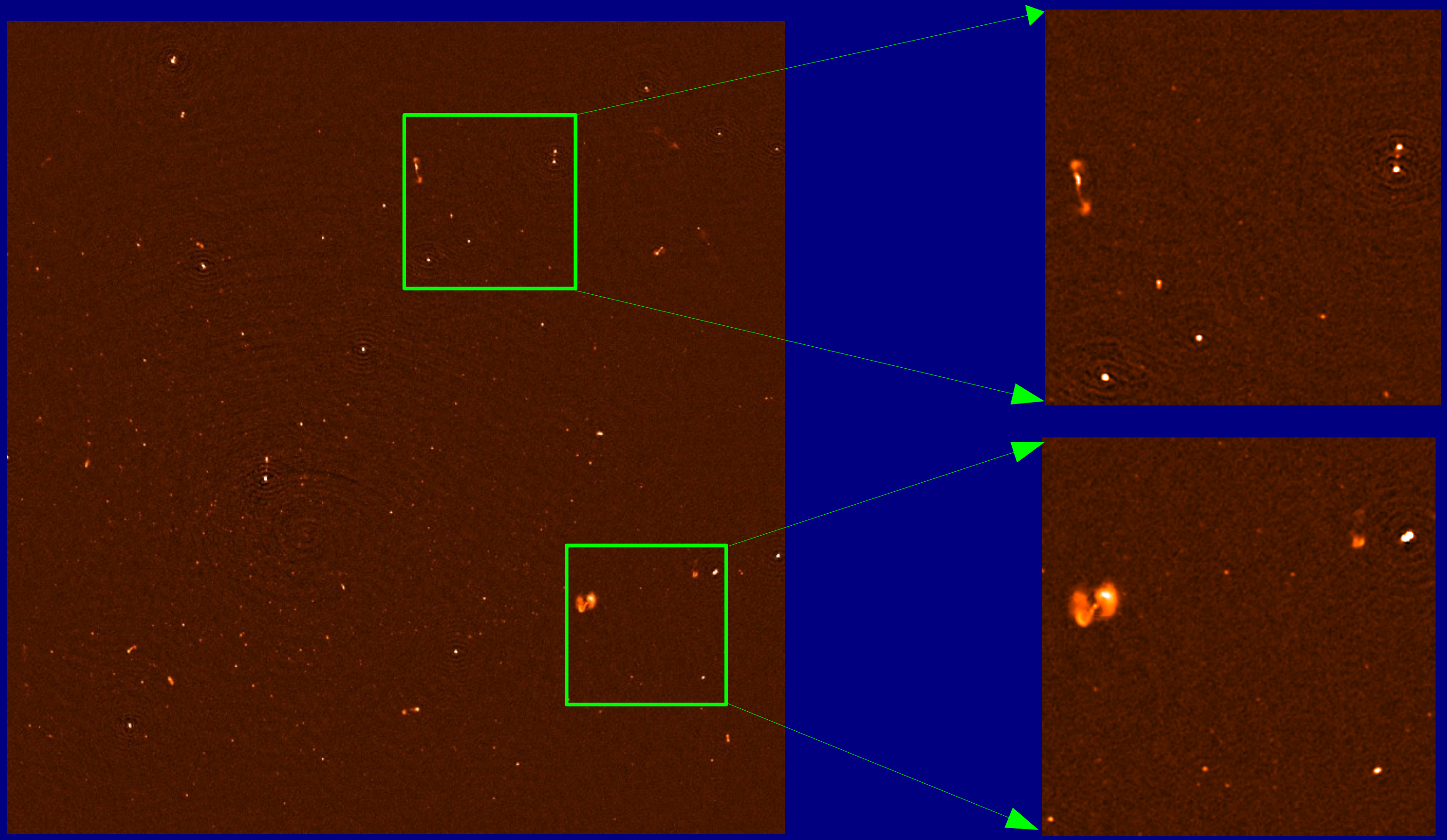
Imaging algorithms and computing



S. Bhatnagar

NRAO

Challenges

- 2:1 Bandwidth ratio
  - Primary beam effects
    - Time and frequency dependent
    - Polarization response
  - Spectral index variations across the sky
  - Deconvolution errors, Pixelation errors
- Direction dependent (DD) effects
  - Pointing errors
  - Long, non co-planar baselines (w-term)
  - Ionospheric phase screen
- Computing and I/O loads

Challenges

- Strong RFI
  - Some algorithms/schemes exist
- Weak RFI
  - Very difficult to detect and remove
  - Will/does affect high dynamic range imaging
- Near field problems
  - Remains correlated
  - Not the same at all baselines
  - Variable in time & frequency
- Self Interference

The Measurement Equation

- Generic Measurement Equation: [HBS papers]

$$V_{ij}^{Obs}(\nu) = M_{ij}(\nu, t) W_{ij} \int M_{ij}^{S}(\mathbf{s}, \nu, t) \; I(\mathbf{s}, \nu) \; e^{\iota \mathbf{s}.\mathbf{b}_{ij}} \; d\,\mathbf{s}$$

Data　Corruptions　Sky　W-term

- Corruptions: $M_{ij} = J_i \otimes J_j^*$:direction independent corruptions

  $M_{ij}^{s} = J_i^{s} \otimes J_j^{s*}$:direction dependent corruptions

- Sky: Frequency dependent sky: $I(s, \nu) = I(s, \nu_o)\left(\frac{\nu}{\nu_o}\right)^{\alpha}$

- W-term: $e^{\iota s.b_{ij}} = e^{\iota[ul + vm + w(\sqrt{1 - l^2 - m^2} - 1)]}$: Not a FT kernel

  (a.k.a. non co-planar array)

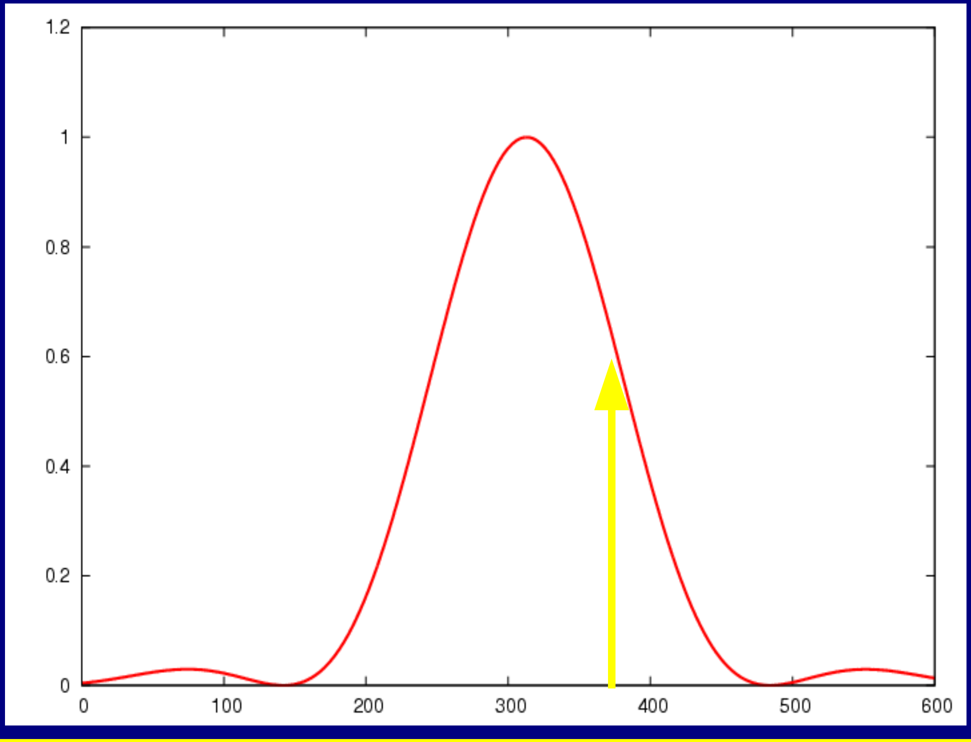
Pieces of the puzzle

- Unknowns:
  - M_{ij},M^s_{ij}: *Electronics, Primary Beams, Antenna pointing, ionosphere,...*
  - I^M : *Extended emission, spectral index variations, polarization,...*
- Need Efficient Algorithms:
  - *Correct for image plane effects*
  - *Decompose the sky in a more appropriate basis*
    - Frequency sensitive (combine with MFS)
  - Solvers for the “unknown” direction dependent effects (pointing, PB shape, ionospheric effects,...)
    - As expensive as imaging!
- Needs (Computing):
  - Parallel computing & I/O
  - Scalable algorithms & software

Pieces of the puzzle: DI gains

- Unknowns: I^M
  - *Ideal stuff: No gain errors, Known Primary beam*
  - *Use image deconvolution to get True Sky Image!*

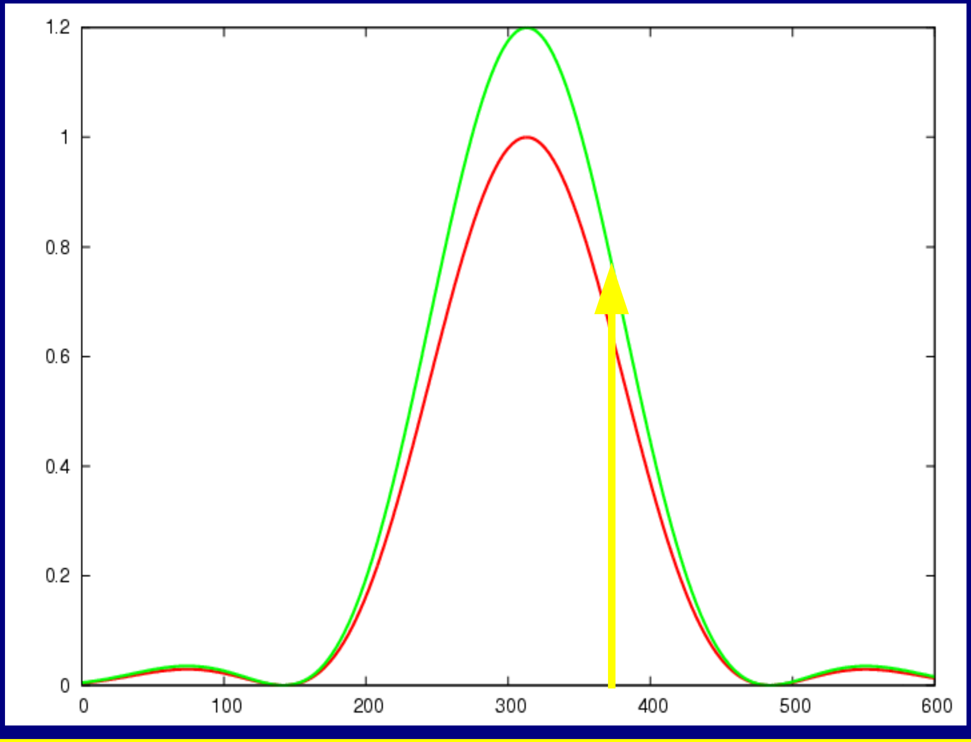
$$V_{ij}^{o} = V_{ij}^{M}$$



Pieces of the puzzle: DI gains

- Unknowns:
  - M_{ij}: *Constant across the Field Of View*
  - I^M : *Extended emission, spectral index variations, polarization,...*

$$V_{ij}^{M} = M_{ij} \; V_{ij}^{o}$$
$$= J_i \otimes J_j \; V_{ij}^{o}$$

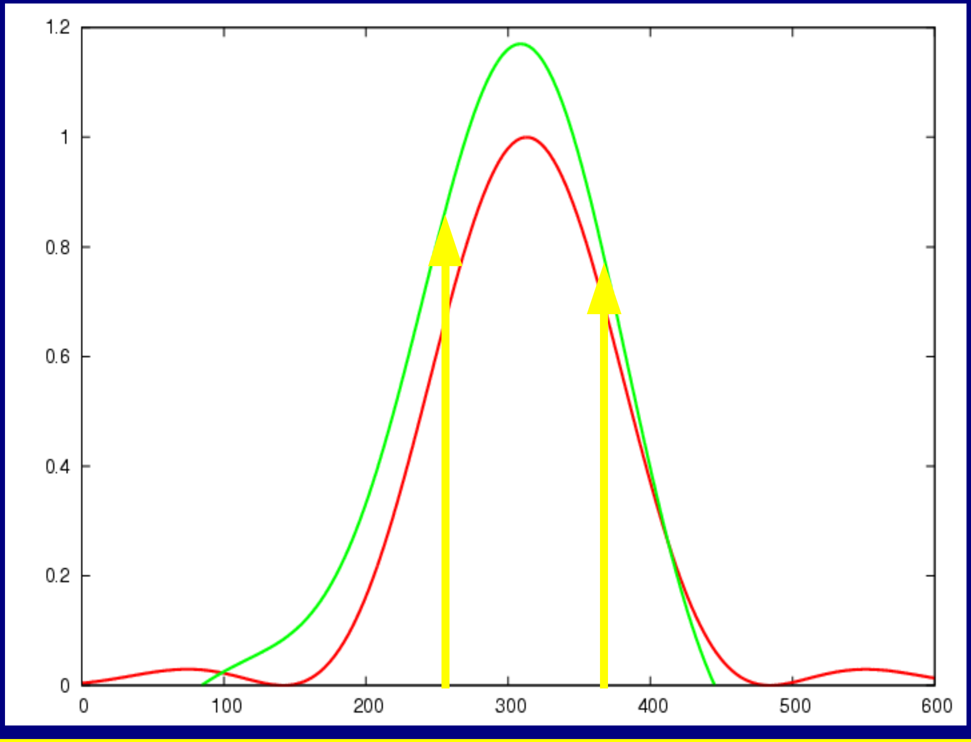


"calib", "bpass" (AIPS)
"gaincal", "bandpass" (CASA)

Pieces of the puzzle: DD gains

- Unknowns:
  - M_{ij}, M^s_{ij}: *Constant Part+ Part Variable across the Field Of View*
  - I^M : *Extended emission, spectral index variations, polarization,...*

$$V^M_{ij} = M_{ij} \; FT\left[M^s_{ij}(s) I(s)\right]$$
$$= M_{ij}\left[M^s_{ij}(s) * V^M_{ij}\right]$$
$$= J_i \otimes J_j \, FT\left[\left(J^s_i \otimes J^s_j\right) \; I(s)\right]$$



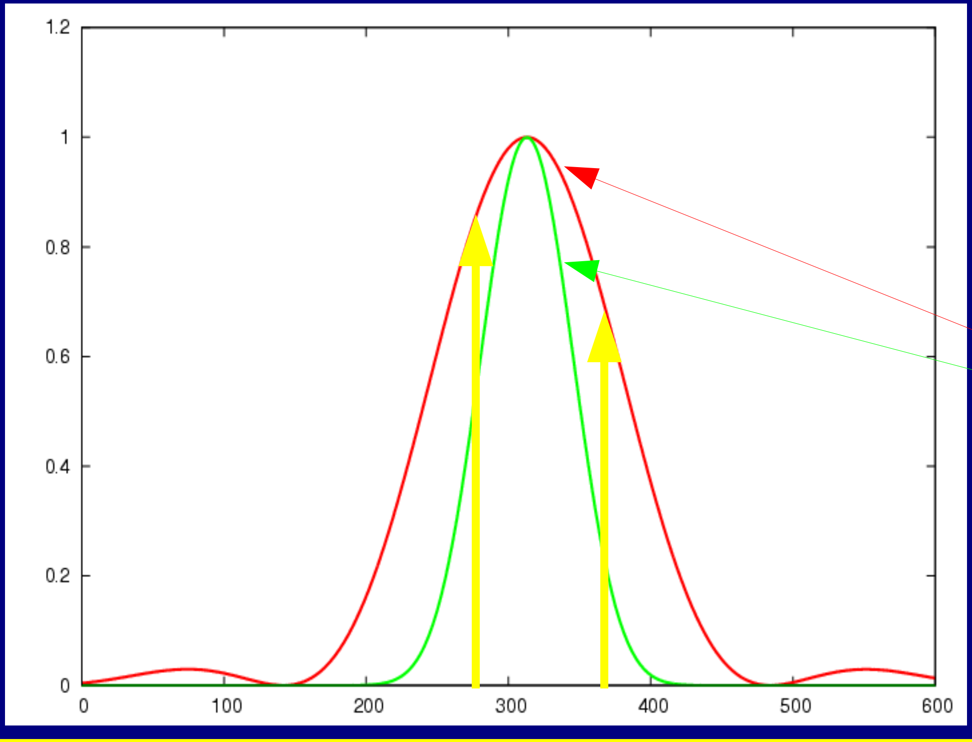
Pieces of the puzzle: DD+Freq. D gains

- Unknowns:
  - M_{ij}, M^s_{ij}: *Electronics, Primary Beams, Antenna pointing, ionosphere,...*
  - I^M : *Extended emission, spectral index variations, polarization,...*

$$V^M_{ij} = M_{ij} \; FT\left[M^s_{ij}(s,\nu) I(s,\nu) \right]$$

$$= M_{ij}\left[M^s_{ij}(s,\nu) * V^M_{ij} \right]$$

$$= J_i \otimes J_j \, FT\left[\left(J^s_i \otimes J^s_j \right) \; I(s) \right]$$



Two ends of the observing band

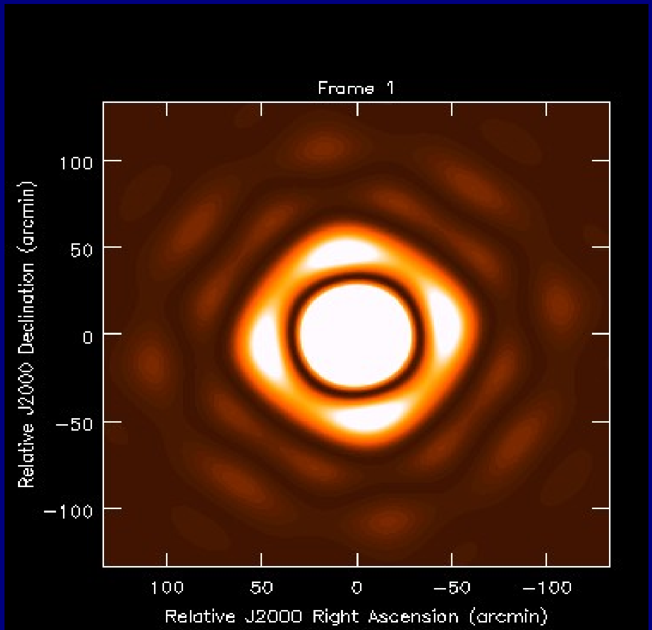
Primary Beam Effects

- EVLA full beam, full band, single feed

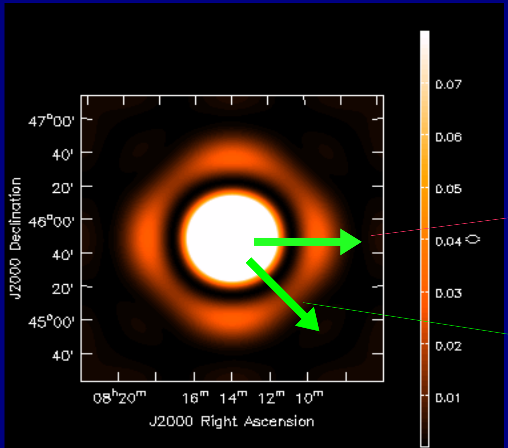
PB variation across the band

EVLA: Sources move from main-lobe to side-lobes

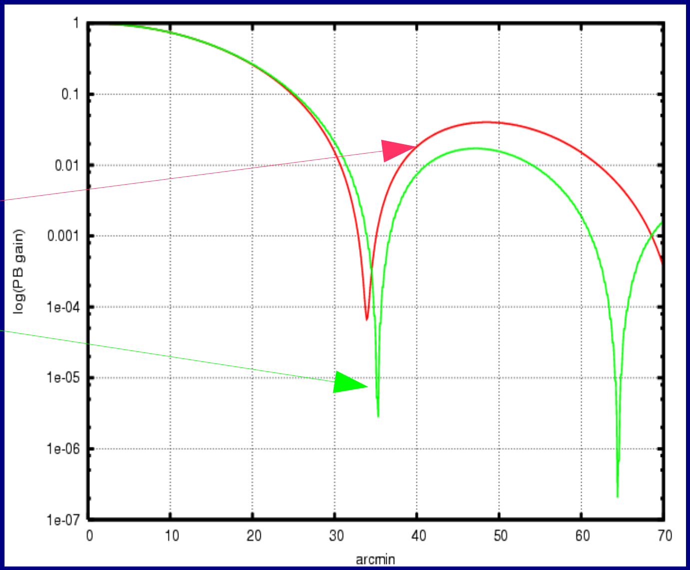
PB rotation, pointing errors

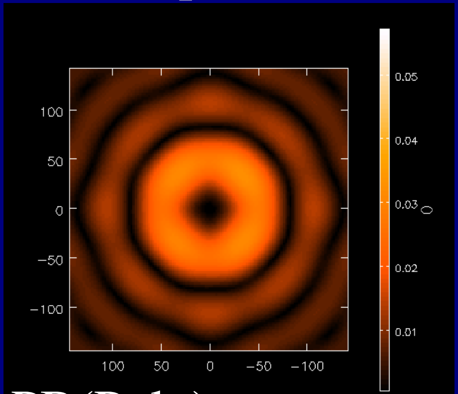


PB(t)



PB(Freq)





PB(Poln)

Cross hand power pattern

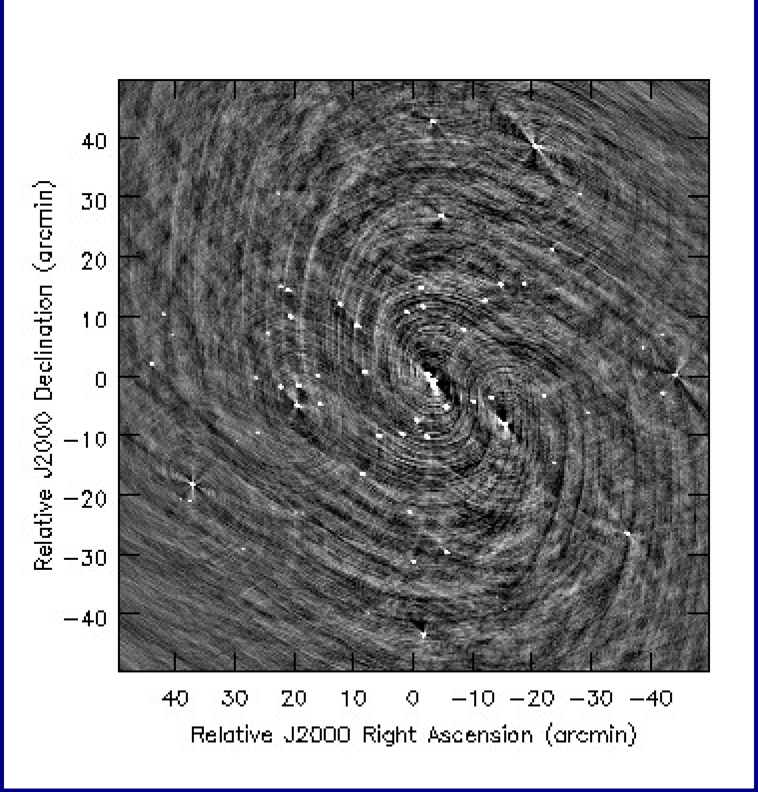
PB gain varies as a function time, frequency and direction in the sky

Dominant sources of error: Single Pointing

Requirements: "...full beam, full Stokes, wide-band imaging at full sensitivity".

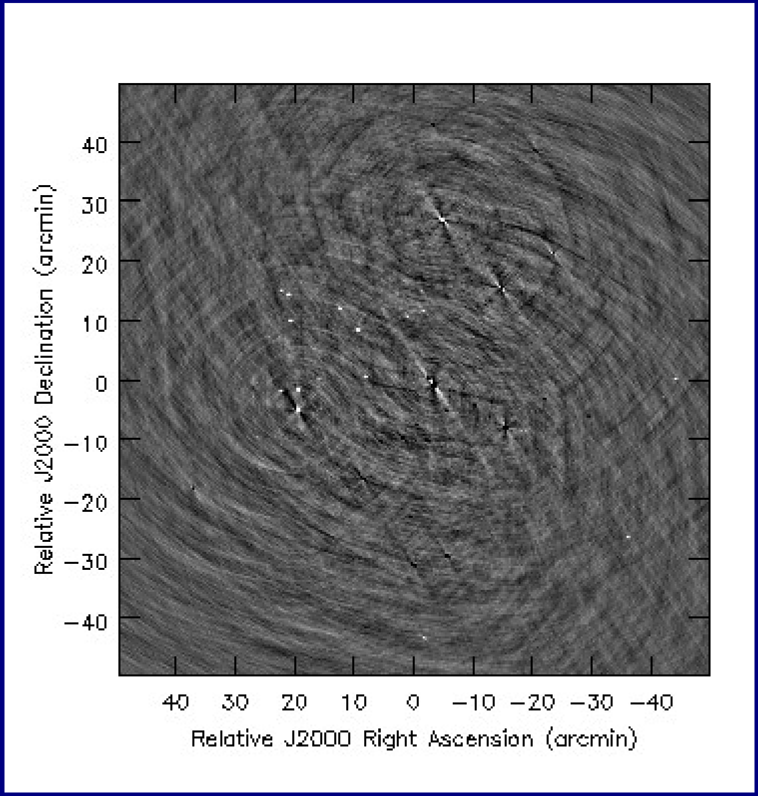
- EVLA full beam
  - Estimated Stokes-I imaging Dynamic Range limit: $\sim 10^4$

Stokes-I



RMS ~15μJy/beam

Stokes-V

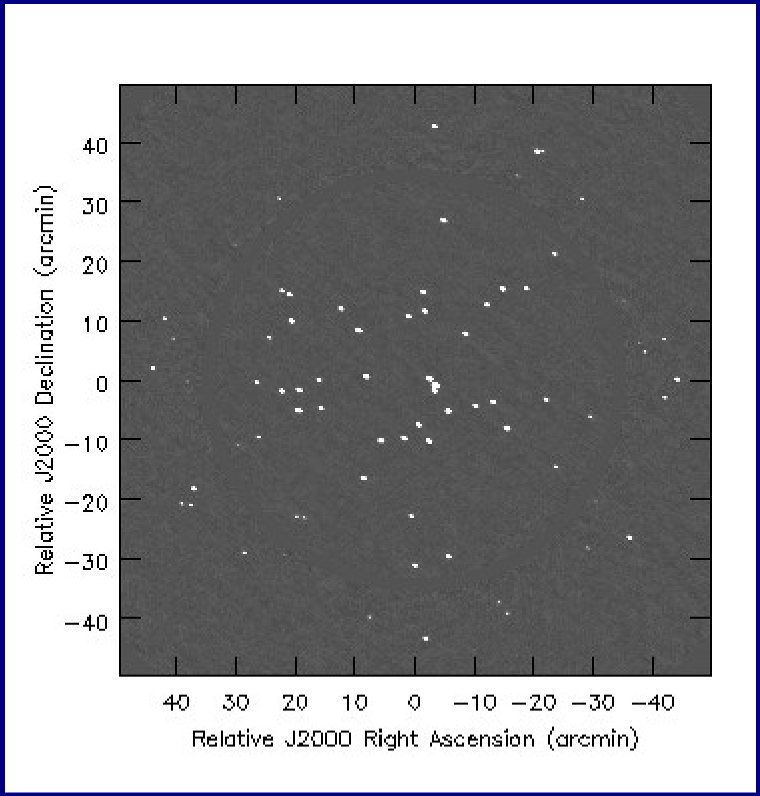


Dominant sources of error: Single Pointing

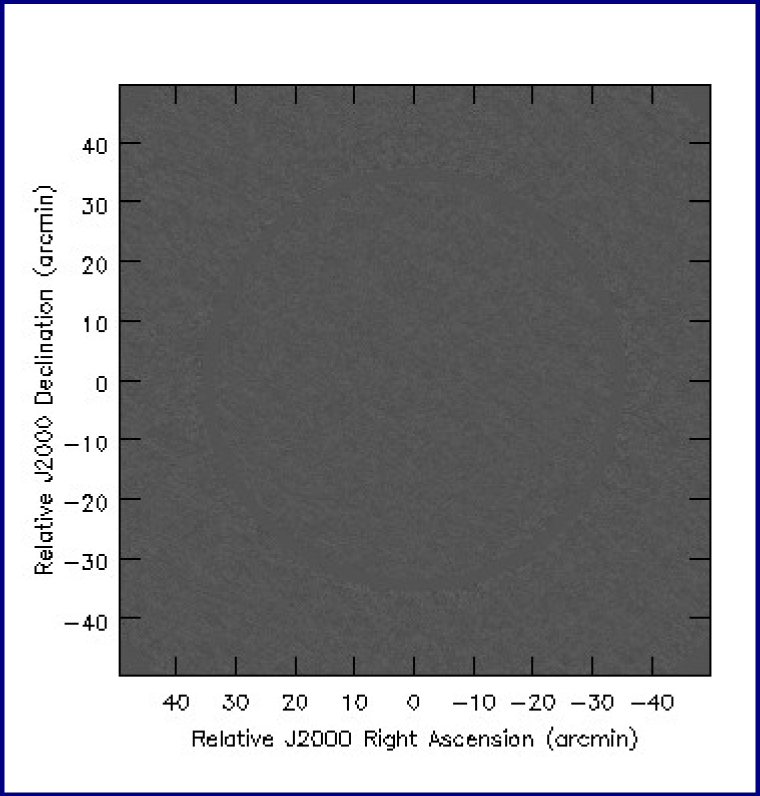
Requirements: "...full beam, full Stokes, wide-band imaging at full sensitivity".

- EVLA full beam
  - Estimated Stokes-I imaging Dynamic Range limit: ~10^4

Stokes-I

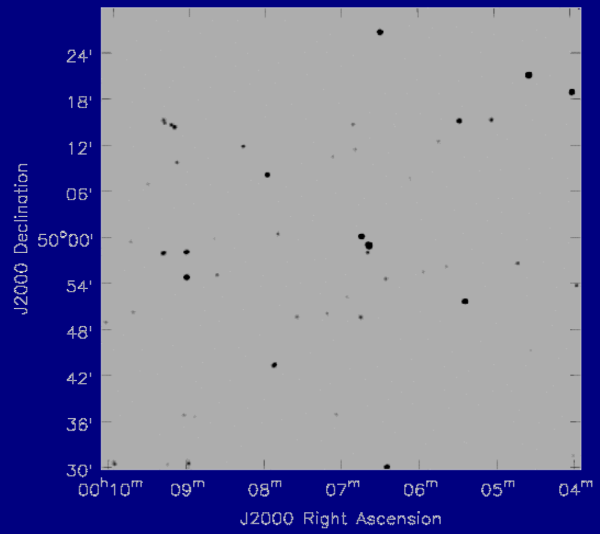


Stokes-V

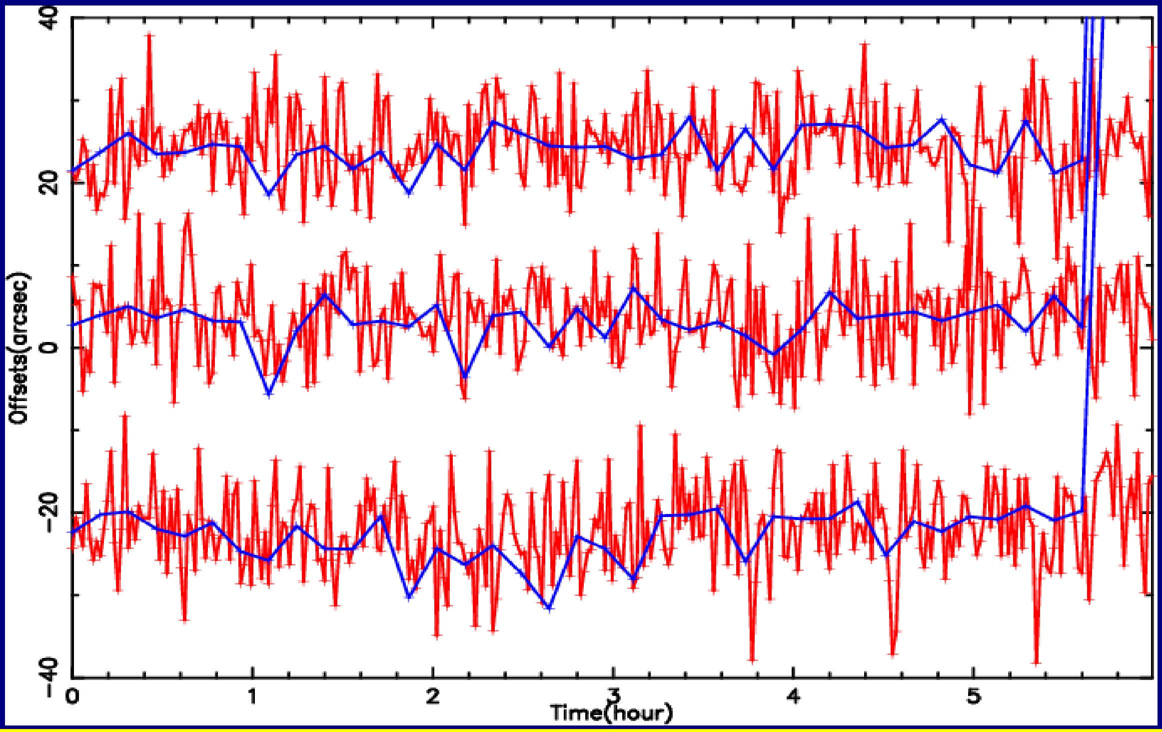


RMS ~1μJy/beam

Pointing SelfCal: Example



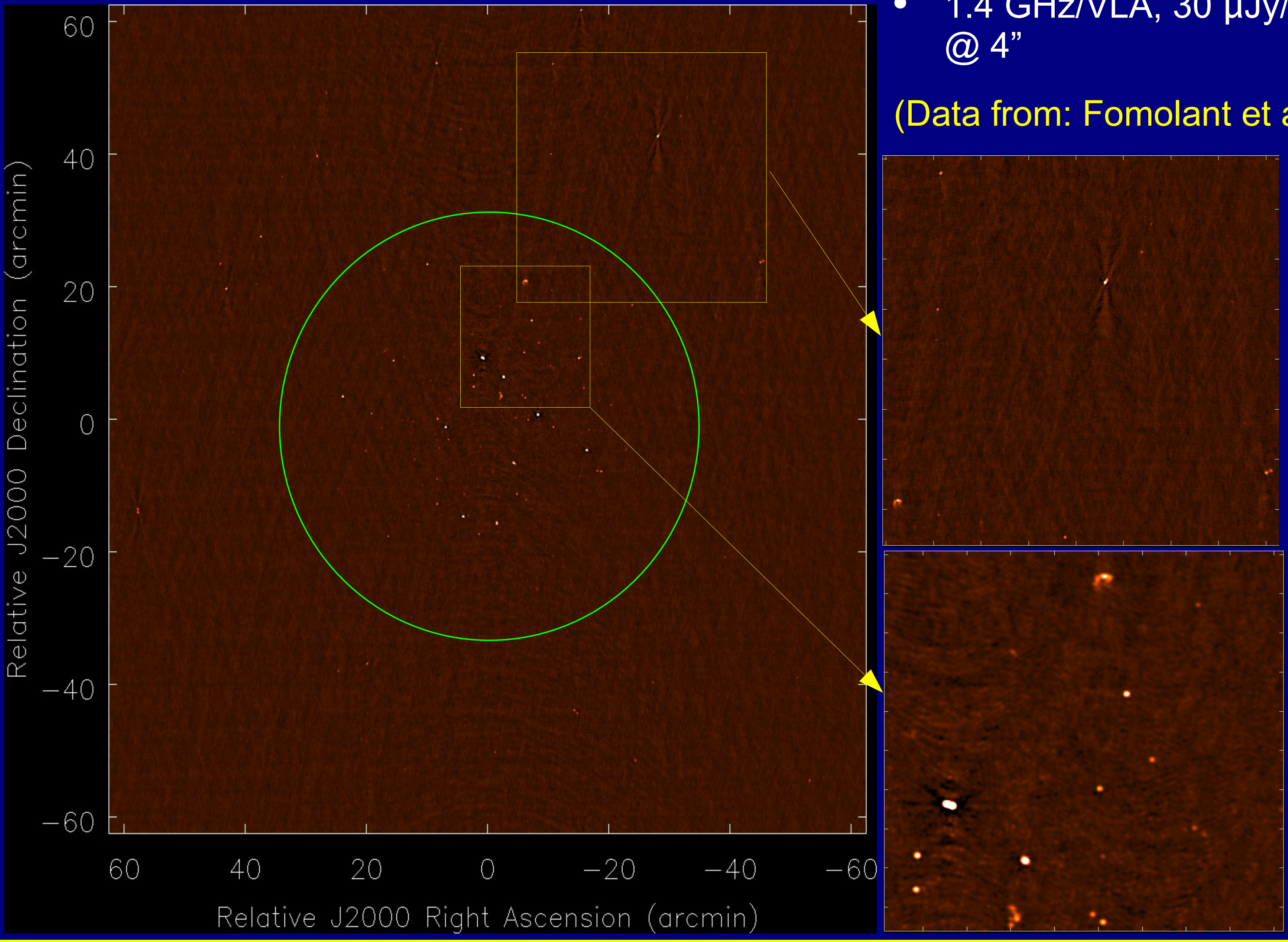
Model image: 59 sources from NVSS. Flux range ~2-200 mJy/beam



Red: Typical antenna pointing offsets for VLA as a function of time

Blue: Solved antenna pointing errors

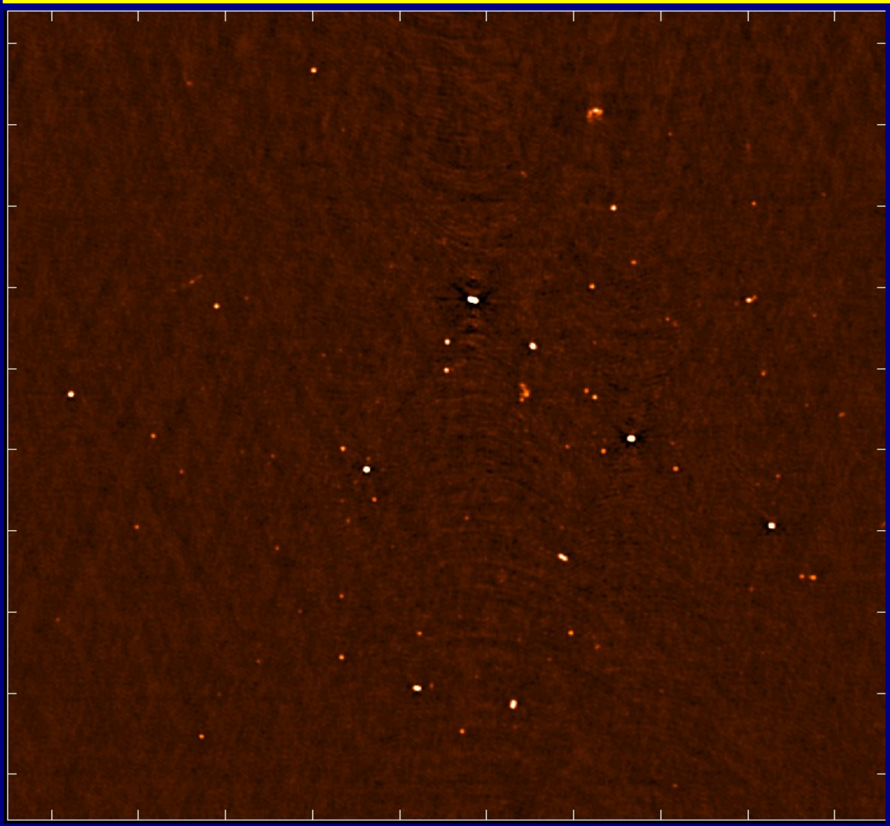
Sky: More complex than point sources



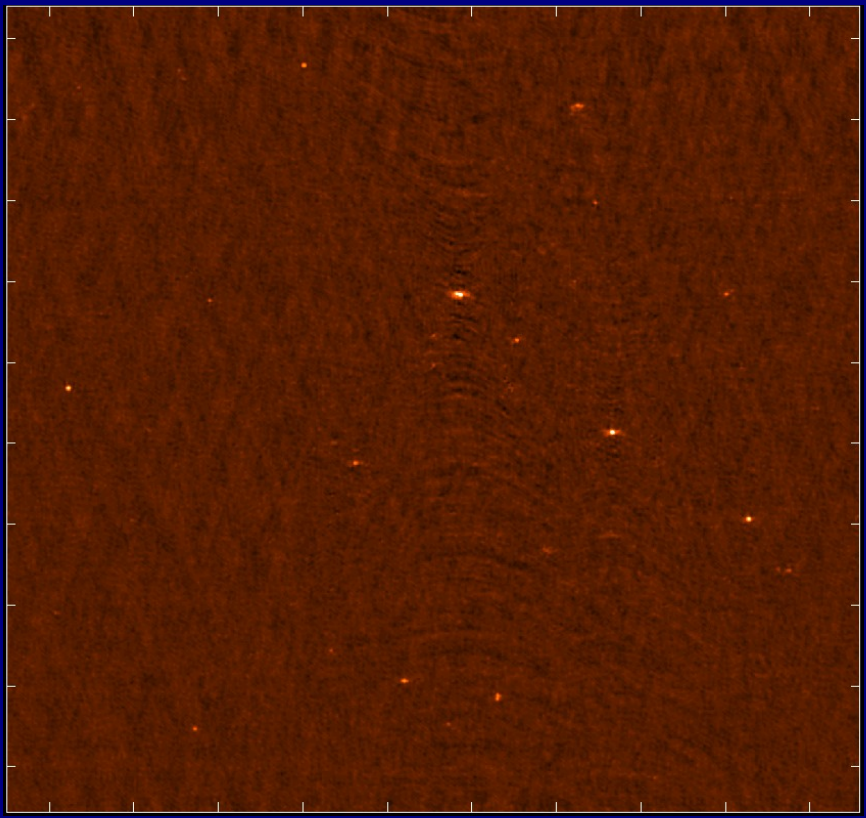
- 1.4 GHz/VLA, 30 μJy/b @ 4”

(Data from: Fomolant et al.)

Sky Frequency dependence



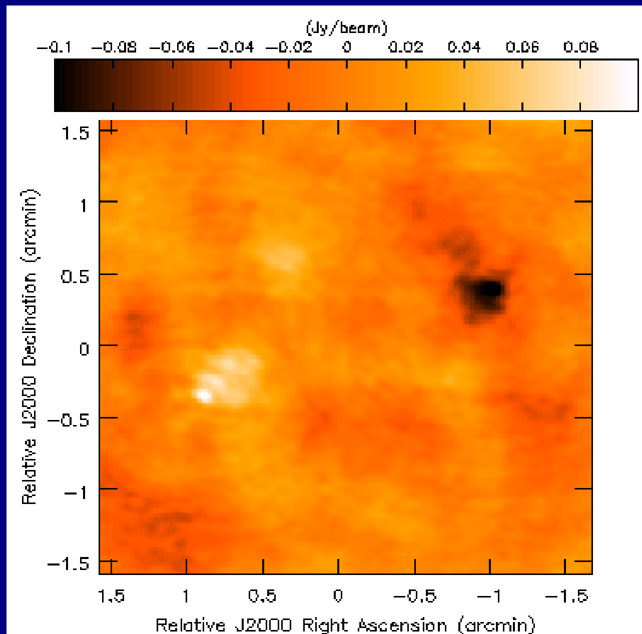
1.365GHz



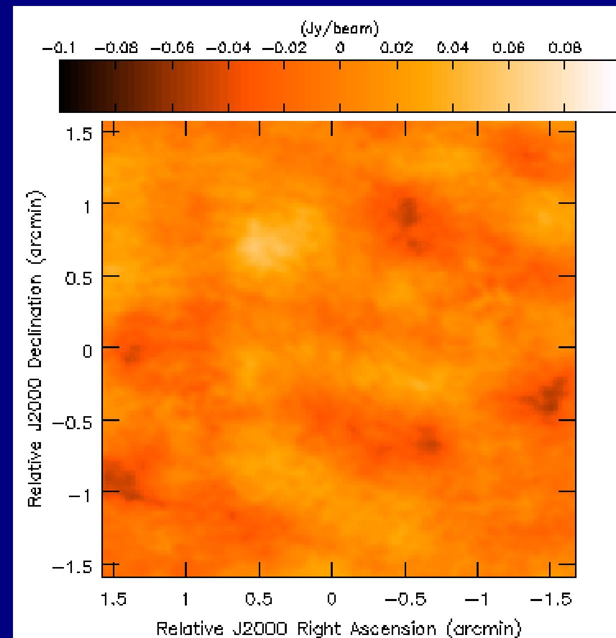
I(1.365GHz)-I(1.435GHz)

- Direction & Frequency Dependent errors
  - Sky spectral index? PB effects? Pointing? Pixelation errors?
- Errors not coherent across frequency
  - Will affect spectral line signals (EoR)

Extended Emission

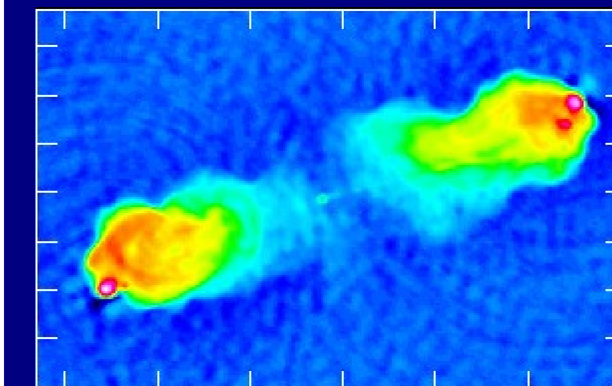


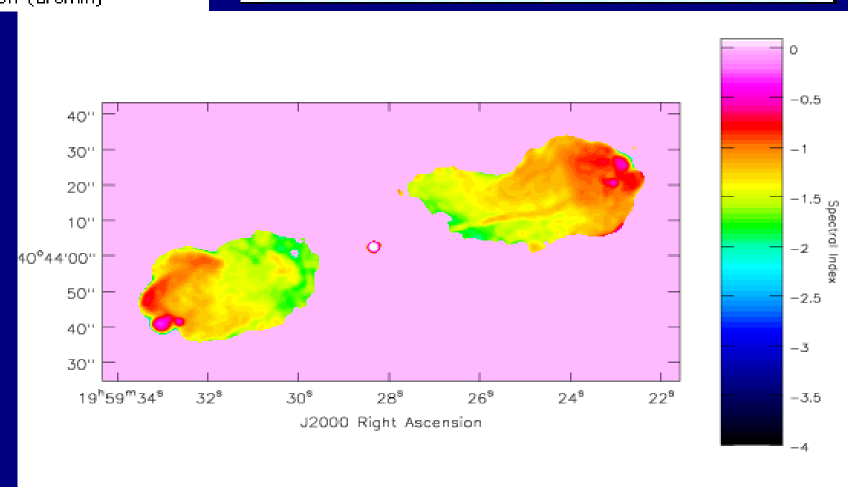
No PB correction



PB correction

(Bhatnagar et al, A&A, June 2008)

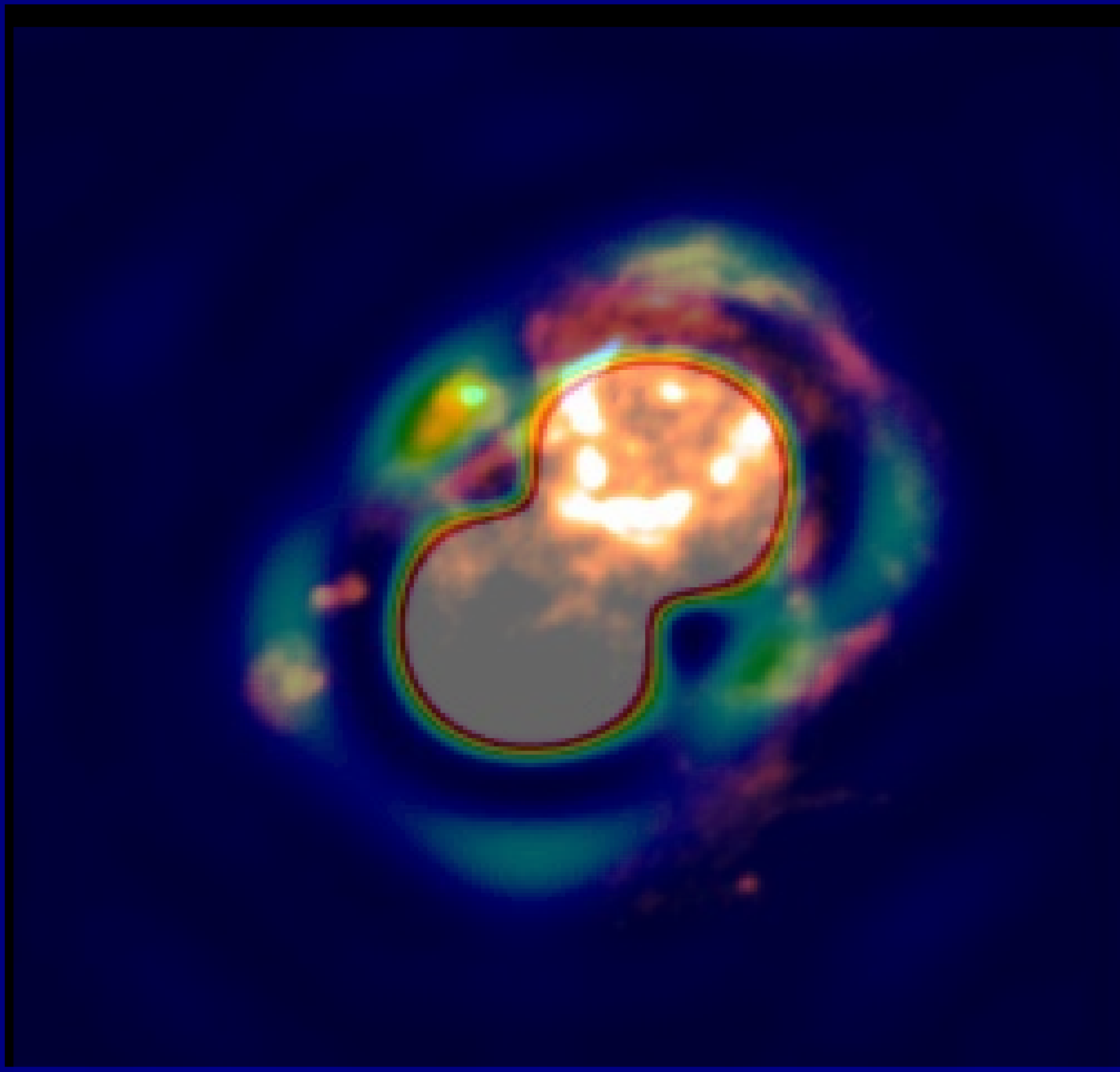


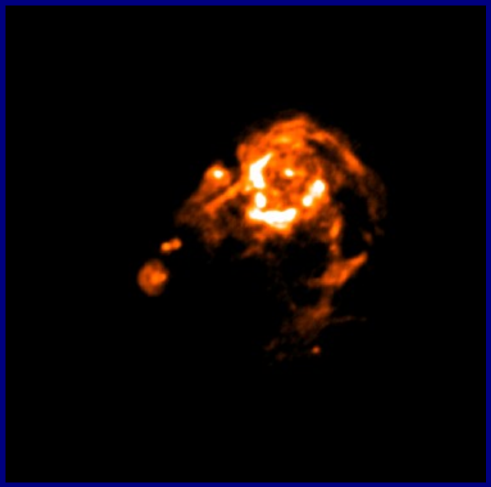


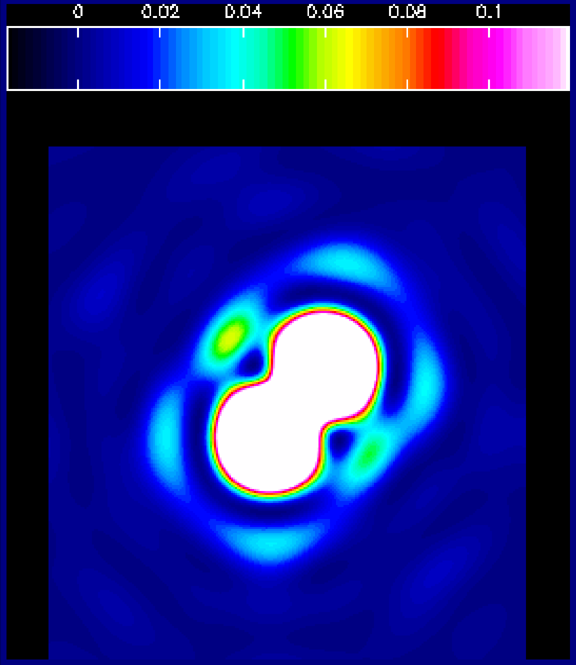
Sp. Index Image (Carilli et al.)

- Stokes-V imaging of extended emission
  - **Algorithms designed for point sources will not work**
  - **Need more sophisticated modeling of the extended emission**

Dominant errors in mosaicing: PB effects







PB errors: Full beam imaging limits

- Limits due to rotation of asymmetric PB
  - In-beam max. error @~10% point
  - DR of few x10^4:1
  - Errors larger in the first sidelobe
- Limits due to antenna pointing errors
  - In-beam max. error at half-power points
  - DR of few x10^{3-4}:1
  - Limits for mosaicking would be worse
    - Significant flux at half-power and side-lobes for many pointing

Computing & I/O costs

- Higher sensitivity ==> more data + correction of more error terms
  - Needs more sophisticated parameterization
  - Significant increase in computing and I/O loads
- Imaging:
  - Correction for PB variations, Pointing errors, ionosphere
  - Better modeling of extended emission
- Calibration: solve for direction dependent effects
  - As expensive as imaging
  - PB shape, pointing, ionosphere
- Processing cost dominated by forward and backward transforms (gridding)
  - I/O time comparable to computing time

Algorithmic issues: Things to think about...

- **Imaging and calibration are coupled**
  - Not possible to produce corrected visibilities independent of imaging
  - Solvers for DD effects requires de-gridding operation
- **Multi-frequency synthesis**
  - Short cuts for DR > 10^4 will not work
- **Mosaic imaging problems are similar in principle**
  - More complicated in practice
- **Near-future data allows higher DR than is allowed by existing algorithms/software**

Algorithm Development needs

- **Algorithm integration**
  - PB-correction: Freq. Scaling, Rotation, Pointing... [Algo. exist]
  - Multi-freq. Synthesis [Algo. exist]
  - Scale-sensitive deconvolution [Algo. exist]
  - Integration required [Requires help!]
    - All the above individually limit DR to Few X 10^4
- **Algorithms R&D for what has been promised!**
  - Full-beam full-sensitivity imaging
  - Some progress, lots of ideas – but require help!
- **Pipeline processing**
  - Auto-flagging (Manual flagging of 1 TB worth of data!? No feasible)
  - Significant research and development required

General Structure of algorithms

- For all iterations
  - For all Channels and Polarization
    1. Compute Residuals [Data – Model] a.k.a. “major cycle”
    2. Compute Gradients
    3. Update Model

- Classical deconvolution
  1. [2 x Gridding Operation + 1 Full data read] per Major Cycle
  2. Minor Cycle: 2x FFT + ...
- Classical Selfcal
  1. 2 Full data read per iteration

- DD Selfcal: 1 Full data read + N_{iter} N_{par} x Gridding operations

Computing & I/O costs

- $$DataSize = \frac{N_a * (N_a - 1)}{2} \frac{T}{\delta T} \left[N_{ch} \;\; N_p [2 * SoF + \frac{SoWt}{N_p}] + 4 \;\; SoF \right]$$
  - For EVLA: 0.5-1.0 TB + 0.5GB

- $$FlOp \;\; per \;\; gridding = \frac{N_a * (N_a - 1)}{2} \frac{T}{\delta T} \left[N_{ch} \;\; N_p \;\; N_{IP} \right] \left[N_{op} S^2 \right]$$
  - One gridding (Major Cycle) will take 1.5-2hrs.
    - Computing efficiency: 10-20% of the rated GFLOPs
  - @100 MB/s, single read of 1 TB data will take ~3hrs.

- Total full data accesses: 10-20

Computing & I/O costs

- Computing scales linearly with N_{ch} , N_p and S^2
  - Convolution support size larger for DD correction (e.g. PB)
- DD calibration
  - Required for what has been promised!
  - N_{iter} N_{par} x [Gridding operations + 2 x full data reads]
- PB-correction+Multi-frequency Synthesis: $$I(\nu)=I(\nu_o)\left(\frac{\nu}{\nu_o}\right)^{\alpha}$$
  - Taylor expantion: N_{terms} depends on the required DR
  - N_{iter} N_{terms} x 2 Gridding Operations + full data read

Computing & Algorithms

- Hard to get away from FFT based forward and inverse transforms
  - Only "peeling" approach not feasible (Noordam, Uson&Cotton,...)
    - Requires 10K-100K components DFT for a 1 TB data base!

- Better understanding of error propagation can lead to efficient algorithms
  - All algorithms (Calibration & Image Deconvolution) are function minimization algorithms (Steepest Descent in fact!)
  - But need to invest and believe in R&D!

- Compute for the allowed dynamic range
  - Computation more accurate than the allowed DR is a waste of resources

Computing Options

- Multi-core multi-CPU machines (4 x dual- or quad-core)
  - Use OpenMP technology to speed up computing (available in GCC 4.x)
    - Least work but requires experimenting
  - Not very helpful beyond 4-6 threads
  - Helps I/O?
- Cluster with multi-core multi-CPU nodes
  - Use MPI at higher software layers
  - Spectral line imaging is embarrassingly parallel (almost)
  - Continuum imaging requires some communication
- Specialized H/W? (FPGA, GPU,...but similar bottlenecks)
  - Bus bandwidth is the bottleneck (Disk->RAM, RAM->CPU)

Data I/O Options

- Central large storage – Disk Raid
  - Can deliver up to 1GB/s (I think)
- Smaller local disks at the cluster nodes
  - Up to 100 MB/s

- Disk-Raid to Node-Disks bandwidth is one of the bottlenecks

- Beyond a certain imaging DR, cluster inter-connect might be a bottleneck (in the non-embarrassingly parallel regime)
  - Astronomical Algorithms R&D required
  - CS'ish R&D might be required

Software Development needs

- **Put together a software architecture and stick with it**
  - Review it periodically
- **Keep it as simple as possible, but not simpler**
  - E.g. Start with the spectral line imaging problem
  - But be careful to not design out the solutions for less straightforward problems
- **Carefully choose technologies and third party tools**
  - Resist the temptation to play with the latest toy on the shelf
  - Resist the "if it was not done here, it is not right" trap!
- **Keep "system level" layers thin**
  - Most popular UI do this : User String --->[UI-layer]---> App-layer
  - Should not require more UI-layer software than App-layer!