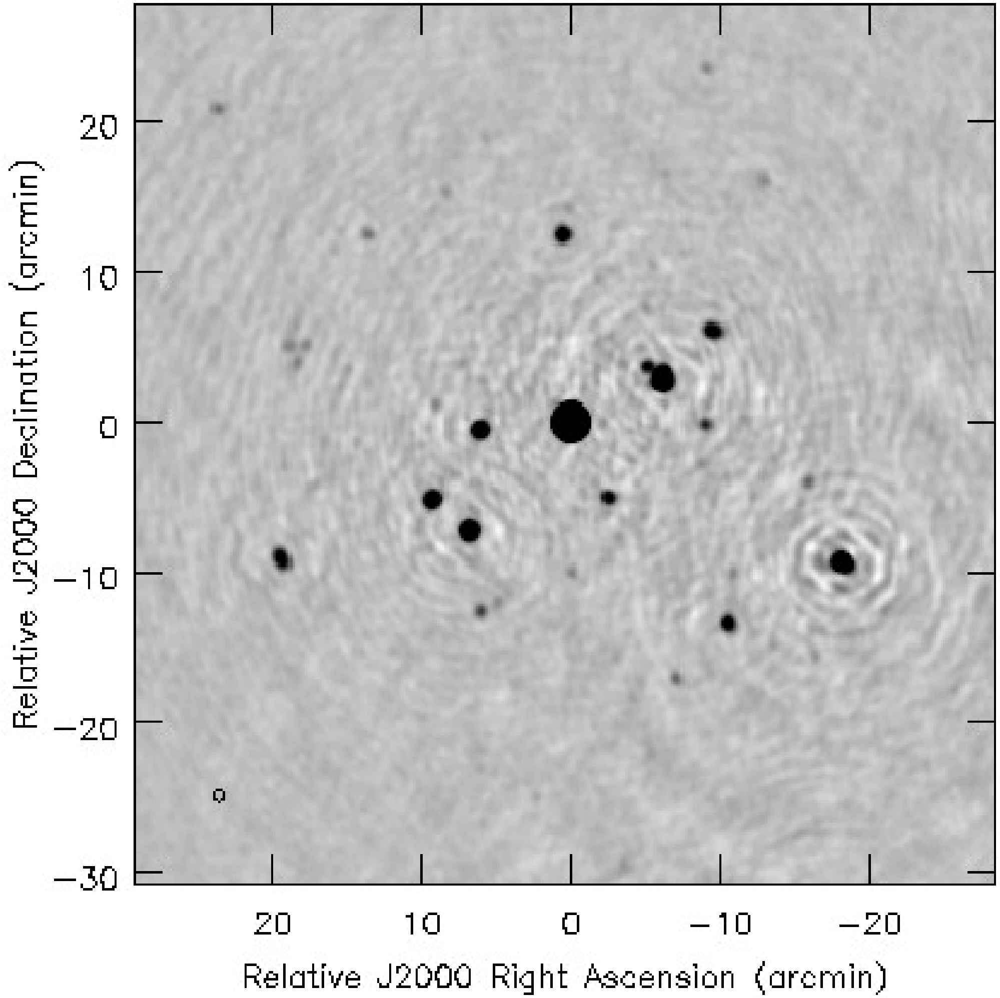


$$\textit{Minimize}: V_{ij}^{O} - E_{ij} * [F I^{M}] \quad \textit{w.r.t.} \quad I^{M}$$

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

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

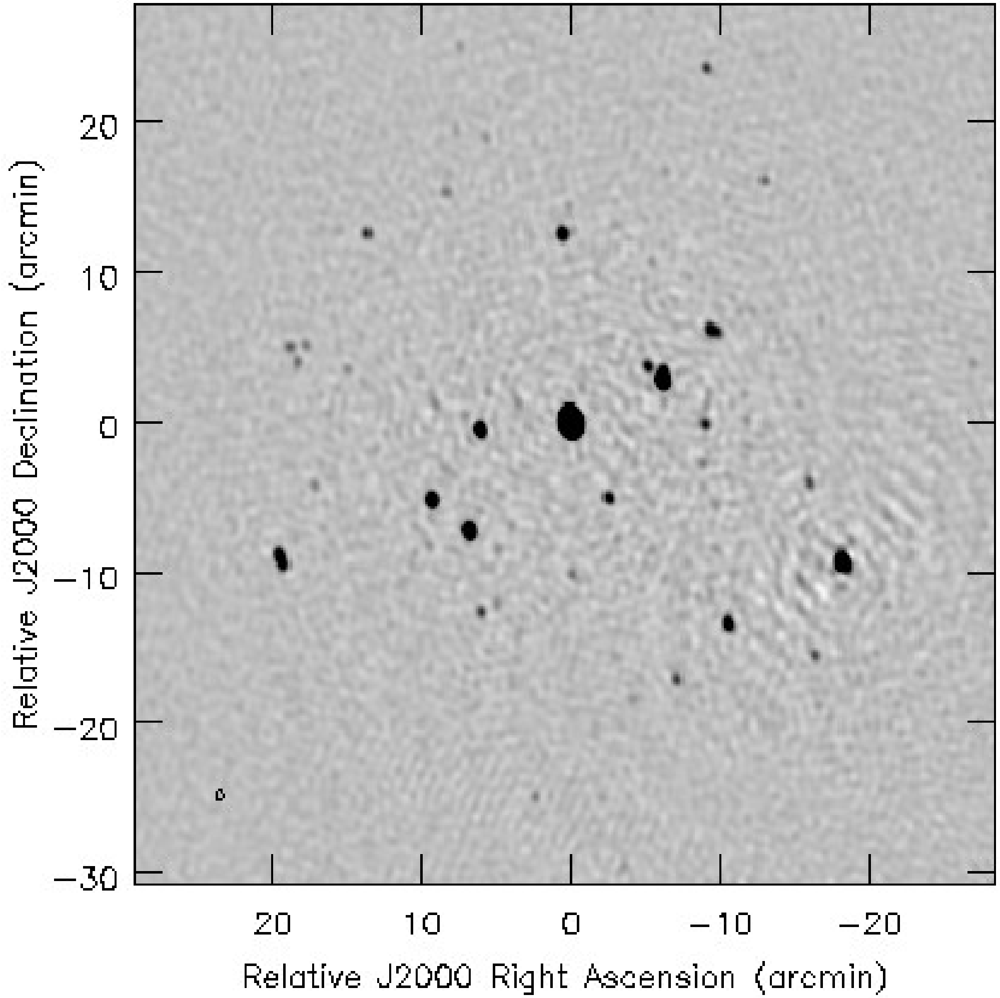
EVLA L-Band Stokes-I: Before correction



- 3C147 field at L-Band with the EVLA
- Only 12 antennas used
- Bandwidth: 128 MHz
- ~7 hr. integration

- Dynamic range: ~700,000:1

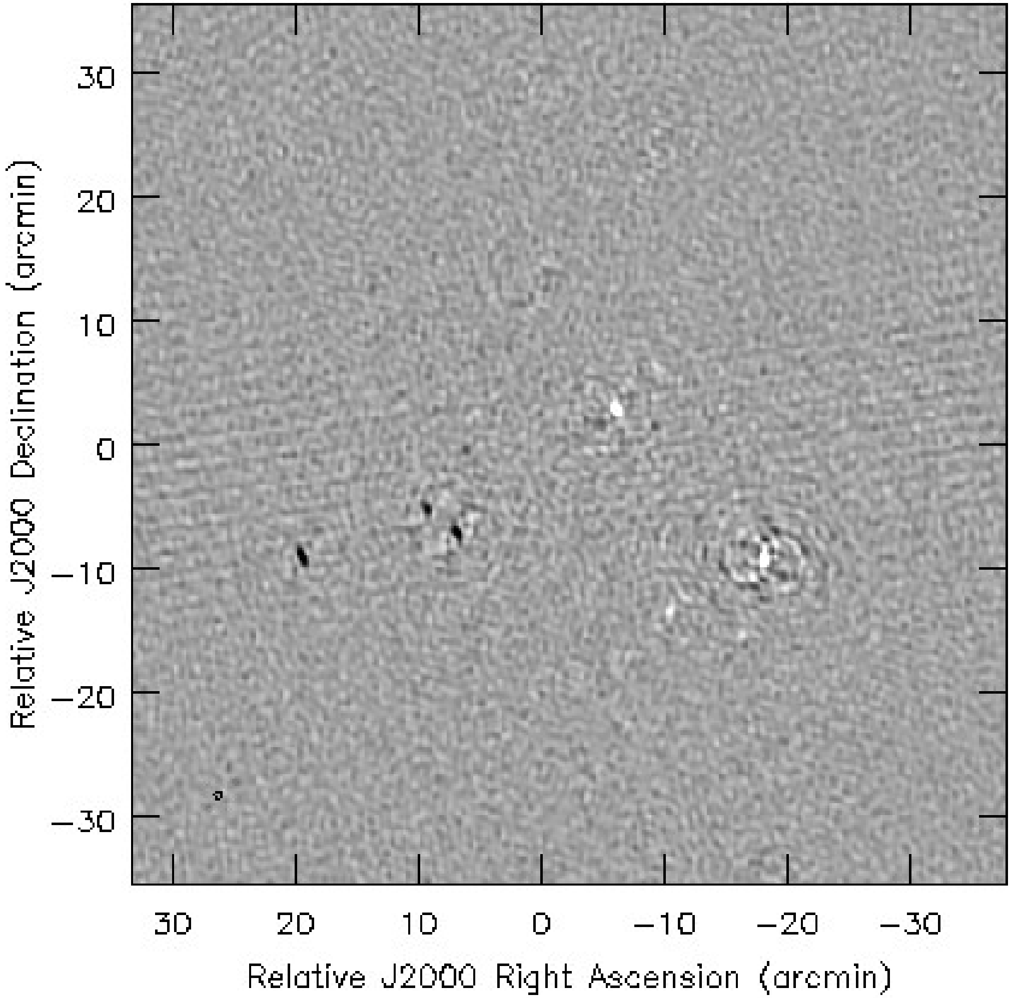
EVLA L-Band Stokes-I: After correction



- 3C147 field at L-Band with the EVLA
- Only 12 antennas used
- Bandwidth: 128 MHz
- ~7 hr. integration

- Dynamic range: ~700,000:1

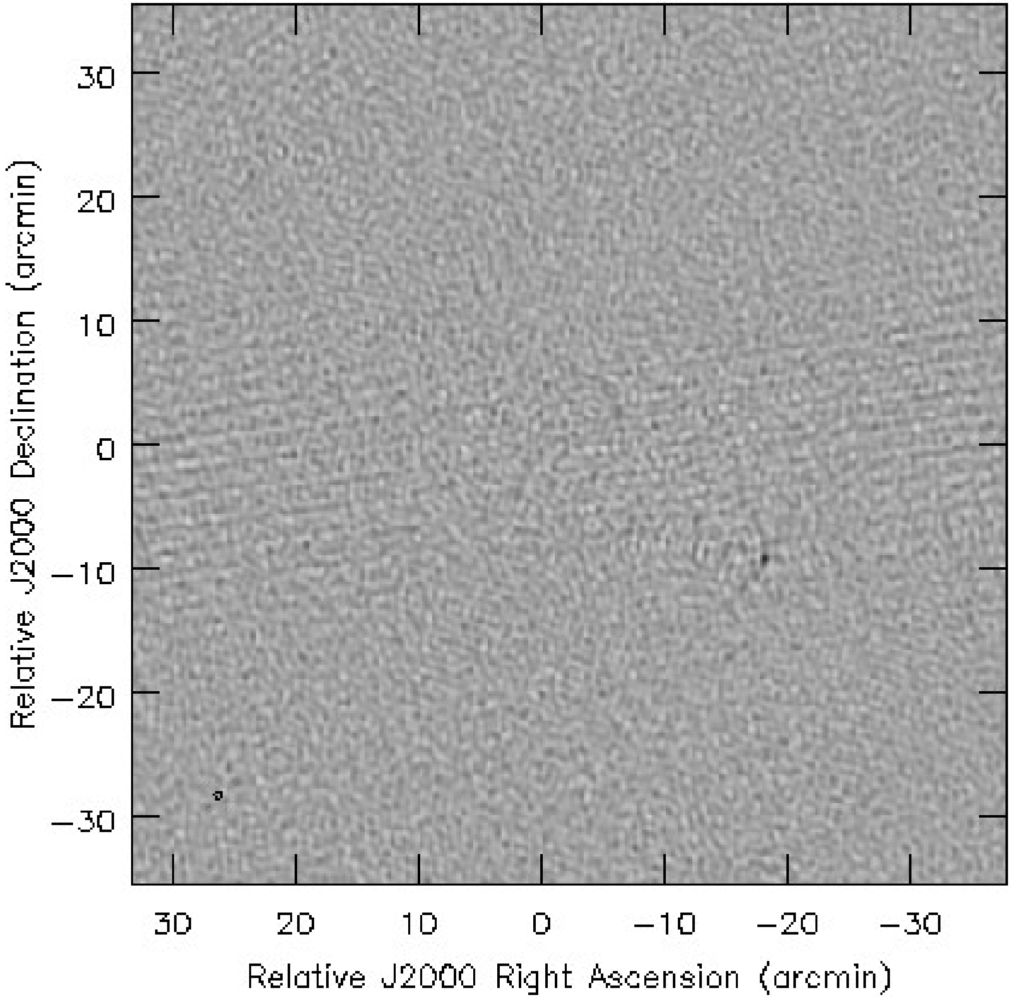
EVLA L-Band Stokes-V: Before correction



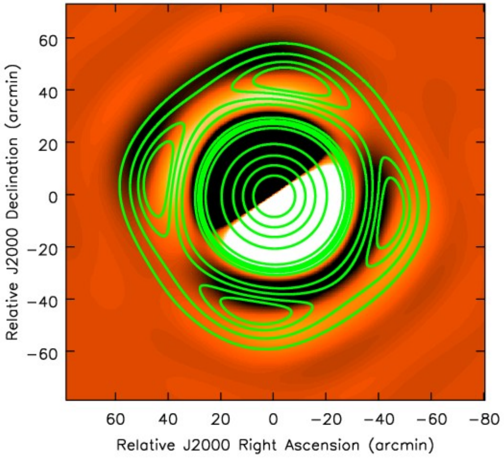
Is it $M(s, Poln)$?

Or is it $I(s, Poln)$?

EVLA L-Band Stokes-I: After correction

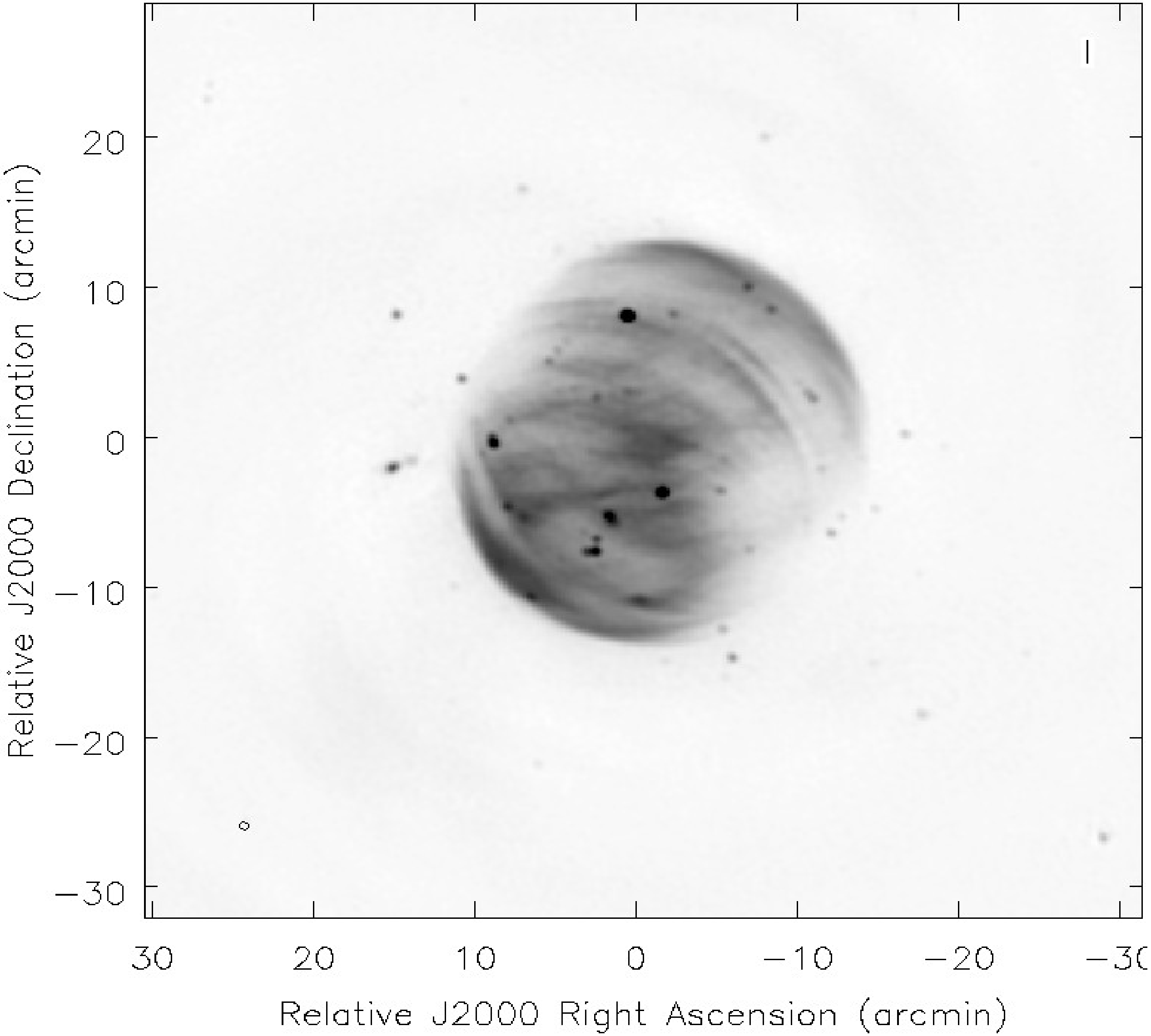


Use physical model for the Stokes-V pattern:



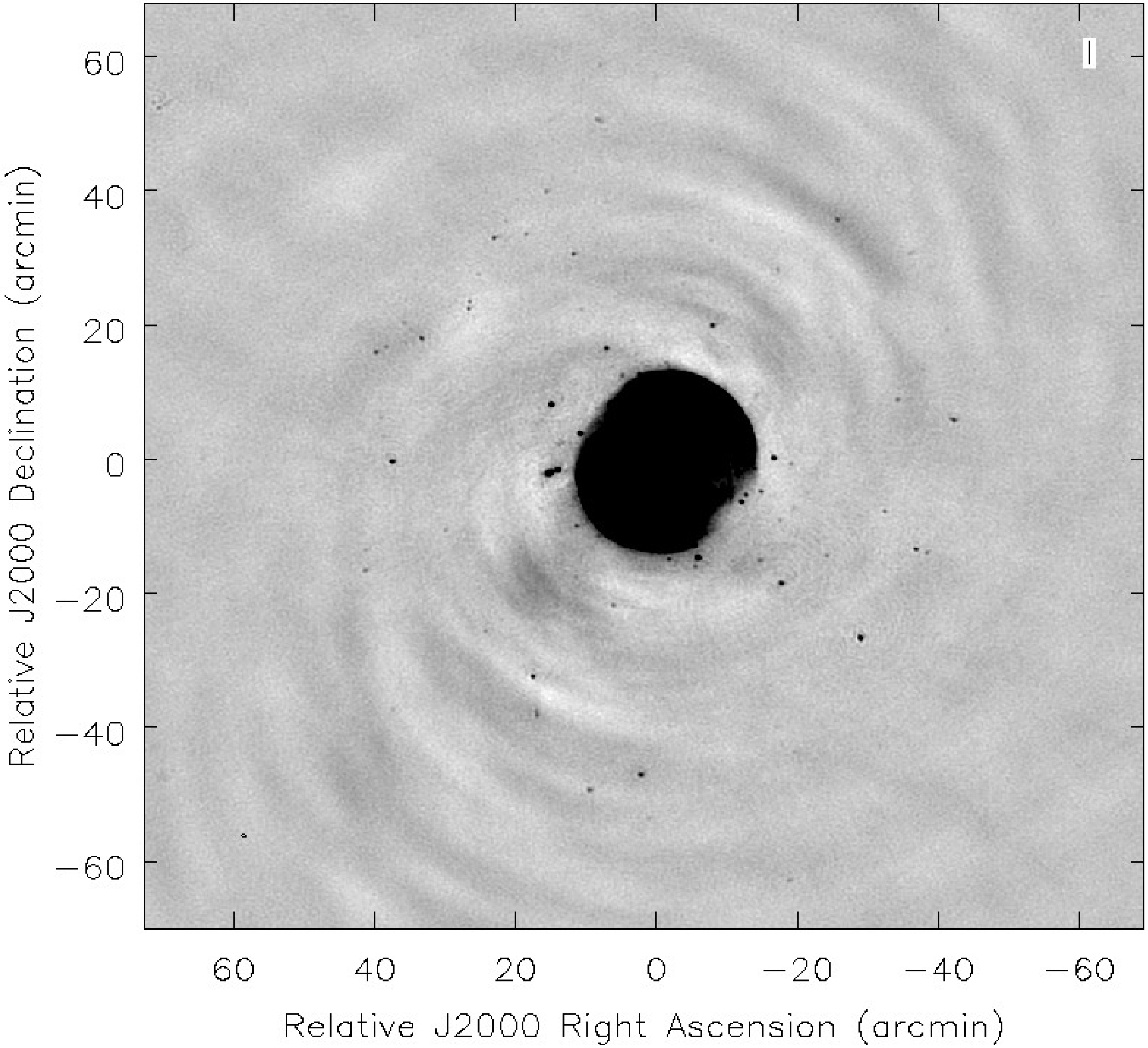
Contours: Stokes-I power pattern
Colour: Stokes-V power pattern

Wide band imaging with the EVLA

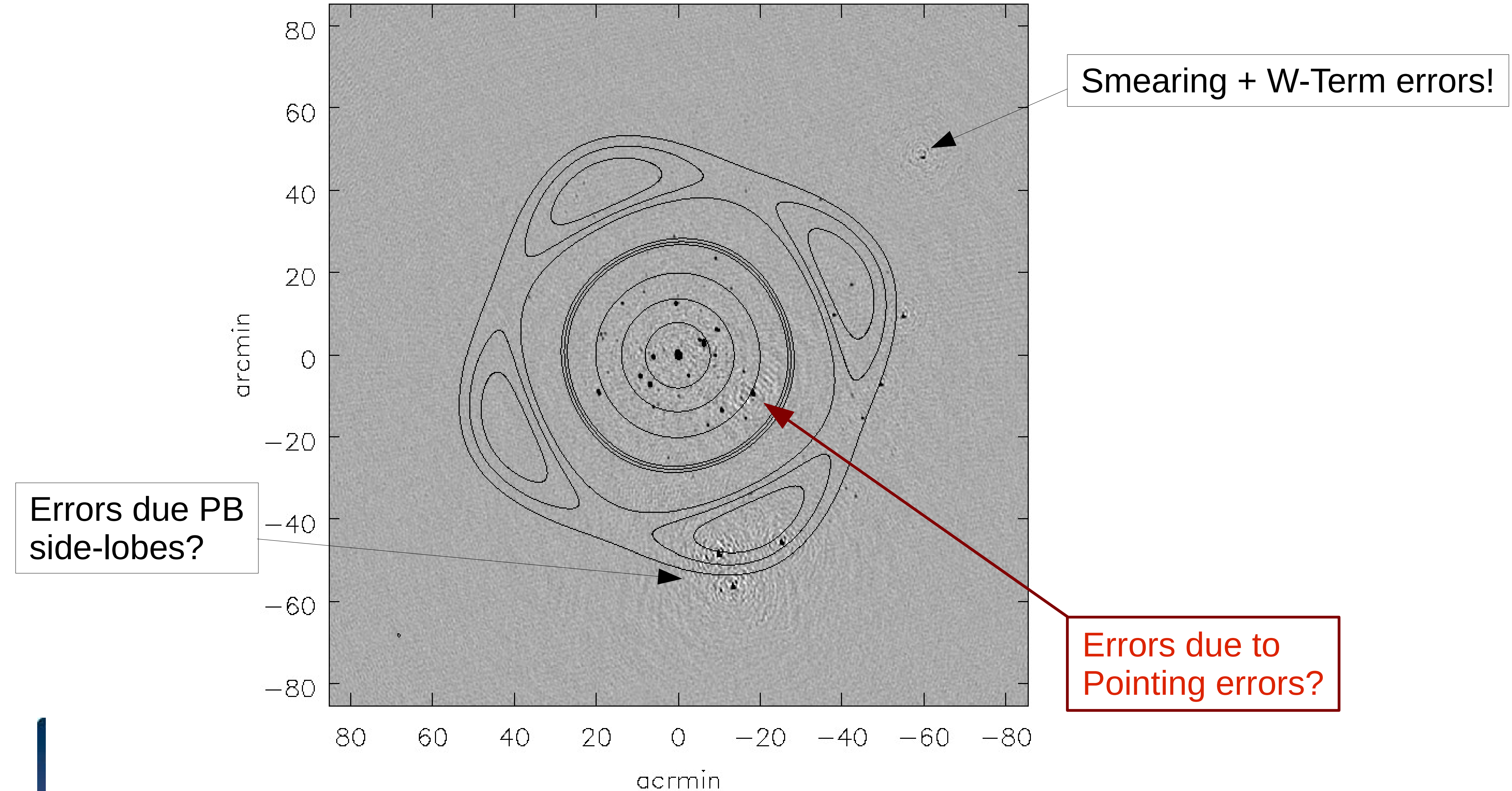


- 1.2-1.8GHz
- ~40 microJy/Beam
- RSRO Projects
  (AB1345, Bhatnagar et al.
  AT374, Taylor et al.)
- Scientific goals
  Spectral Index imaging
  RM Synthesis

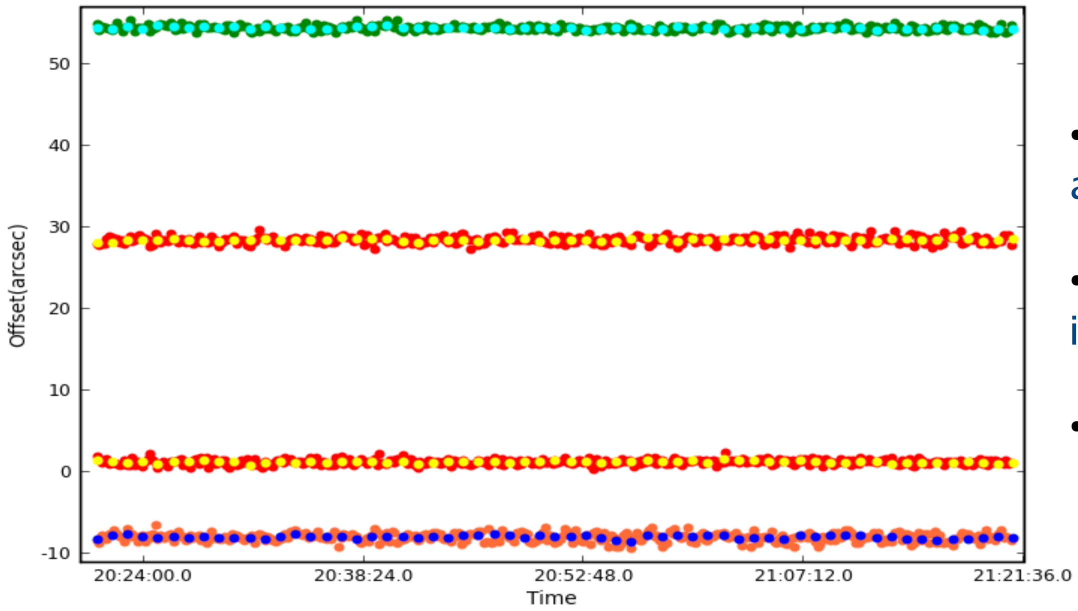
Wide band imaging with the EVLA



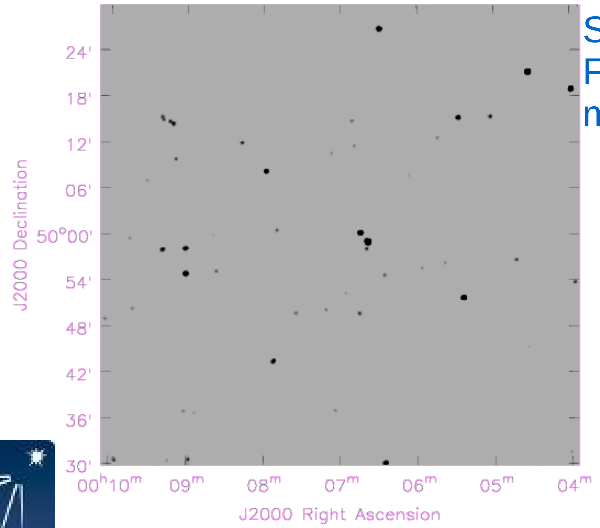
3C147: Residual errors in full field



DD SelfCal algorithm: Simulations



- Typical antenna pointing offsets for VLA as a function of time
- Over-plotted data: Solutions at longer integration time
- Noise per baseline as expected from EVLA



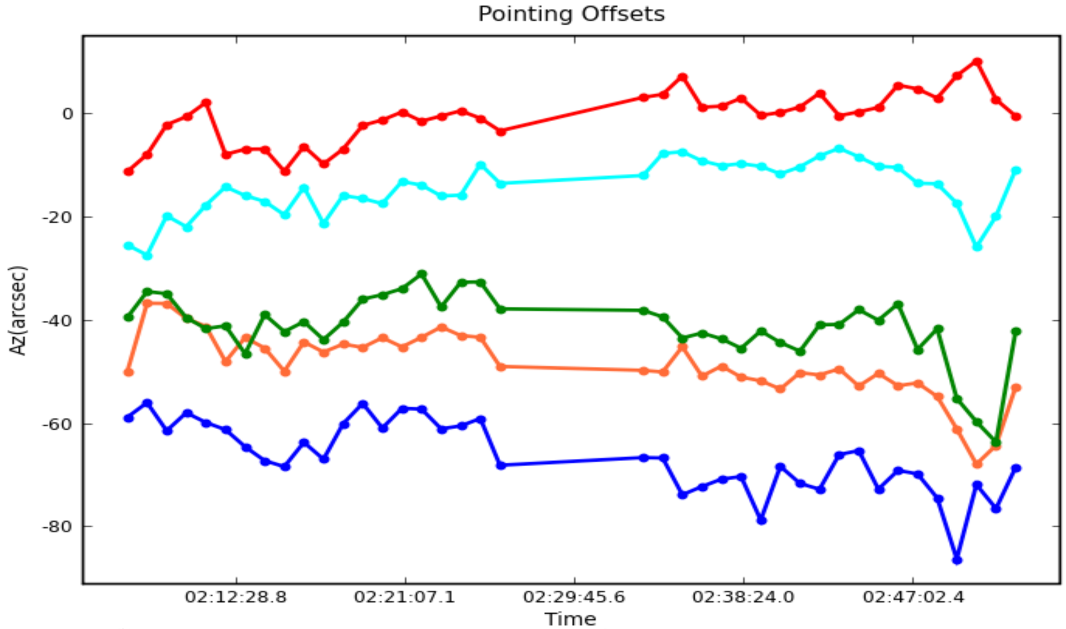
Sources from NVSS. Flux range ~2-200 mJy/beam

$$\textit{Minimize}: V_{ij}^{O} - E_{ij} * V_{ij}^{M} \quad w.r.t. \quad E_i$$

$$\left[\frac{\partial E_{ij}(p_i^k, p_j^k)}{\partial E_i} \frac{\partial E_i}{\partial p_i^k} \right] * V_{ij}^{M} = 0$$

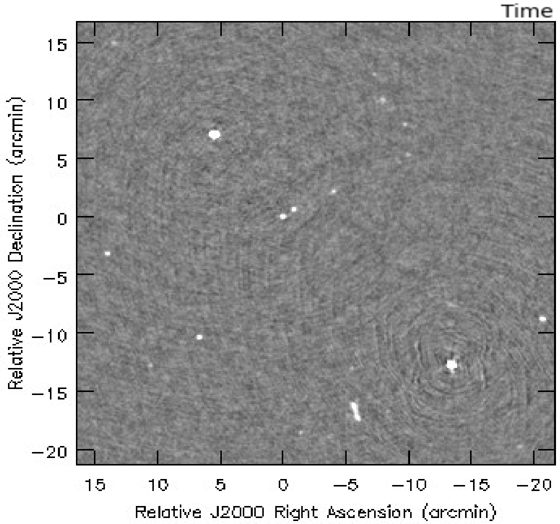
[Bhatnagar et al., EVLA Memo 84, 2004]

DD SelfCal algorithm: EVLA Data



- Typical solved pointing errors for a few antennas
- Solution interval: 2min

- Using ~300 MHz of bandwidth.



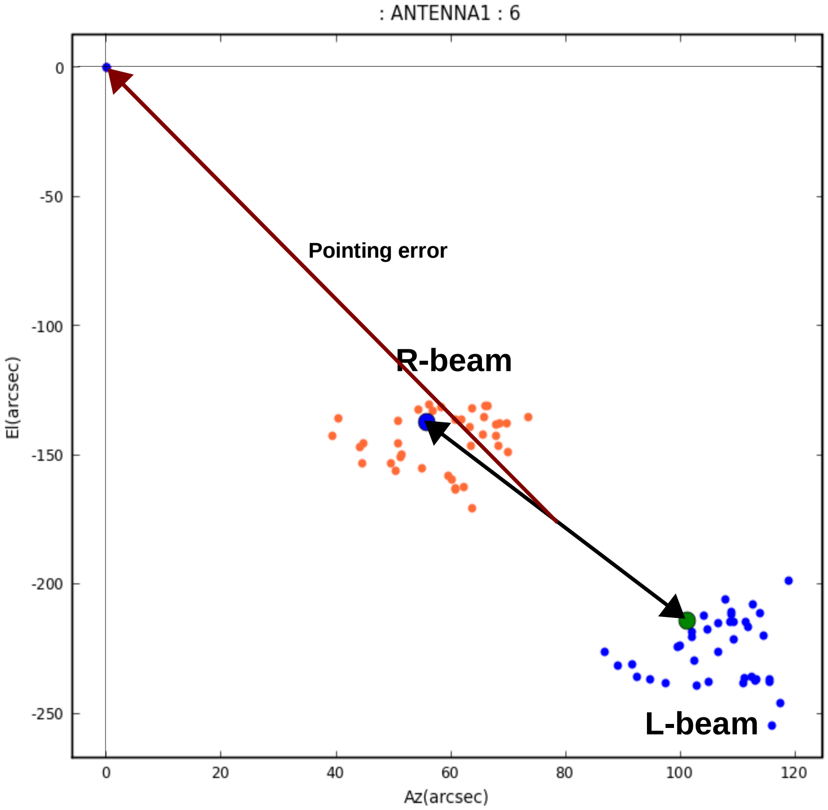
IC2233 field in the 1-2 GHz band with the EVLA

$$Minimize: V^O_{ij} - E_{ij} * V^M_{ij} \quad w.r.t. \quad E_i$$

$$\left[\frac{\partial E_{ij}(p_i^k, p_j^k)}{\partial E_i} \frac{\partial E_i}{\partial p_i^k}\right] * V^M_{ij} = 0$$

[Bhatnagar et al. (paper in preparation)]

DD SelfCal algorithm: EVLA Data



- El-Az mount antennas
- Polarization squint due to off-axis feeds
  - The R- and L-beam patterns have a pointing error of +/- ~0.06 $\frac{\lambda}{D}$

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

$$\sigma(p) = \left[\frac{2 k_b T_{sys}}{\eta_a A \sqrt{N_{ant} \nu_{corr} \tau_{corr}} \sqrt{N_{SolSamp}}} \right] \frac{1}{S}$$

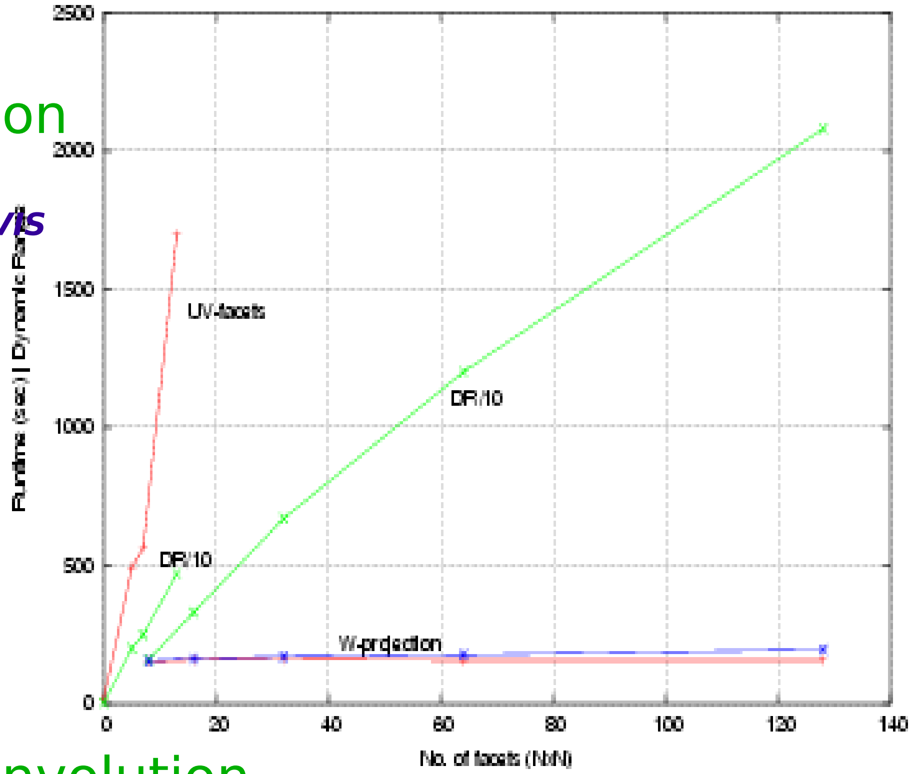
$$\text{where } S = \int \frac{\partial E_i(\mathbf{s}, p)}{\partial \mathbf{s}} E_j^*(\mathbf{s}, p) I^M(\mathbf{s}) e^{2\pi\iota \mathbf{s}.\mathbf{b}_{ij}} d\mathbf{s}$$

- EVLA polarization squint solved as pointing error (optical pointing error).
- Squint would be symmetric about the origin 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.

[paper in preparation]

Computing load

- Scaling laws for imaging
  - Non co-planar baseline correction
  - W-Projection: $\boldsymbol{(N^2_{wproj} + N^2_{GCF})N_{vis}}$
  - Faceting: $\boldsymbol{N^2_{facets} N^2_{GCF} N_{vis}}$



- Combine with Scale-sensitive deconvolution
- N_{vis}: $10^{10\text{-}12}$, N^2_{GCF}: 7-50 , N_{comp}: $10^{4\text{-}5}$

I/O load

- Near future data volume (0-1 years)
  - Recent data with the EVLA: 100-500 GB
- Next 5 years
  - 100X increase (in volume and effective I/O)

- Non-streaming data processing
  - Expect 20-50 passes through the data (flagging + calibration + imaging)
    - Effective data i/o: few TB
  - Exploit data parallelism
    - Distribute normal equations (SPMD paradigm looks promising)
  - Deploy *computationally efficient* algorithms (‘P’ of SPMD) on a cluster

Computing challenges

- Calibration of direction dependent terms
  - As expensive as imaging
- Significant increase in computing for wide-field wide-band imaging
  - E.g. convolution kernels are larger (up to 50x50 for single facet EVLA A-array, L-band imaging)
  - E.g. Multiple terms for modeling sky and aperture for wide-band widths
- Terabyte Initiative: 4K x 4K x 512 x 1Pol tests using 200 GB data set
  - Timing
    - Simple flagging : 1h
    - Calibration (G-Jones) : 2h15m
    - Calibration (B-Jones) : 2h35m
    - Correction : 2h
    - Imaging : 20h
  - Compute : I/O ratio : 2:3

Parallelization: Initial results

- Continuum imaging: (No PB-correction or MFS)
  - Requires inter-node I/O (Distribution of normal equations)
  - Dominated by data I/O
  - 1024 x 1024 imaging: (Traditional CS-Clean; 5 major cycles)
    - 1-node run-time : 9hr
    - 16-node run-time : 70min (can be reduced up to 50%)
      : 60min (MS-Clean)
      (residual CPU-power available for projection algorithms)
- Imaging deconvolution is most expensive step
- DD Calibration as expensive as a deconvolution major-cycle
  - CPU bound (a good thing!)

[Golap, Robnett, Bhatnagar]

Parallelization: System Design

- Matching data access and in-memory grid access patterns is critical
- Optimal data access pattern for imaging and calibration are in conflict
  - Freq-Time ordered data optimal for imaging
  - Time-ordered data optimal for calibration
- SS deconvolution + MFS might make FLOPS per I/O higher: A good thing!
- Production Cluster
  - 32 nodes, 2x4 cores, 12 GB RAM, InfiniBand
  - Data served via a Luster FS
    - Measured I/O throughput: 800-900 MB/s
  - Multiple processes per node gets I/O limited
  - I/O handler separated from compute processes

General comments

- Algorithms with higher Compute-to-I/O ratio
  - Moor's law helps
- Pointing SelfCal and MS-MFS solutions demonstrate the need for minimizing the DoF per SNR (?)
- Exact solutions in most cases is a mathematical impossibility
  - Iterative solvers are here to stay: Image deconvolution, calibration
  - Baseline based quantities are either due to sky or indistinguishable from noise.
    - Modeling of calibration terms is fundamentally antenna-based
- Data rate increasing at a faster rate than i/o technology
  - Moor's law does not help!
    - Moving 100s of GB EVLA data can take up to a weak
    - More time spent in i/o-waits than in computing
- Need for robust algorithms for automated processing that also benefit from and can be easily parallelized
- Need for robust pipeline heuristic