Direction Dependent Calibration

Wed. Lunch, Socorro, Aug. 22nd 2010

S. Bhatnagar

NRAO

High sensitivity imaging

- Sensitivity $\propto \frac{N_{ant}(\eta A_{ant})\sqrt{(N_t \tau)(N_{chan} \Delta \nu)}}{T_{sys}}$
- 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
    - Smaller scale variations over larger parameters space to be accounted for
  - $\mathbf{N_{channels}}$ = 1-10GHz/KHz-MHz and $\mathbf{N_t}$ = 10hr/(1-10sec)
  - 10-100x increase in the number of samples to achieve the required sensitivities
    - Algorithm efficiency remains a critical parameter
  - Wider range of angles on the sky (→ Direction Dependence)

NRAO

Synthesis Imaging

- Measurement Equation

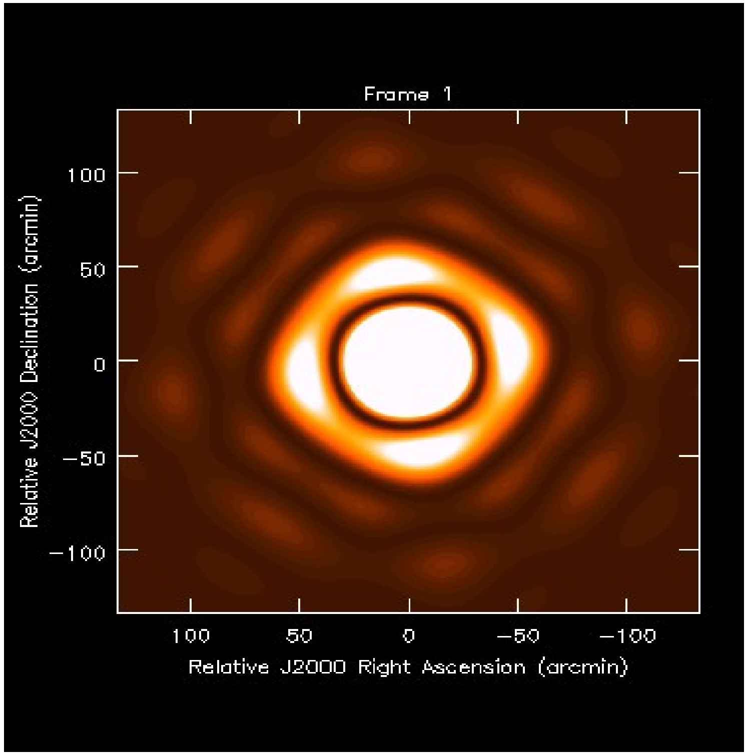
$$V_{ij}^{Obs}(\nu) = M_{ij}(\nu,t) W_{ij} \int M_{ij}^{S}(s,\nu,t) I(s,\nu) e^{2\pi\iota(b_{ij}.s)} ds$$

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

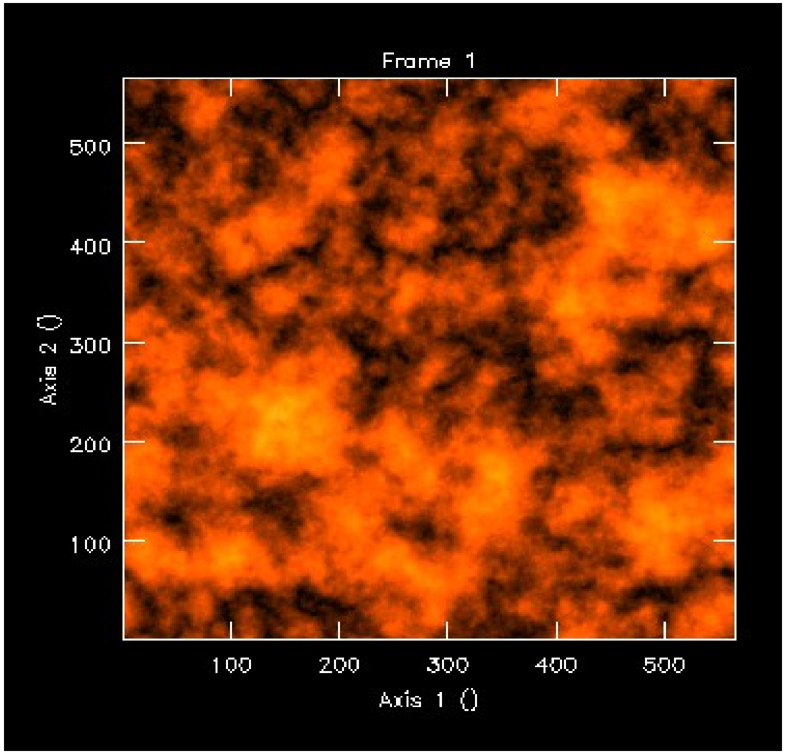
$M_{ij}^{S}(s,\nu,t) = J_i(s,\nu,t) \otimes J_j^*(s,\nu,t)$:Direction dependent (DD) gains

- 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.

Examples of DD effects

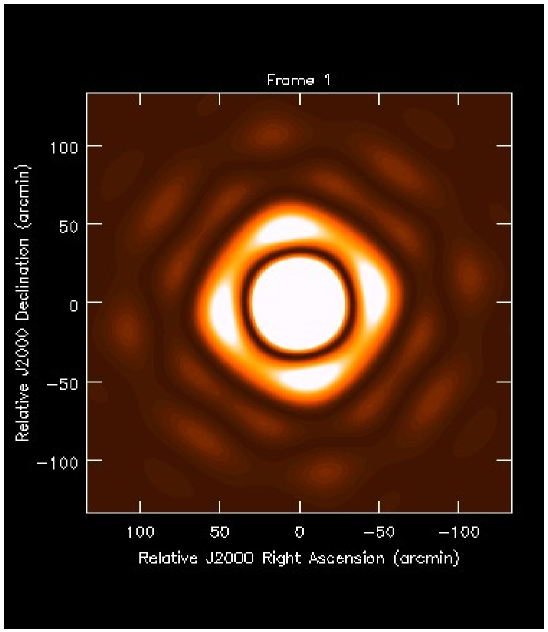


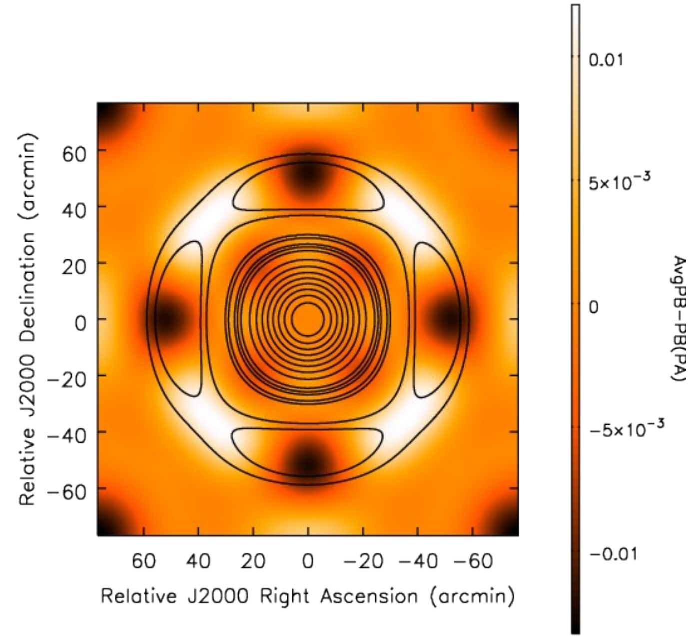
Time and DD Primary Beam: EVLA



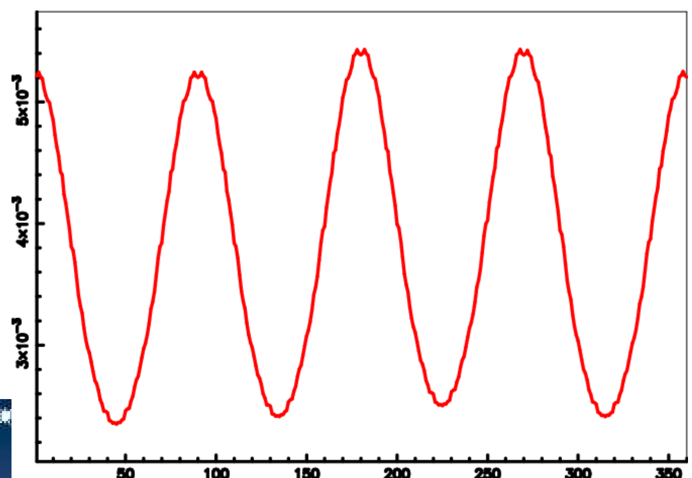
Ionospheric Phase Screen

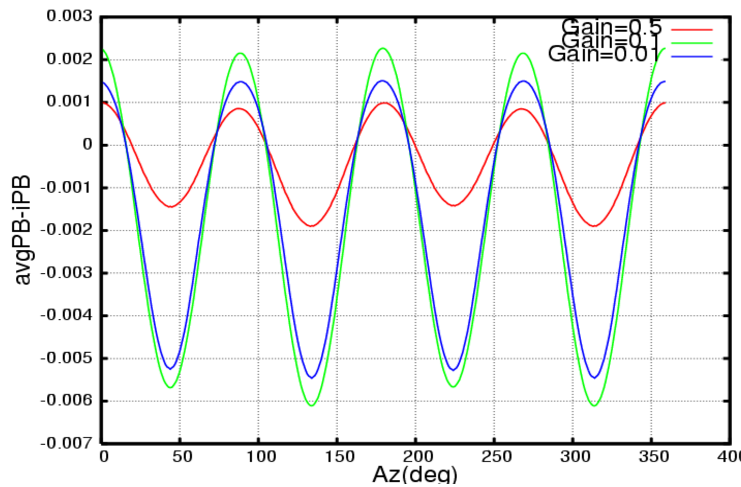
Time varying DD gains due to PB





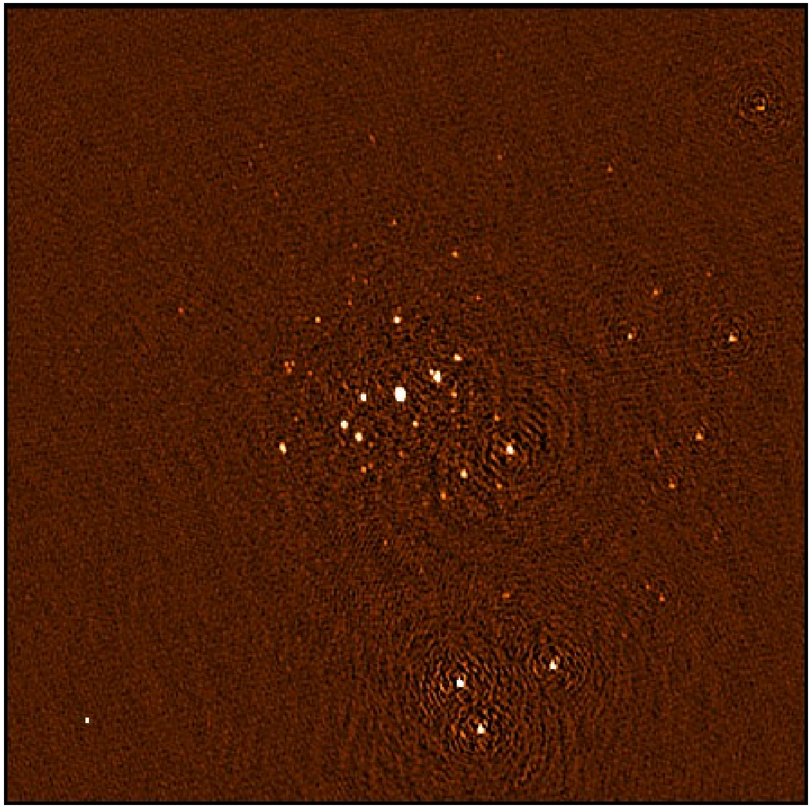
ΔPB



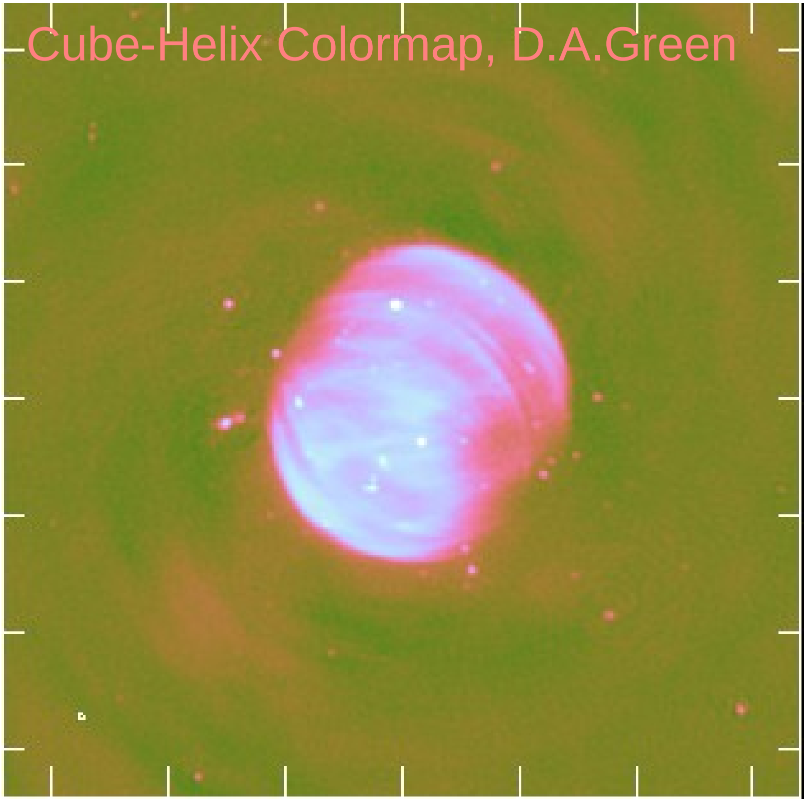


$$\Delta I = PSF * (I \Delta PB)$$

Range of imaging challenges



Field with compact sources filling the FoV



Compact + extended emission filling the FoV

Used mostly auto-flagging + some manual flagging

Parametrized Measurement Equation

- Need more sophisticated parametrization of the ME
- Two approaches
  - Faceting: Partition the data & apply DI techniques per facet
    - » Issues: + More familiar, trivial parametrization

      – DoF vs. SNR, run-time efficiency, high algorithm complexity
  - Global/Projection methods: Include DD terms in the Measurement Equation
    - » Issues: + Mathematically better formulated, optimal use of the available SNR, run-time efficiency, lower algorithm complexity

      – Less familiar, non-trivial parametrization (?)
- Noise per antenna based DoF:

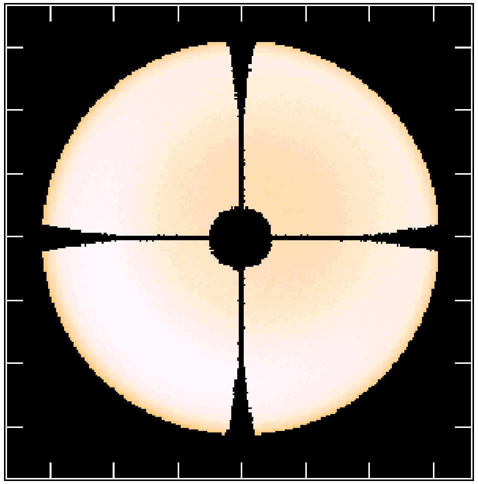
$$\sigma(p) = \left[\frac{2 k_b T_{sys}}{\eta_a A \sqrt{N_{ant} \nu_{corr} \tau_{corr}} \sqrt{N_{SolSamp}}} \right] \frac{1}{S}$$

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

The A-Projection algorithm

$$V^o(u,v,w) = V^M(u,v) * J_i(u,v;s) * J_j^*(u,v;s)$$

- 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)

- A-Projection is the first term of the series expansion of the Aperture Illumination pattern.

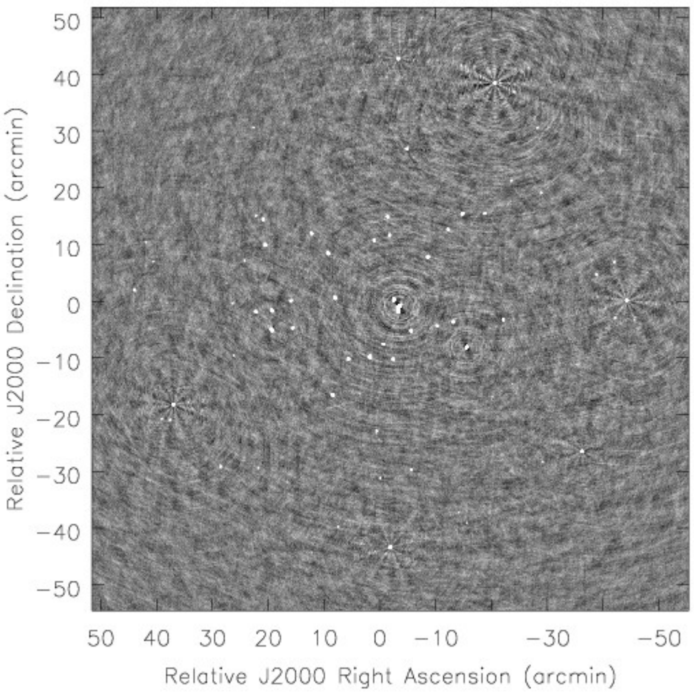
$$A(u) = A_o(u)[1 + a_o Z_o(u) + \ldots]$$

- Projection formulation delivers efficient solvers to solve for parametrized models (Pointing SelfCal and its extensions)

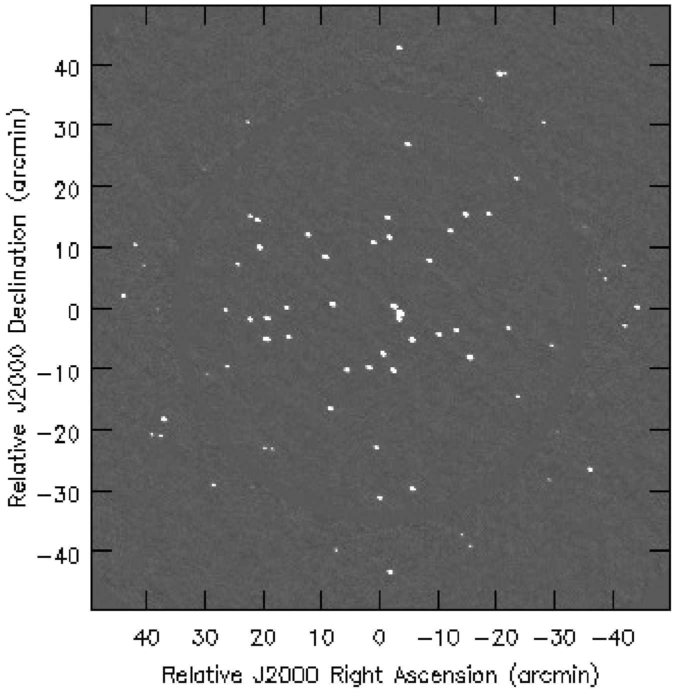
A-Projection algorithm, A&A 2008

A-Projection algorithm: Simulations

Before Correction



After Correction

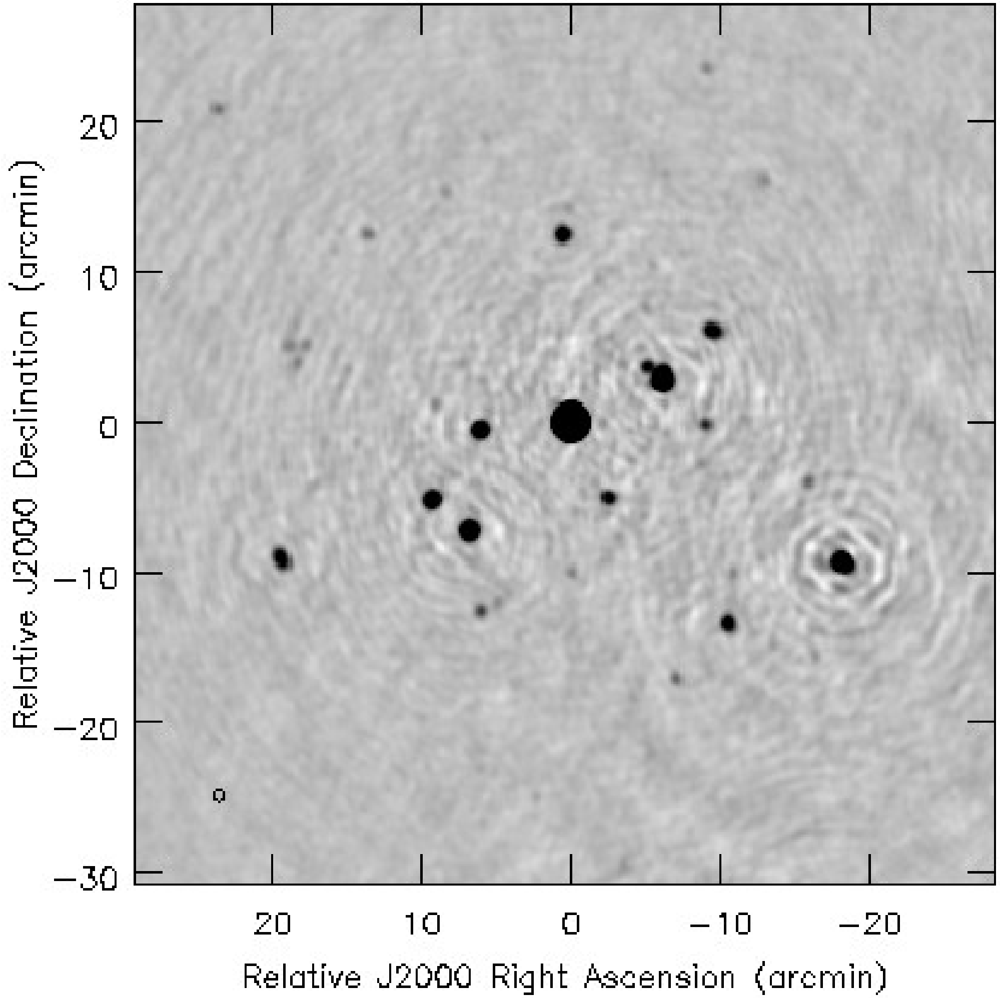


$$\textit{Minimize}: V_{ij}^{O} - E_{ij} * [F I^{M}] \quad w.r.t. \quad I^{M}$$

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

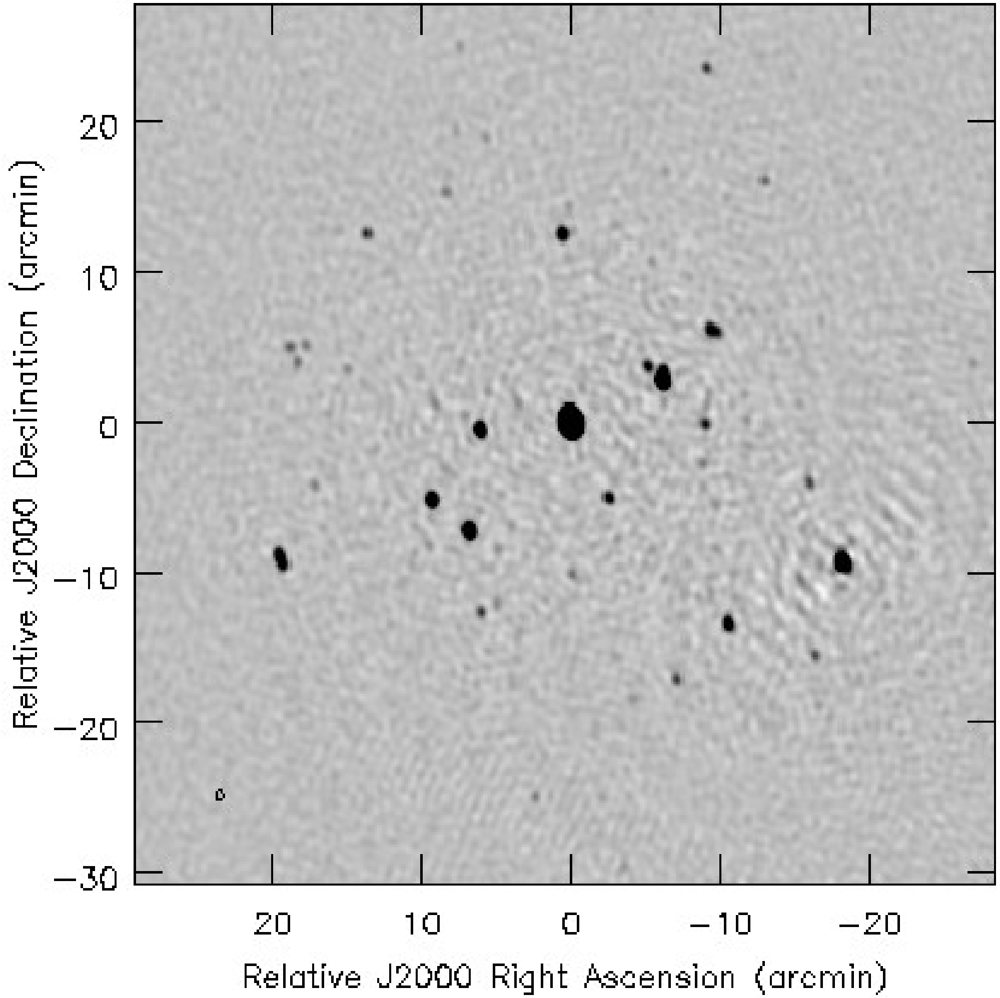
A-Projection: Bhatnagar et al., A&A,487, 2008

EVLA L-Band Stokes-I: Before correction



- 3C147 field at L-Band
- Pre-OSRO Mode
  WIDAR0!
- Only 12 antennas used
- Bandwidth: 128 MHz
- ~7 hr. integration
- Dynamic range: ~700,000:1
- A single baseline based correction was applied

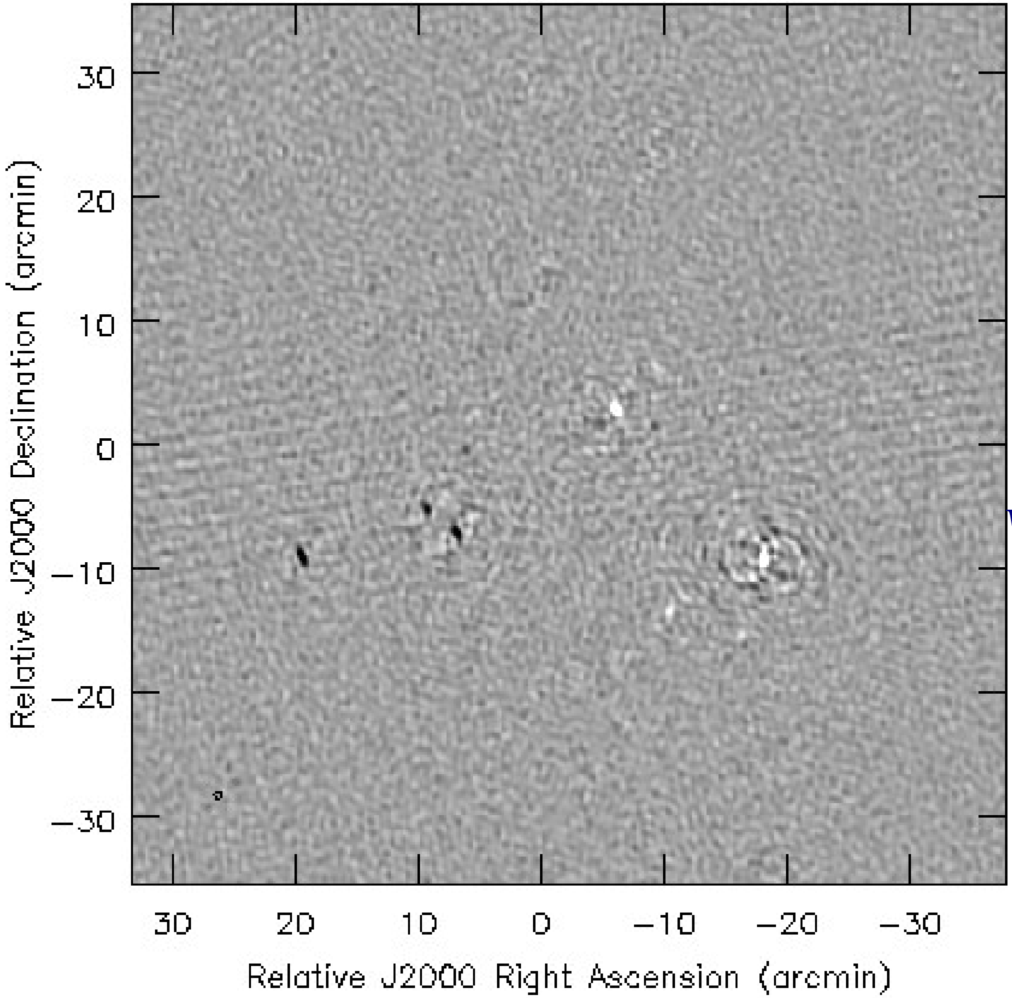
EVLA L-Band Stokes-I: After correction



- 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

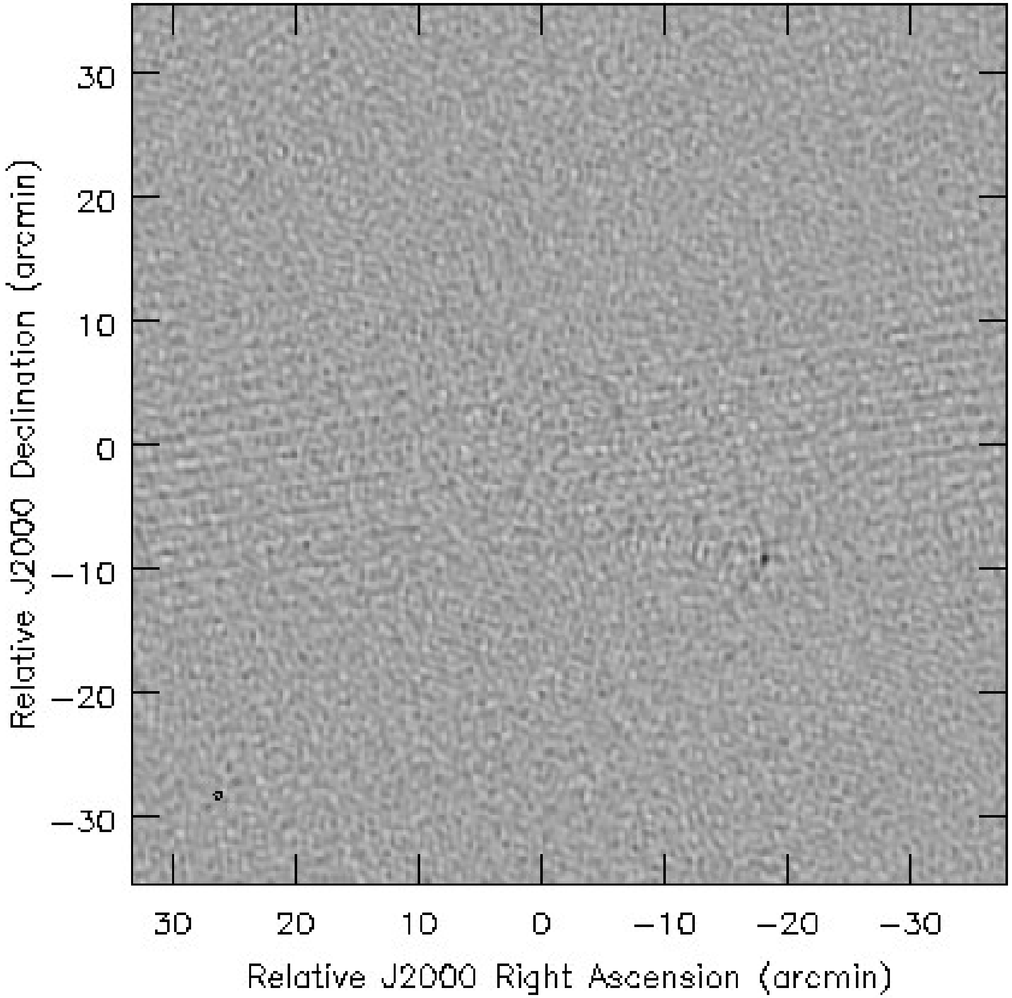


Is it $M(s, Poln)$?

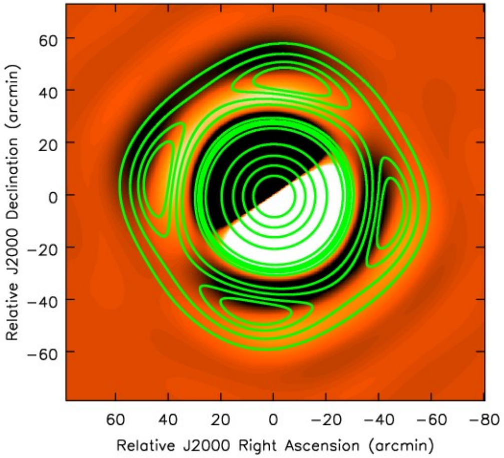
Or is it $I(s, Poln)$?

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

EVLA L-Band Stokes-I: After correction

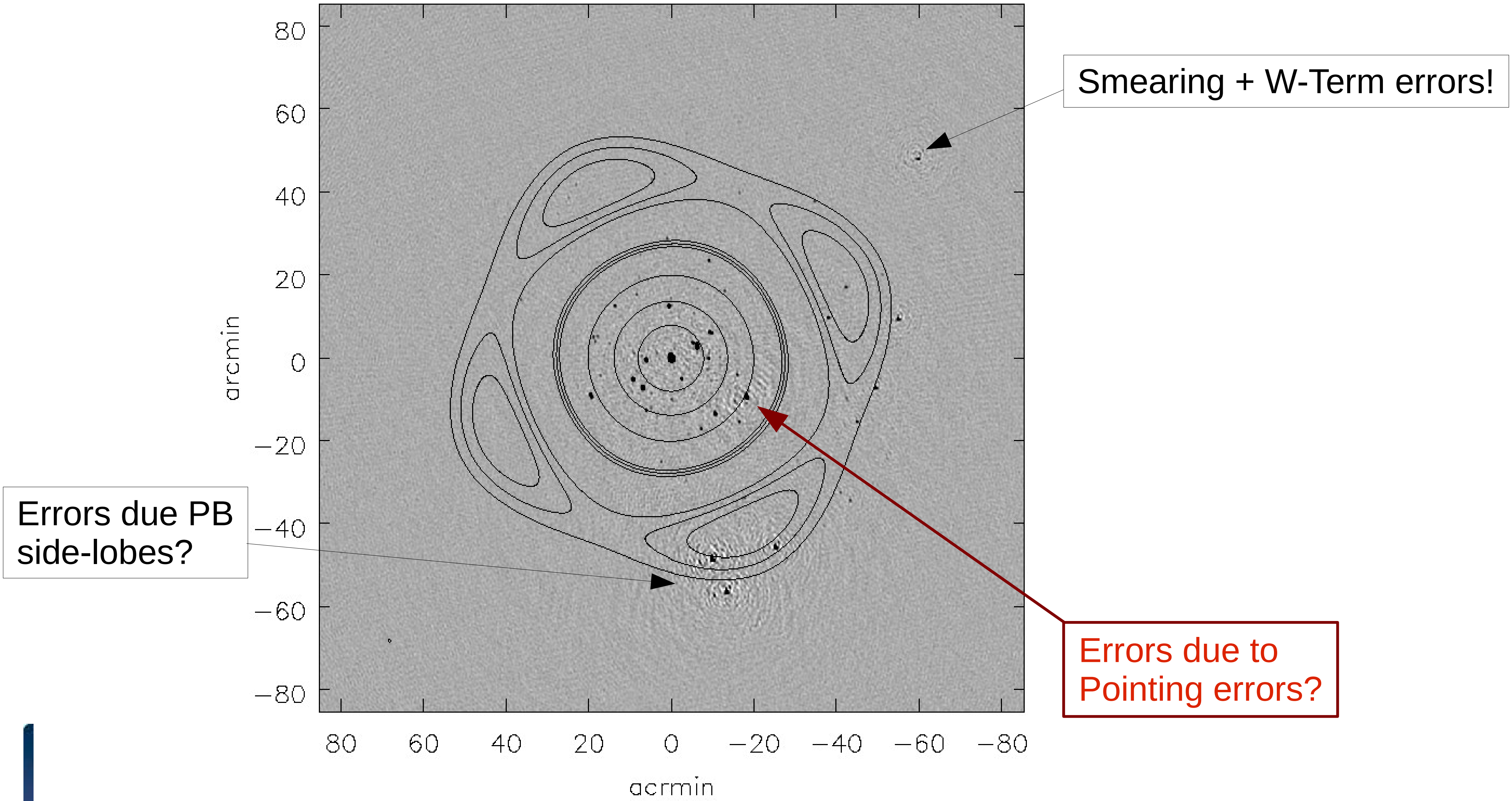


Use physical model for the Stokes-V pattern:

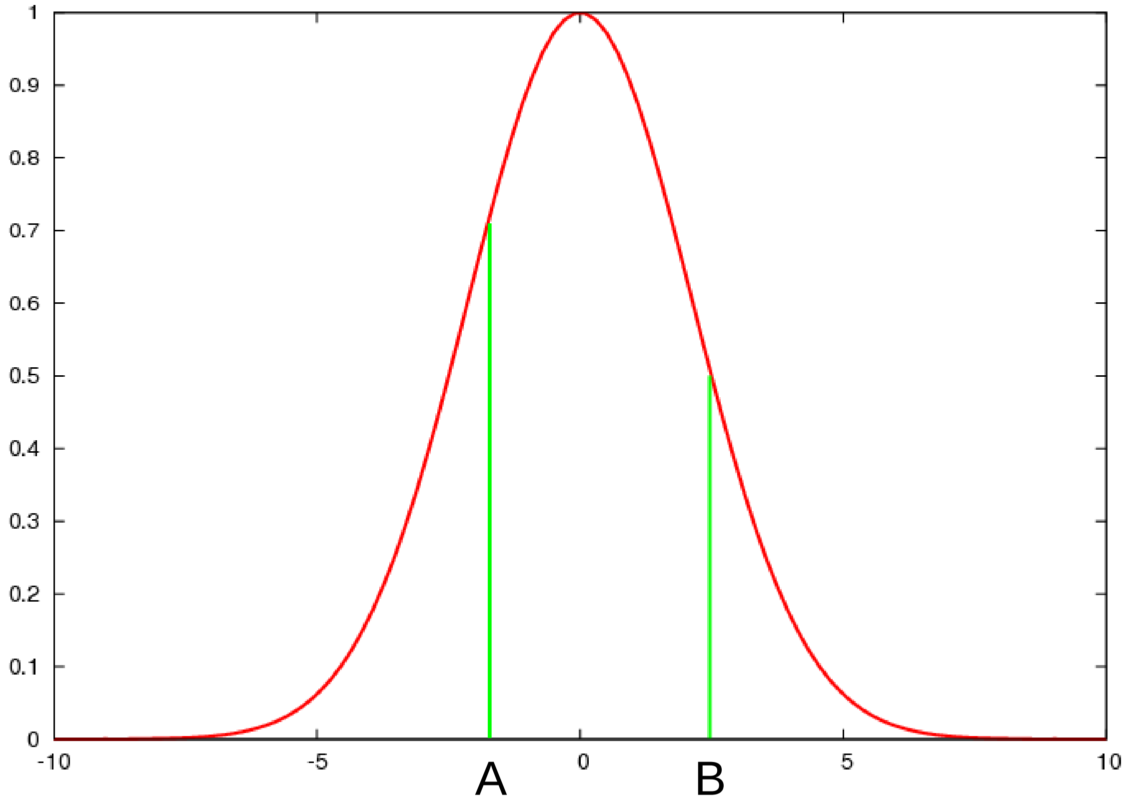


Contours: Stokes-I power pattern
Colour: Stokes-V power pattern

3C147: Residual errors in full field

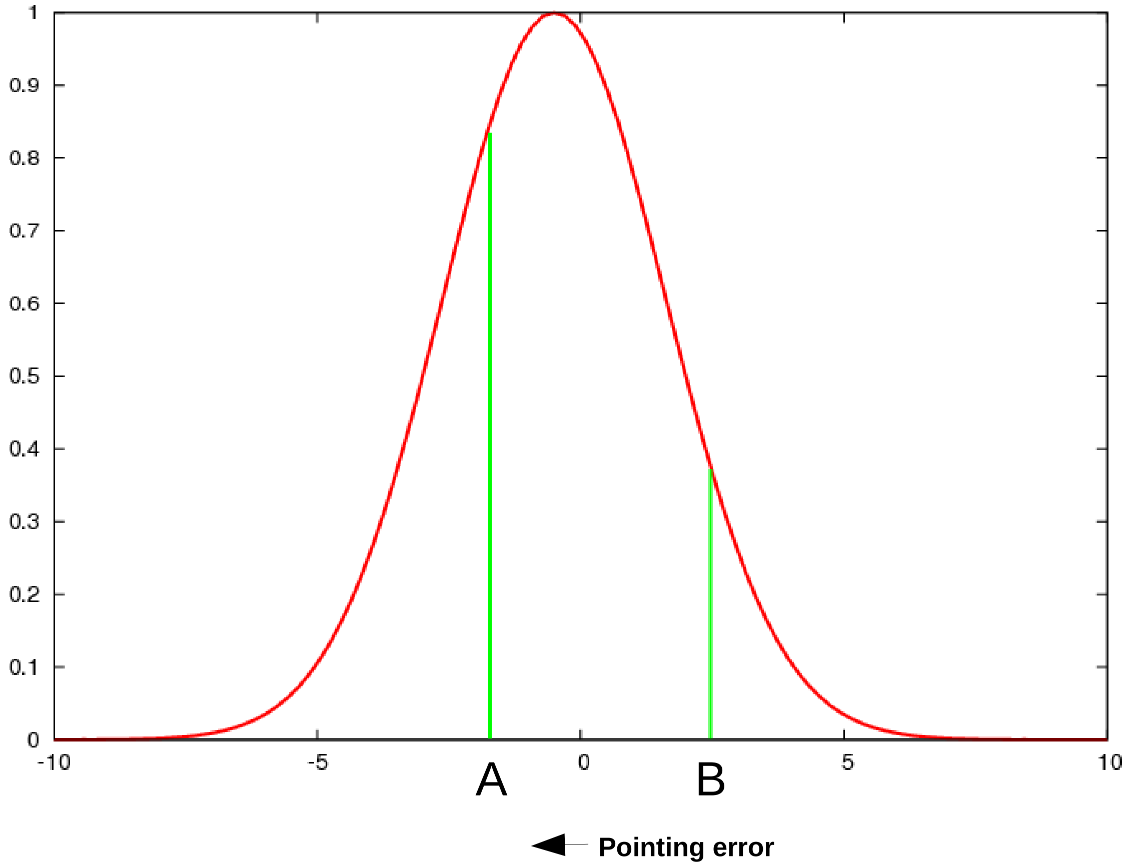


Effect of antenna pointing errors



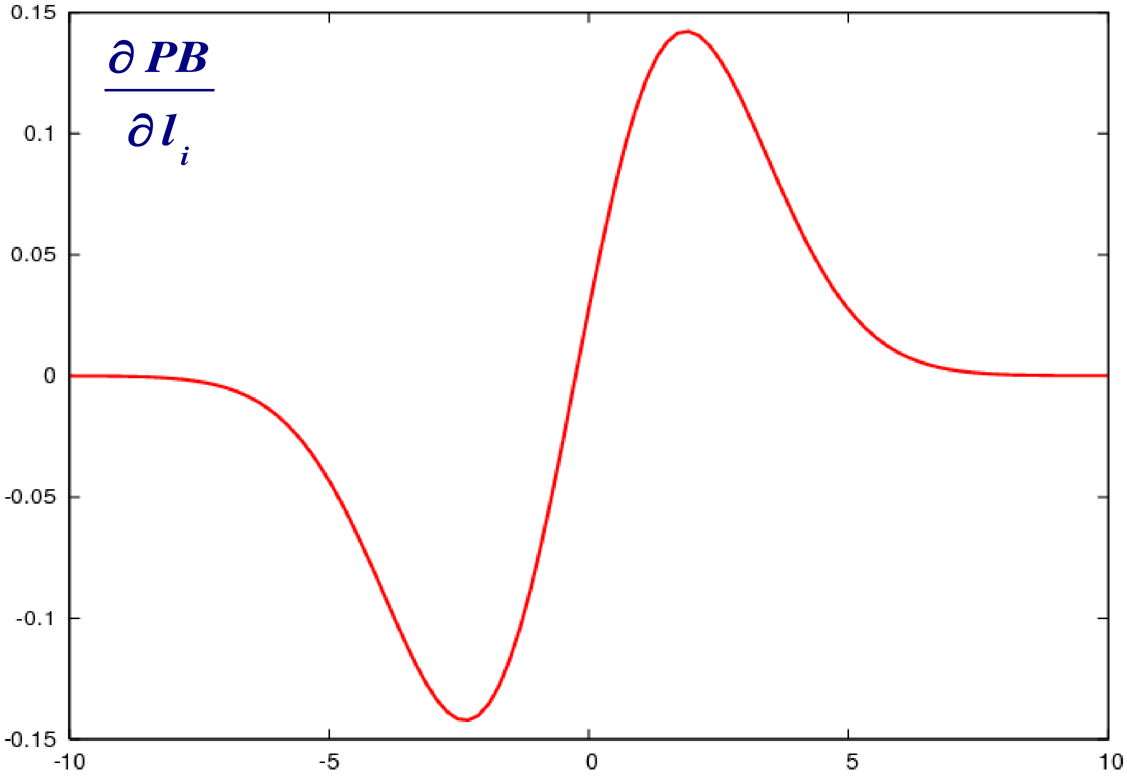
- Effect of antenna pointing error is a direction dependent effect
- A purely Hermitian effect in the data domain, in the absence of DI gains
  - To the first order, amplitude-only error in image domain
- However, there is significant in-beam phase structure –particularly for wide-field, full-Stokes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

Effect of antenna 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
  - Solve for the shape of the function which best-fits the gain variations at A and B

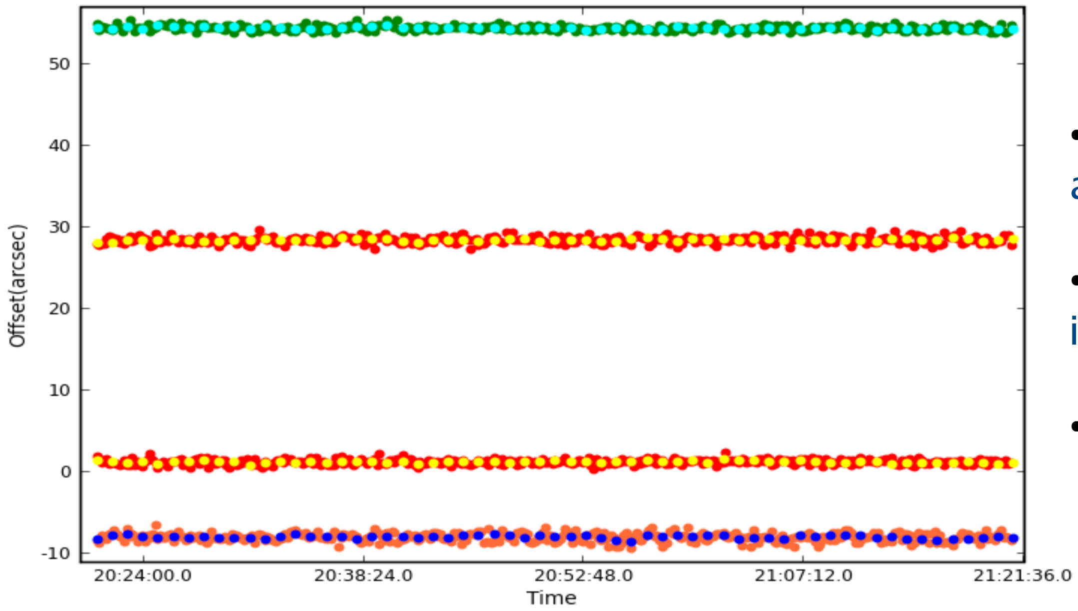
Error pattern: Derivatives



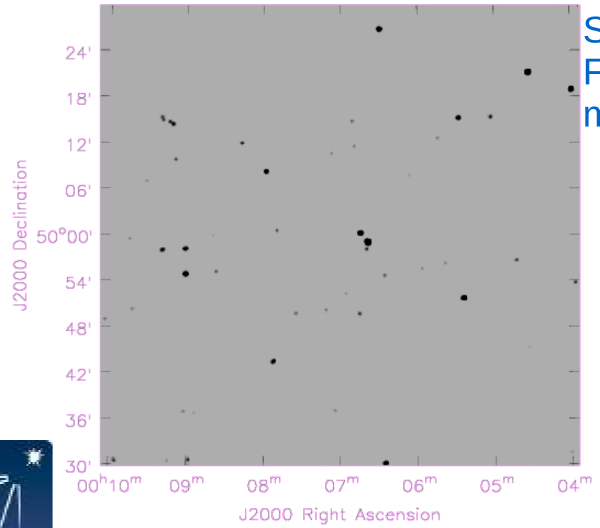
$$Image\ plane\ error \sim PSF * \left(I(s) \frac{\partial PB(s)}{\partial l_i} \right)$$

- Peak error signal around the half-power point
  - Points of largest derivative
- Significant for fields with emission distributed throughout the Primary Beam
  - Single pointing or mosaicking at low frequencies
  - Mosaicking at high frequencies
- Realistic Primary Beam patterns are wide-spread in image plane
  - Band-limited in the data domain
  - Fortuitously, it is also a data domain effect!
- Hint: Think of sinc(x) vs Pill-box functions

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



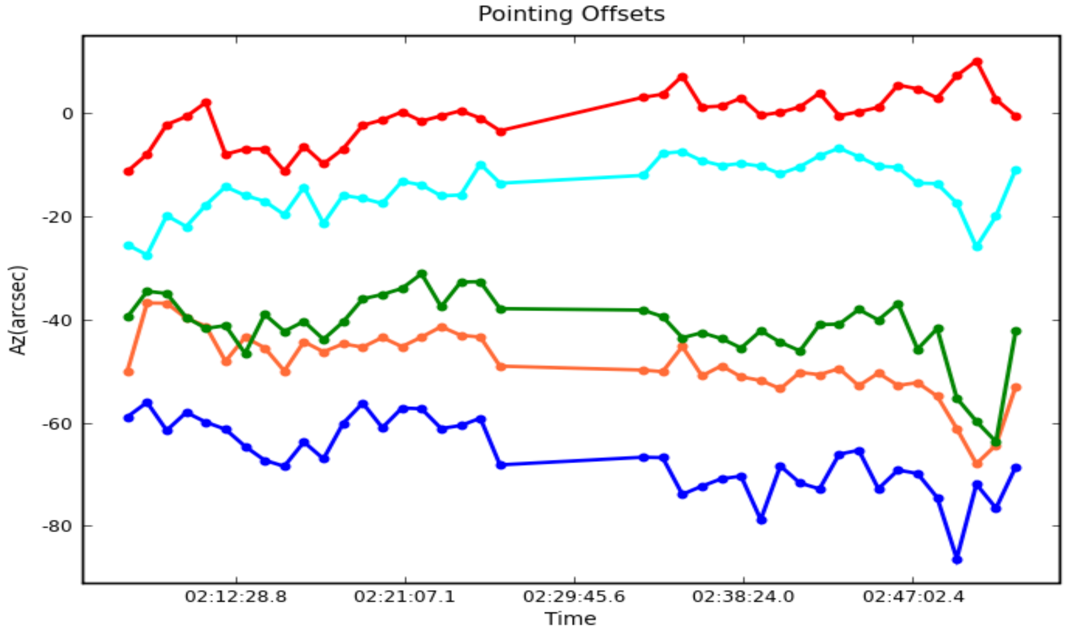
Sources from NVSS. Flux range ~2-200 mJy/beam

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

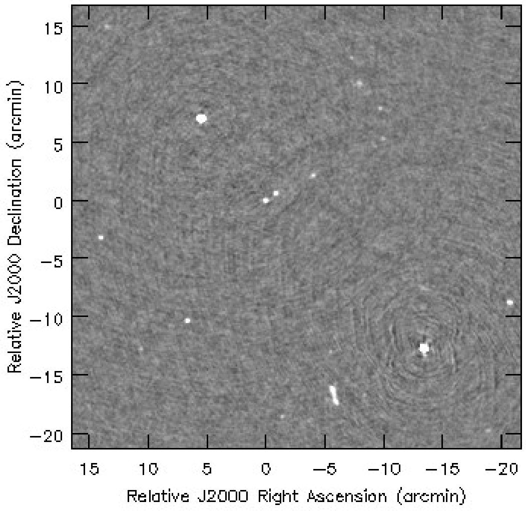
$$\left[\frac{\partial E_{ij}(p_i^k, p_j^k)}{\partial E_i} \frac{\partial E_i}{\partial p_i^k} \right] * V_{ij}^{M} = 0$$

[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.
- Data from TALG0002



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

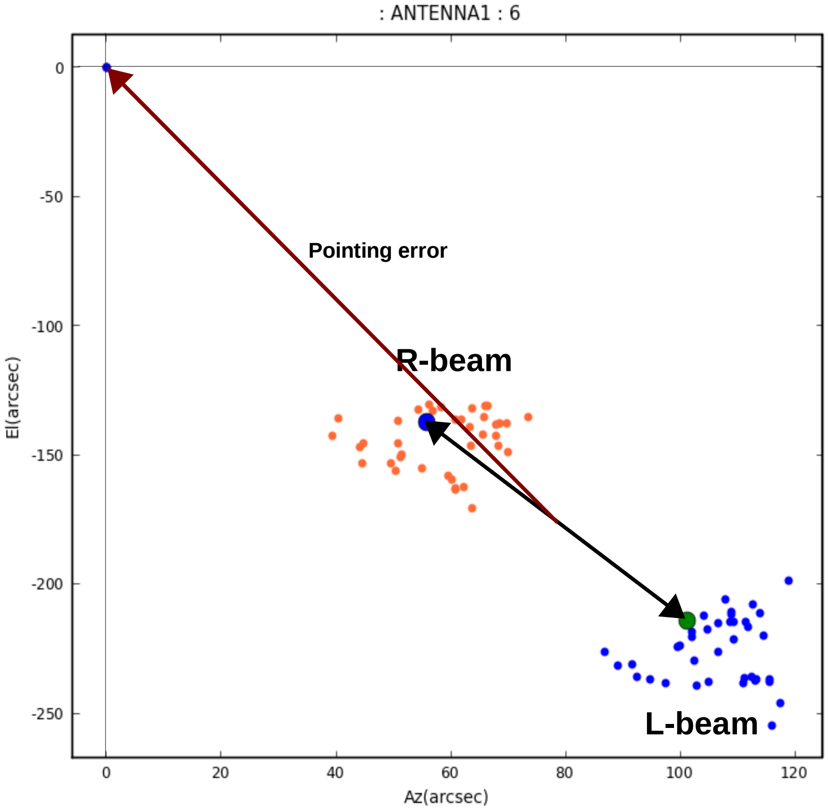
Noise Budget:

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

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

[paper in preparation]

DD SelfCal algorithm: EVLA Data



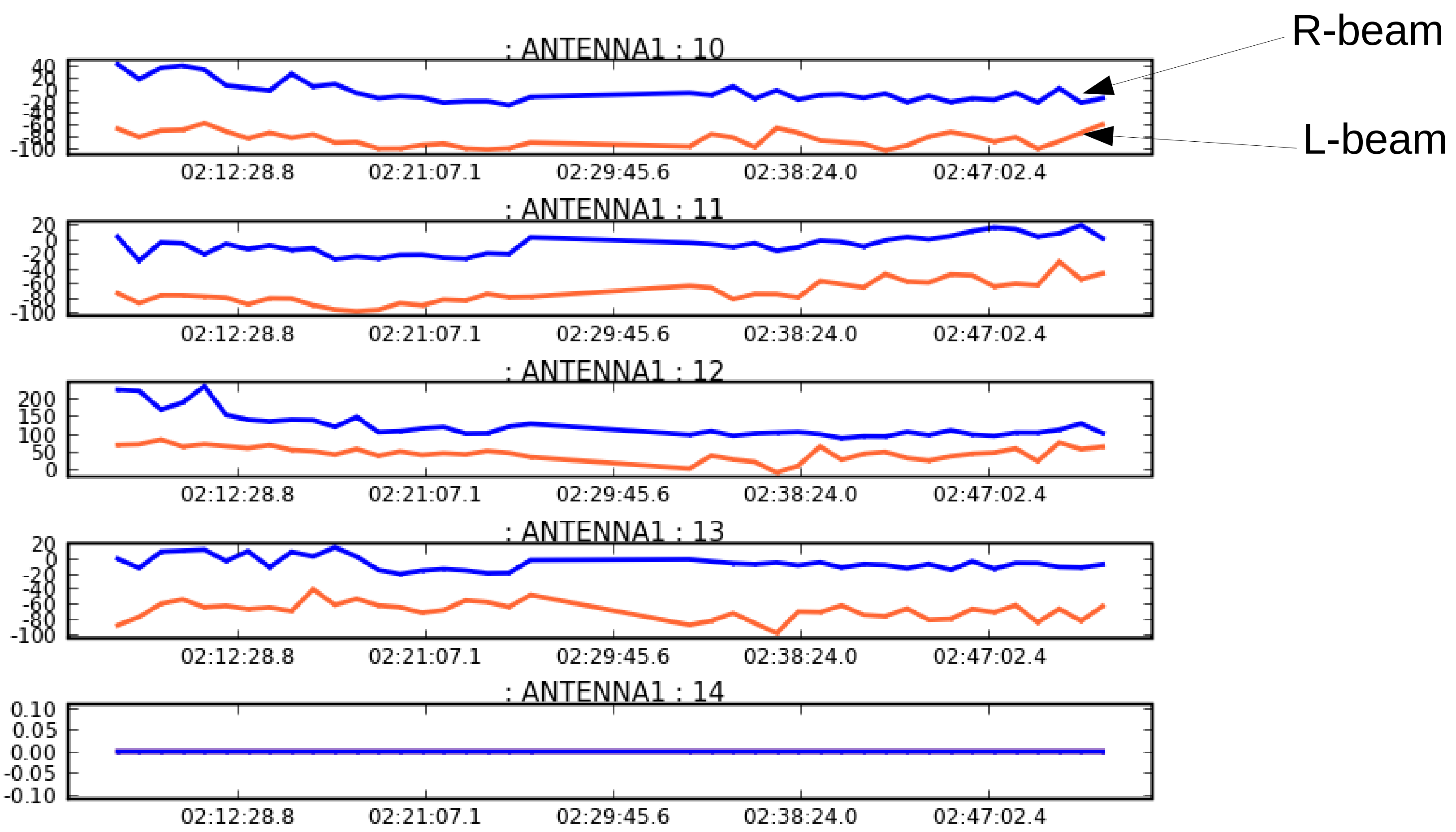
- El-Az mount antennas
- Polarization squint due to off-axis feeds
  - The R- and L-beam patterns have a pointing error of +/- ~0.06 $\frac{\lambda}{D}$

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

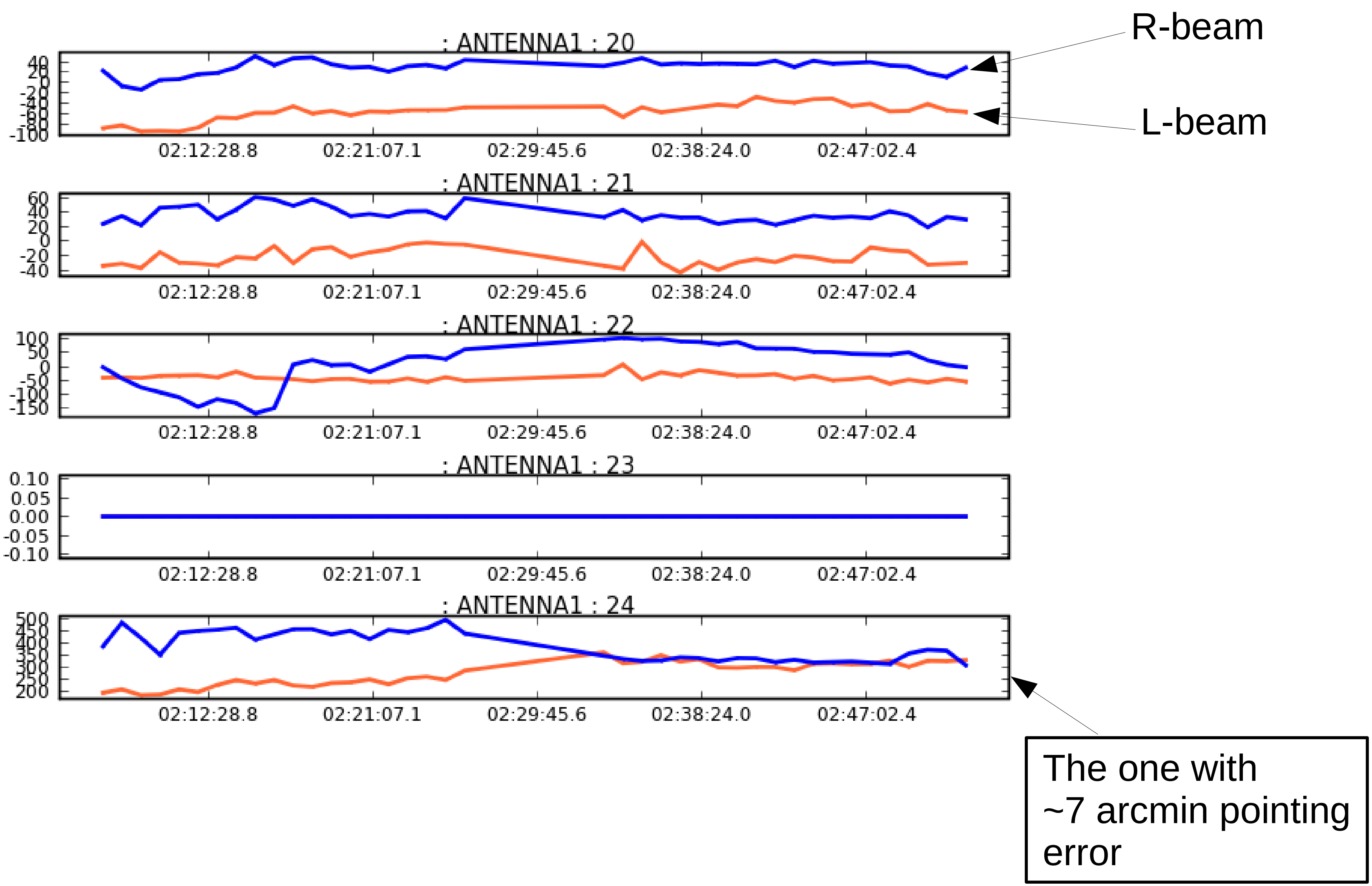
- EVLA polarization squint solved as pointing error (optical pointing error).
- Squint would be symmetric about the origin in El-Az plane in the absence of antenna servo pointing errors.
- Pointing errors for various antennas detected in the range 1-7 arcmin.
- Pointing errors confirmed independently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

[paper in preparation]

Time dependent solutions



Time dependent solutions



DD SelfCal: General comments

- Pointing SelfCal formulation is generalization of DI SelfCal

Standard SelfCal (DI): $$V_{ij} = (G_i \otimes G_j^*) V_{ij}^M$$

Pointing SelfCal: $$V_{ij} = (J_i^S \otimes J_j^{S^*}) * V_{ij}^M$$

- Effects of PB/antenna pointing is purely Hermitian in the data domain – in the absence of DI gains or in-beam phase etc.
  - I.e., amp-only effect in the image plane
- Fundamentally an antenna based effect
  - Difficult to decouple/interpret in the image plane
- Fundamentally a data-domain effect
  - Not an “image plane effect”
  - Unlike, e.g., effects of sky spectral index variations (a DD error)

- Clean works, but scale-sensitive methods work better
- Similarly, Partitioning/SelfCal works, but DD SelfCal should work better!

Next steps

- Fold these solutions into A-Projection
  - Test with simple fields (e.g. IC2233)
  - Test for Stokes-V : A strong instrumental effect
- Is polynomial expansion a better parametrization

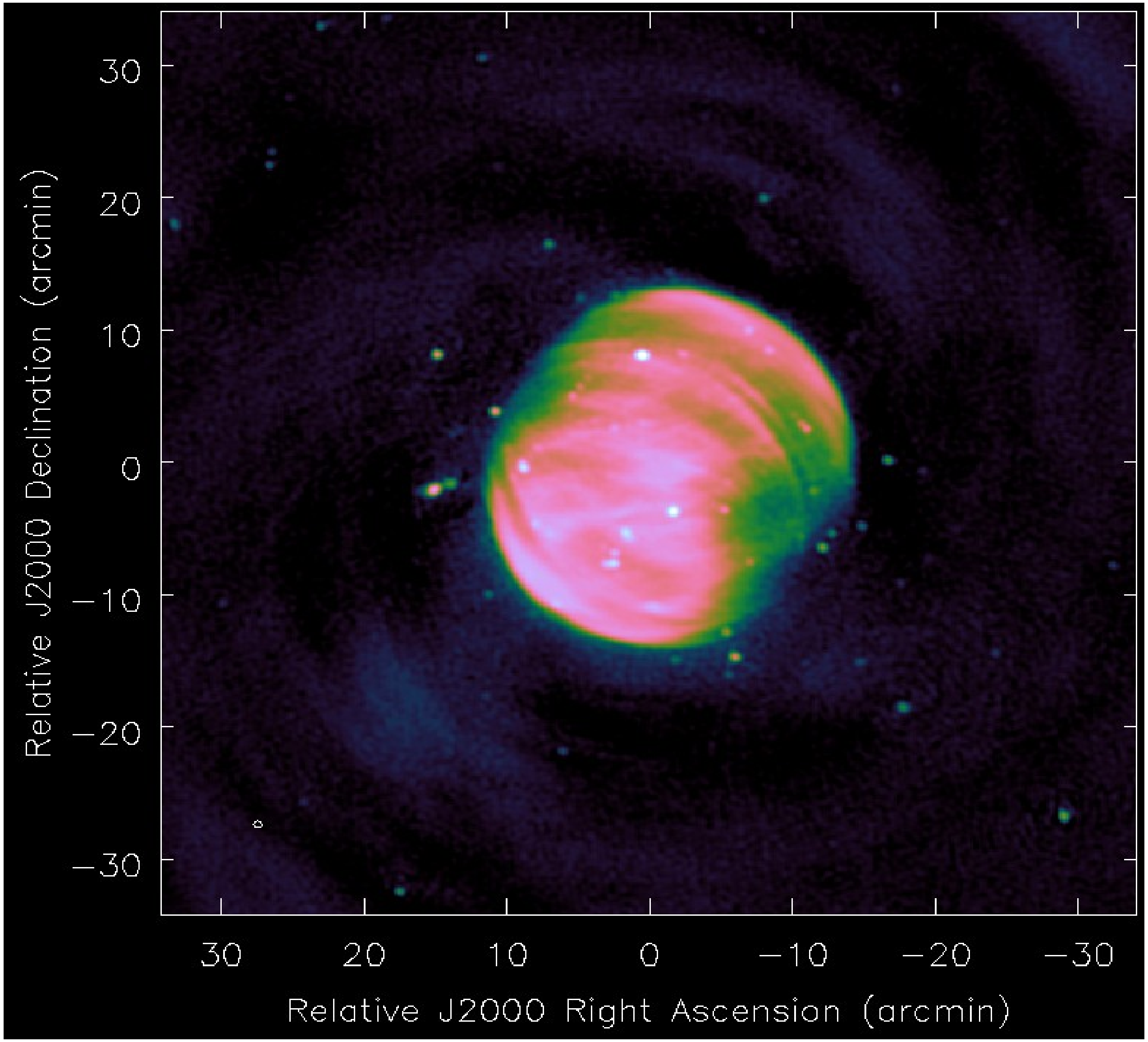
$$A(u,v)=A_o(u,v)\,e^{\iota(a_o u+a_1 v)}$$

vs.

$$A(u,v)=A_o(u,v)[a_o Z_o(u,v)+a_1 Z_1(u,v)]$$

- A good target algorithm for GPU computing
- Include it in mosaicking
  - Generalize A-Projection for mosaicking
  - Naturally includes corrections for PB effects
- Solve for shape!
  - Strongest “term” here is the elevation dependence.
  - Slow and smooth changes with time (a good thing!)

Wide band imaging with the EVLA

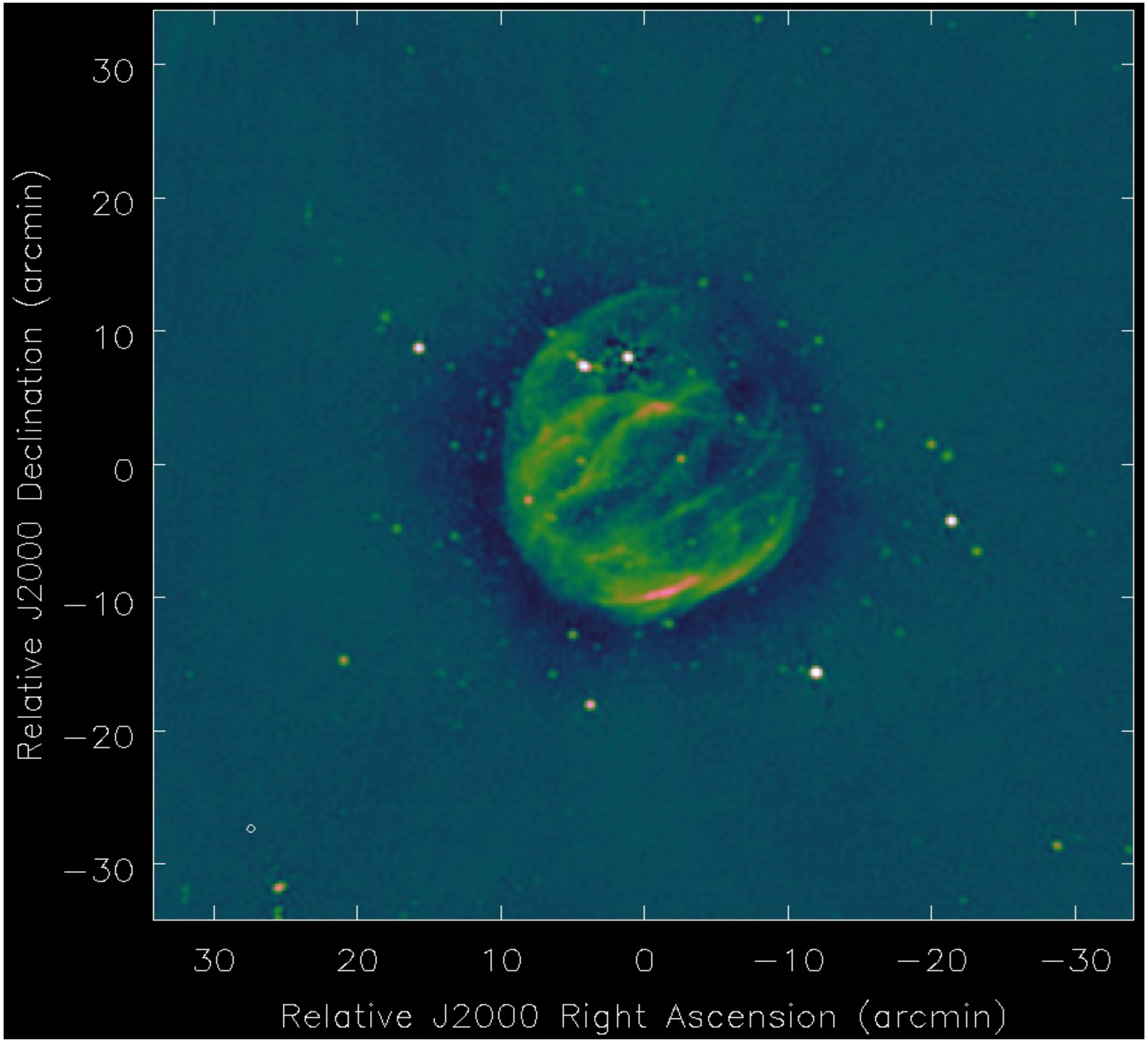


- Emissions fills the PB
- Extended emission with superimposed compact sources

- Partitioning approach designs-out such fields!

- 2X128 MHz

Wide band imaging with the EVLA

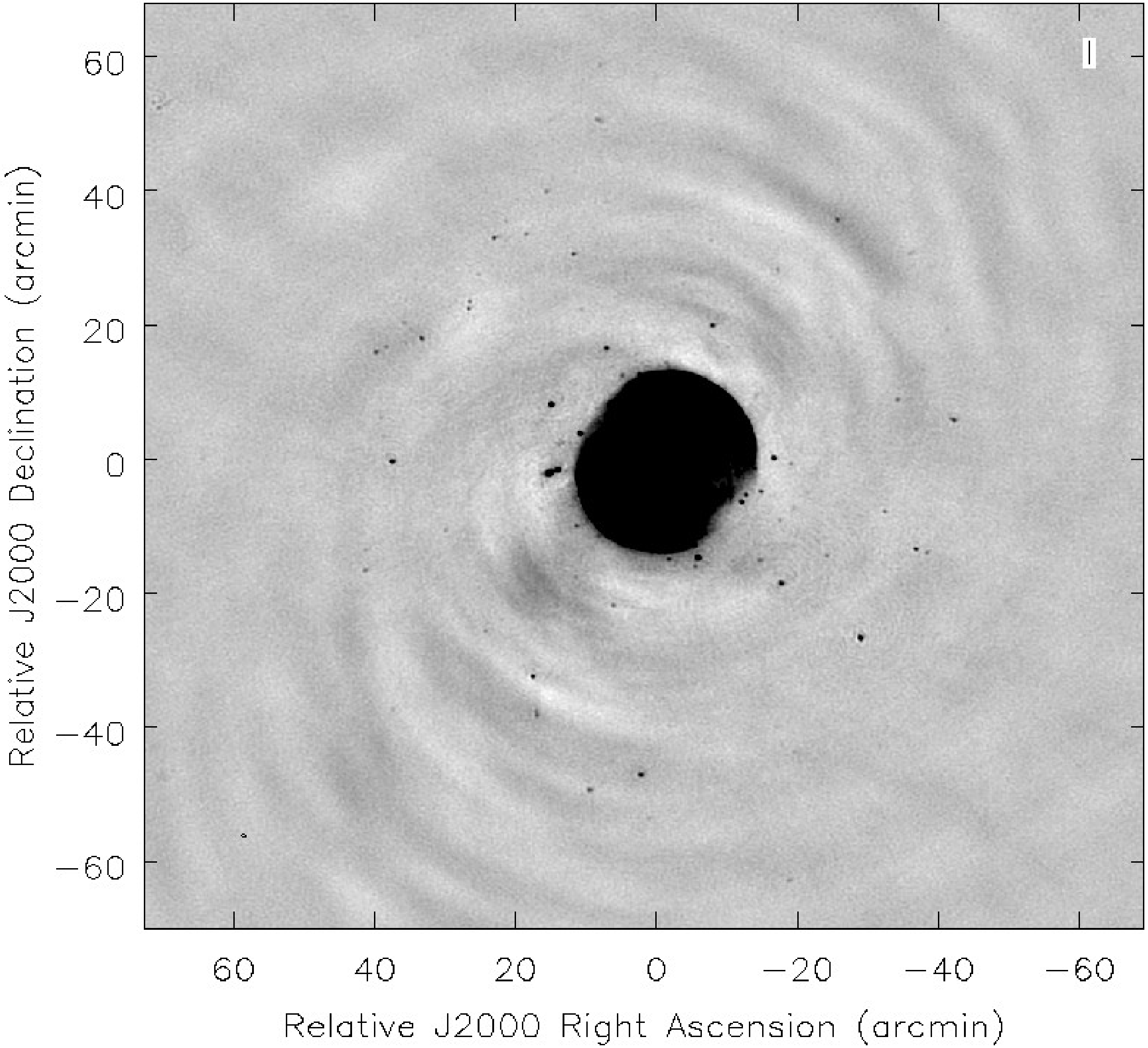


- Emissions fills the PB
- Extended emission with superimposed compact sources

- Partitioning approach design-out such fields!

- 4X128 MHz

Wide band imaging with the EVLA



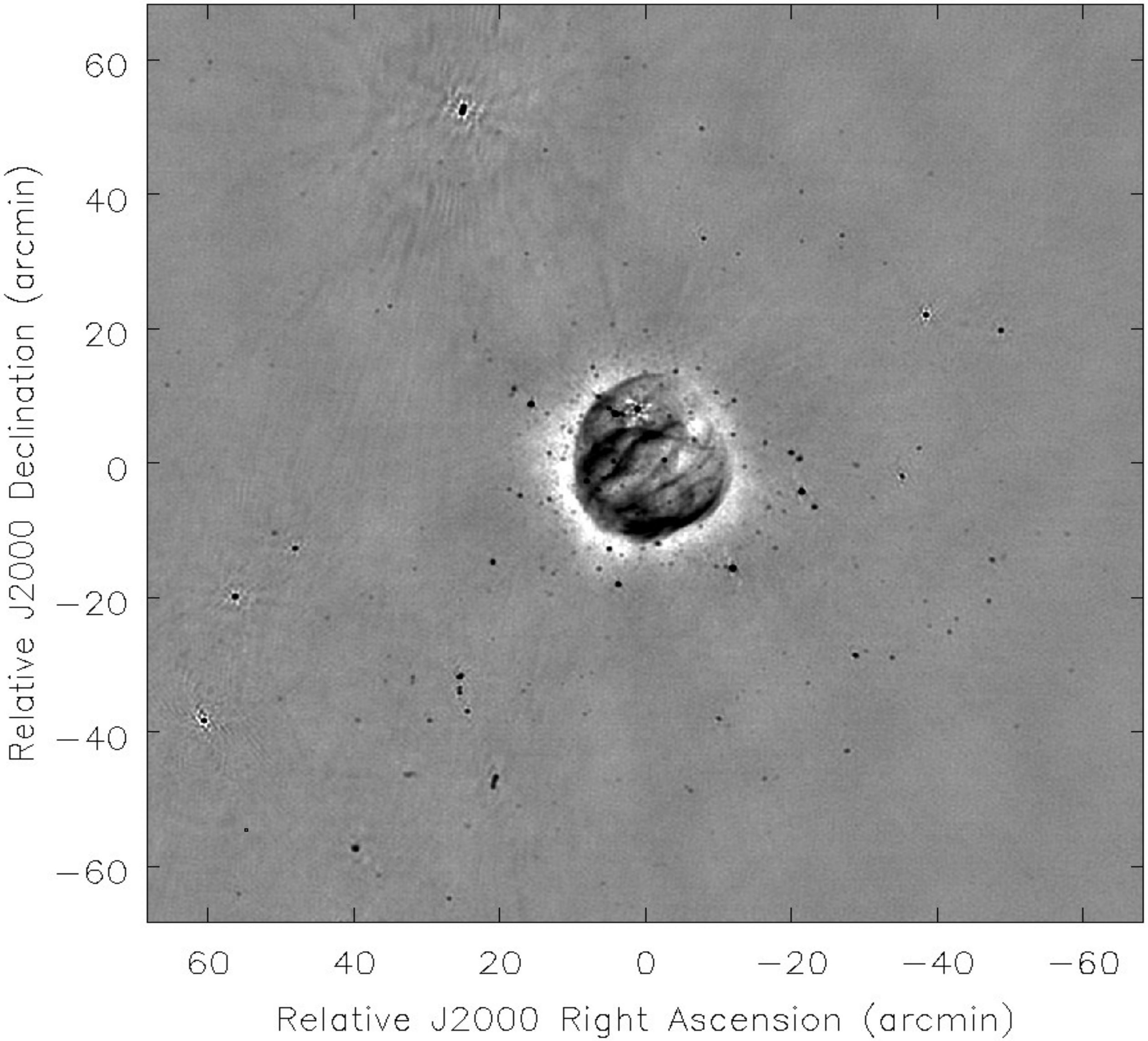
- 1.2-1.8GHz
  (2x128 MHz)
- ~40 microJy/Beam

- RSRO Projects
  (AB1345, Bhatnagar et al.)

- Scientific goals
  - Spectral Index imaging
  - RM Synthesis

  - Wide-band, wide-field imaging
  - HPC

Wide band imaging with the EVLA



- 1.2-1.8GHz
  (4x128 MHz)
- ~25 microJy/Beam

- RSRO Projects
  (AB1345, Bhatnagar et al.)

- Scientific goals
  - Spectral Index imaging
  - RM Synthesis

  - Wide-band, wide-field imaging
  - HPC

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

NRAO

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

NRAO

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]

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!
    - 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