Post-processing Algorithmic Challenges

NG-VLA Technical Workshop ,
Socorro, Dec. 8th 2015

S. Bhatnagar
NRAO

Current Challenges

- Current telescopes that provide useful experience for the frequency range of NG-VLA: SKA1, EVLA (1 – 50 GHz) and ALMA (100s of GHz)

- Unique to NGVLA
    - Higher sensitivity
    - Wider-band, higher-N, longer-B
    - Unexplored frequency range

Imaging: Basic parameters

- Frequency range : 1(?) – 100 GHz (EVLA: 1 – 50 GHz)
- Imaging dynamic range: $10^{3-4\,(7?)}$ (EVLA: few x 10^{6} : 1)
- FoV: 10s arcmin – 10s arcsec
- Mosaic imaging
- Resolution: $10^{-\{1-3\}}$ arcsec
- Imaging of resolved sources

- NG-VLA is in-between EVLA and ALMA
  - Current Sc. Cases: > 10 GHz, DR < 10000:1
  - For 1 – 10 GHz range, capable of DR > 10x EVLA

- Computing, memory footprint:
  - Images: 10x larger (100s Kilo pixels on a side)
  - Data volume: 3 – 5 orders of magnitude larger than EVLA

Current Challenges

- Summary of current challenges/R&D
  - Freq. Dependence of sky brightness distribution
  - Effect stronger of NGVLA (thermal emission)

  - Wide-band Wide-field imaging
    - Effects of WB PB, Pointing Errors, W-Term

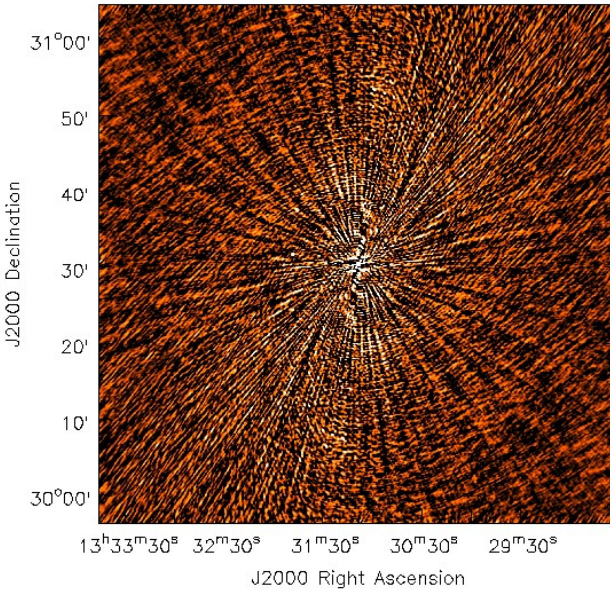
    - Wide-field wide-band polarimetric mapping
      - Corrections for in-beam polarization (WB)
      - Faraday Rotation Synthesis

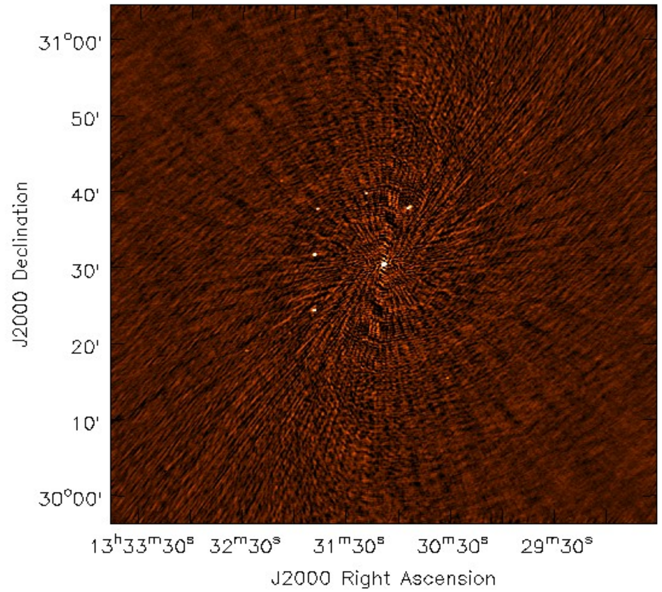
  - Computing load, Memory footprint, use of heterogeneous-HPC

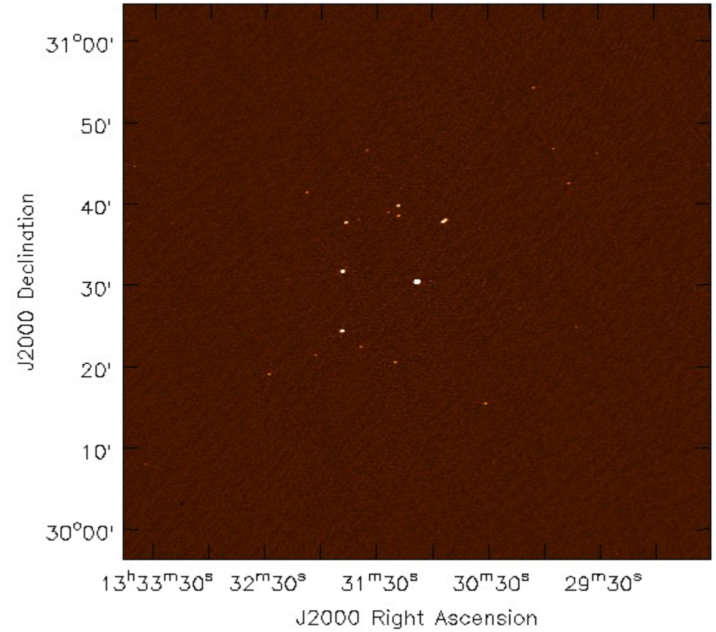
  - Non-isoplanatic atmospheric effects at high frequencies

Wide-band Effects

- Wider-band observations: 10—20 GHz across 1(?) – 100 GHz
  - Affects all continuum imaging: wide-field and narrow-field
  - Spectral Index effects limits DR to few x 1000:1
  - WB imaging of extended emission requires MS and MT reconstruction : NGVLA memory footprint is prohibitive







Wide-band Effects

- NG-VLA: Spectral variations stronger/complex (non-thermal + thermal emission)
  - Better techniques to model spectral variations
  - Better techniques to model spatial variations

  - Current tests up to 4 Taylor Terms for simple fields
    - Model spectral variations as a polynomial in frequency
    - Can become numerically unstable for large number of terms.

- Current implementations proven to be numerically advantageous but has high memory footprint and computing load that is difficult to parallelize
  - Possible alternate approaches (component based)

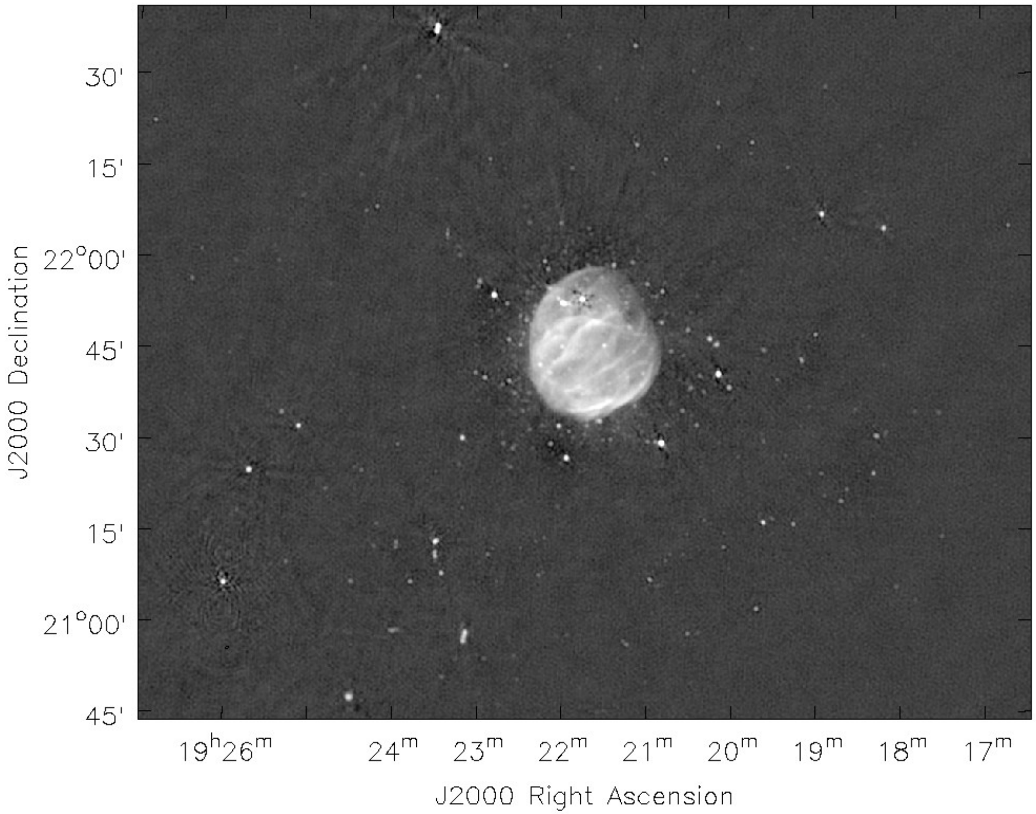
Wide-field Effects: W-Term

- Continuum sensitivity pattern spreads FoV to 2x – 3x PB

No. of facets due to W-Term

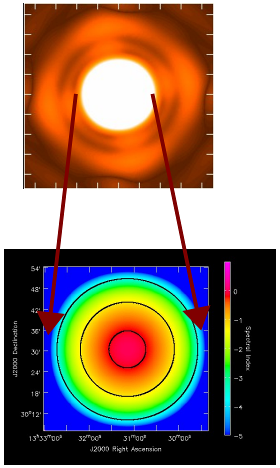
GHz	EVLA	NGVLA
1	15	400
10	1.5	40
50	0.3	8
100	-	4

- Benefit from developments for EVLA and other low frequency telescopes

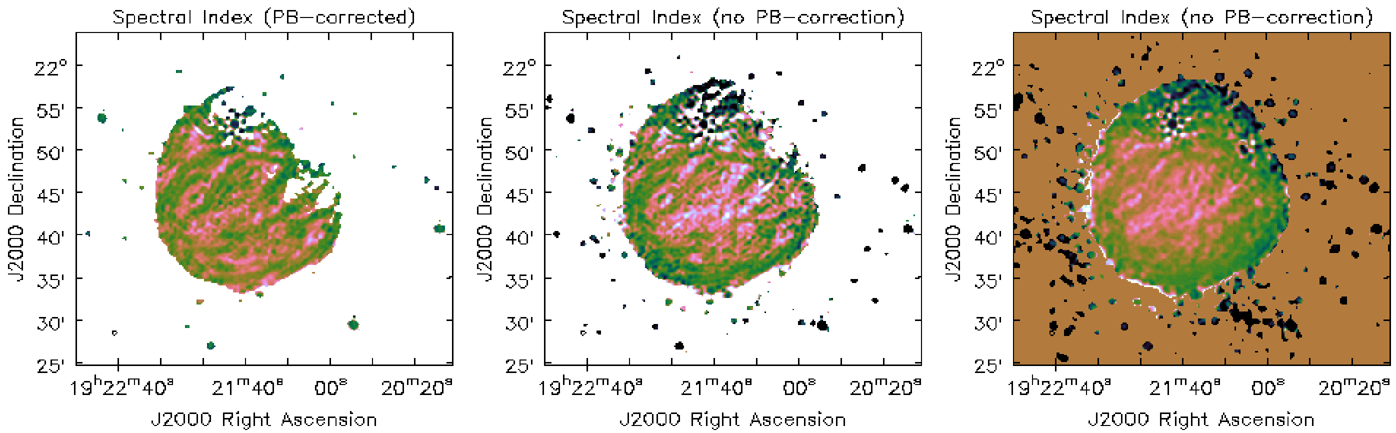


Wide-field Effects

- Frequency dependence of PB
  - DR limits: few x 1000:1
  - Is idealized model (scaling with freq) sufficient for NGVLA?

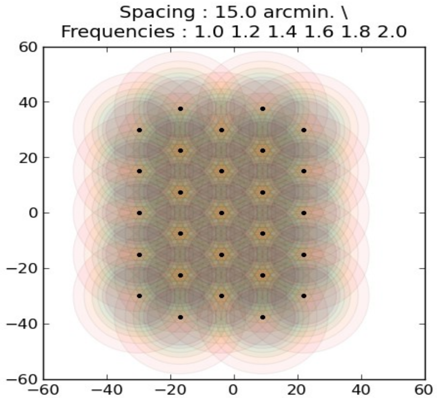


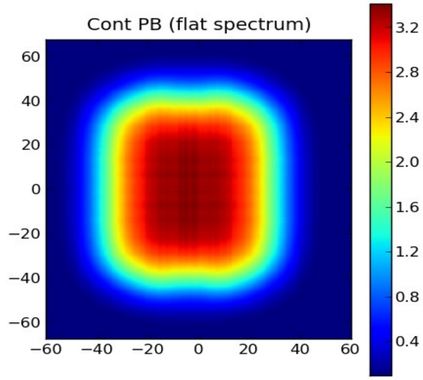
PB “Spectral Index”

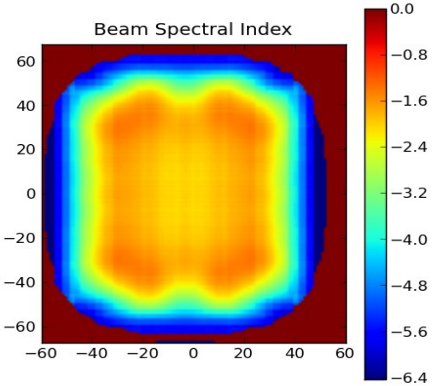


Wide-field Effects

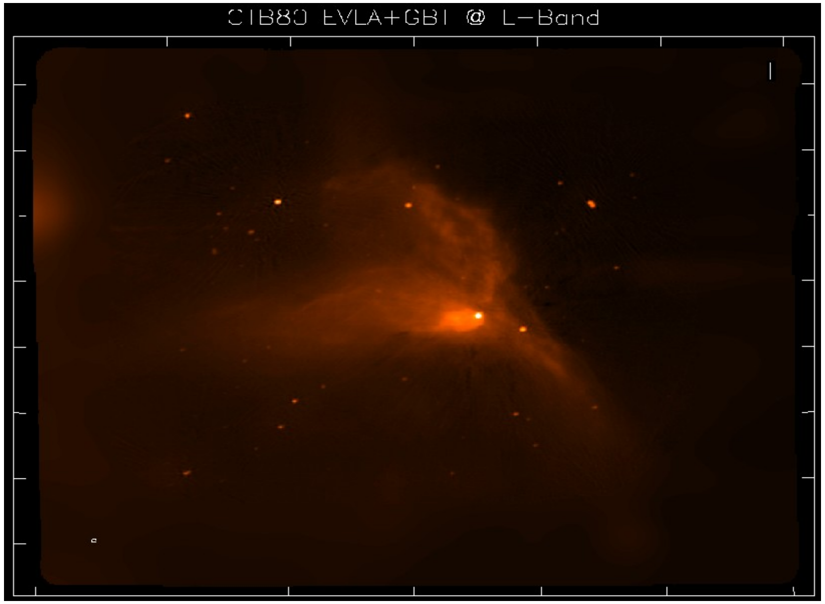
- Wide-band continuum mosaic imaging







PB frequency dependence Spreads across the FoV



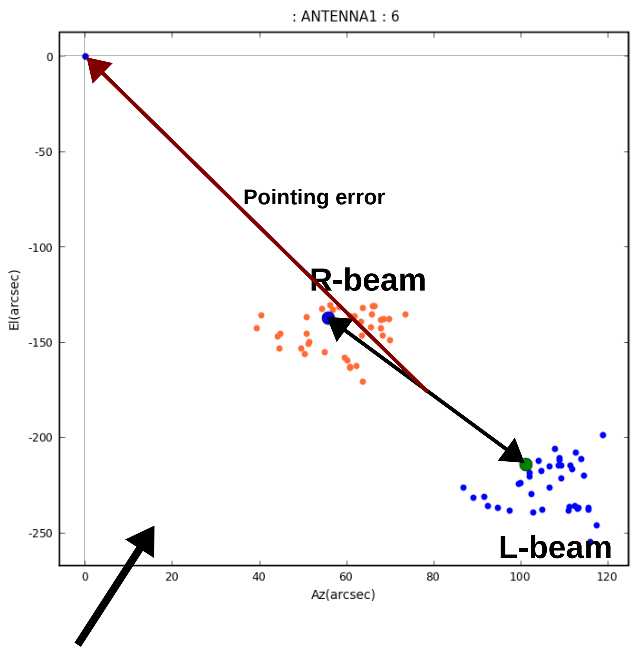
- **100-pointing EVLA mosaic**
- **Requires combining WB single dish data to map spectral index across the field**
  - **A work in progress...**

Wide-field Effects

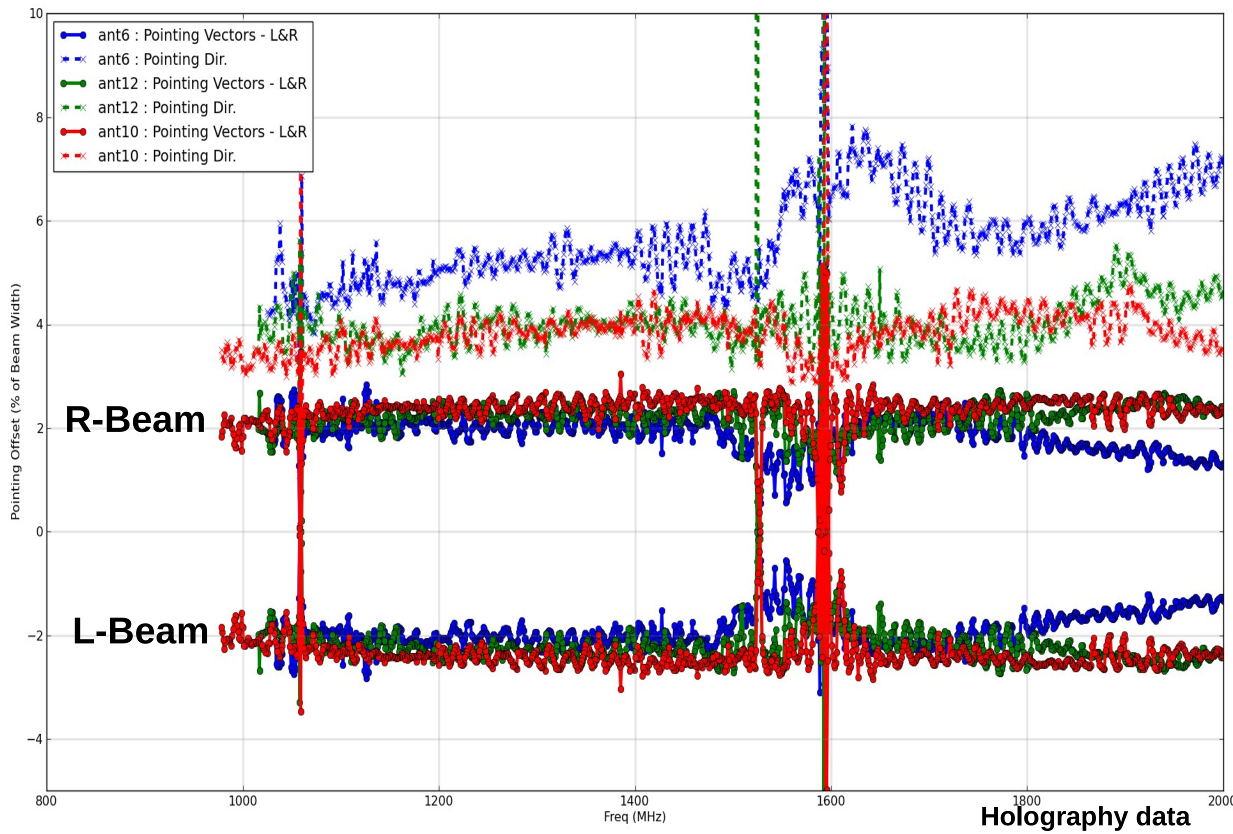
- Pointing errors: E.g., squint: ~5.6% of PB (EVLA)

Varies across the band. Does it matter?

Limits Stokes-I DR ~10000:1



P.S. Solutions at L-Band



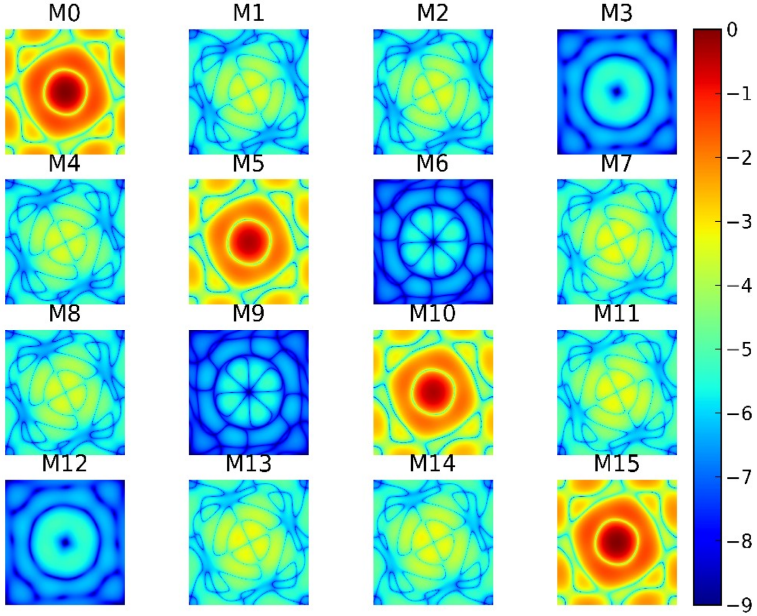
Holography data
Courtesy R. Perley
Analysis: Jagannathan

Full pol. Imaging: In-beam effects

$$V_{ij}^{Obs} = \left[J_i \otimes J_j^*\right] . \left[V_{ij}^o\right] = \left[M_{ij}\right] . \left[V_{ij}^o\right]$$

$$\begin{bmatrix} V_{pp}^{Obs} \\ V_{pq}^{Obs} \\ V_{qp}^{Obs} \\ V_{qq}^{Obs} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix} . \begin{bmatrix} V_{pp}^{o} \\ V_{pq}^{o} \\ V_{qp}^{o} \\ V_{qq}^{0} \end{bmatrix}$$

- Limits fidelity: strong effect at < 20-25% PB-level



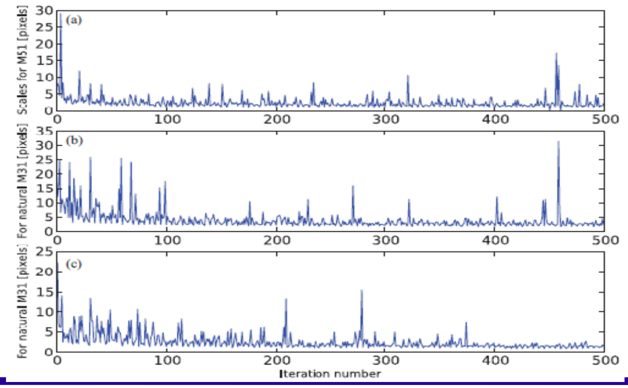
(Jagannathan, PhD Thesis)

R&D for NGVLA Imaging

- Robust pipelines!
  - Needs systemic move away from processing “by hand”
  - Seamless HPC
  - Condense information for human consumption/intervention
  - Fault tolerant: in the input data, output products
    - Fault reports for human consumption (as opposed to “computer brain dump”)
  - Heuristics to trigger optimal algorithms (not always the most expensive ones)
  - Develop algorithms to take advantage of large-N

PSF far sidelobes ~1/N_{ant}

Near sidelobes are higher



Zhang et al., ApJ(submitted)

R&D for NGVLA Imaging

- Wide-band wide-field imaging is expensive (but required!)
  - A- + W-Projection
    - Computing load and memory footprint for filters (CF) is prohibitive
    - Solutions: GPUs, FPGAs. Other approaches?
  - Is current approach to MS + MT-MFS sufficient (model spectra as a polynomial)?
  - Is it efficient?

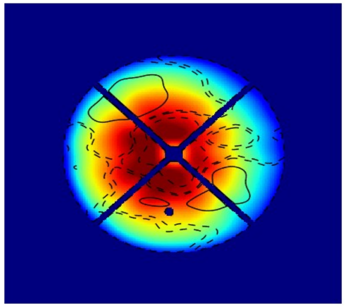
  - Wide-band Interferometric + Single Dish Imaging
  - Necessary to achieve imaging performance for diffused extended emission

- Time-domain
  - What is the optimal approach?
  - Account for time-variable sky brightness? Bi-spectrum? Combination? **(Prototype algorithm, Rau et al.)**

R&D for NGVLA Imaging

- Effects of wide-band antenna far-field pattern

- 300 antennas, higher frequency spread over 300Km: most likely antenna-to-antenna variations (ALMA example)
    - Develop parameterized models that also capture these variations to the appropriate level
    - What is the appropriate level for NGVLA?



- Full-stokes imaging
    - Characterize in-beam effects.
    - Will learn about limits from work for EVLA/ALMA
- Atmospheric effects
    - Characterize (phase structure function)
    - A hard, largely unsolved problem
    - Synergy with existing telescopes

R&D for NGVLA Imaging

- Antenna pointing errors
  - Pointing SelfCal: Required? Possible? Solution interval?

Noise Budget:

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

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

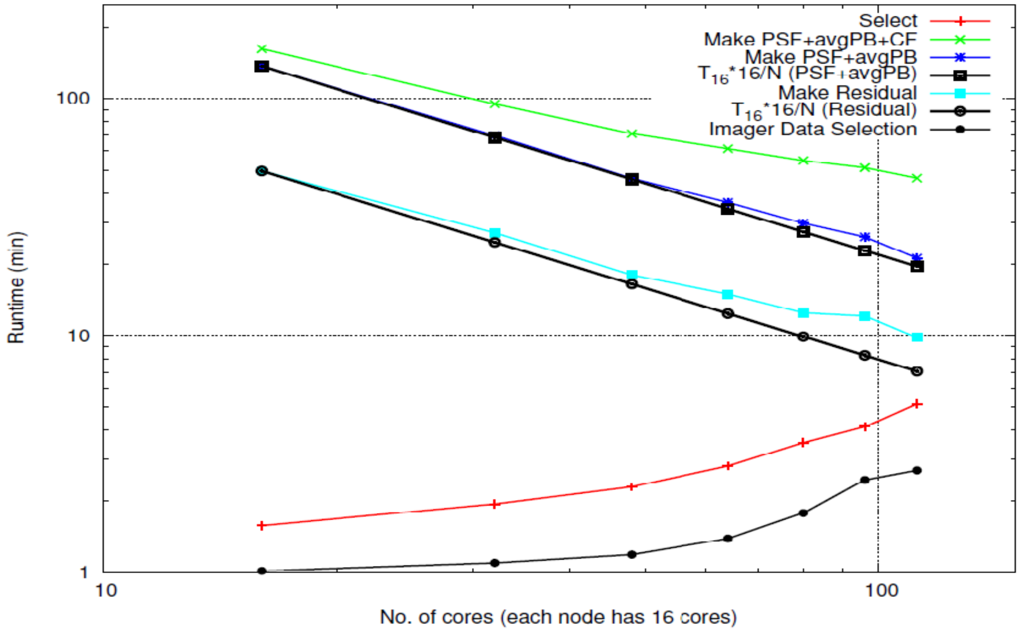
  - Characterization (Kundert et al. IEEE Ant&Prop, in prep.)
    - Order of errors: PB shape/freq. Dep., Pointing errors/(rotation?), Ant-to-ant variations

- Shape variations (temporal, spectral)
  - In general more serious (non-hermitian)
  - Develop low-order models for PB

R&D for NGVLA Imaging

R&D for NGVLA Imaging

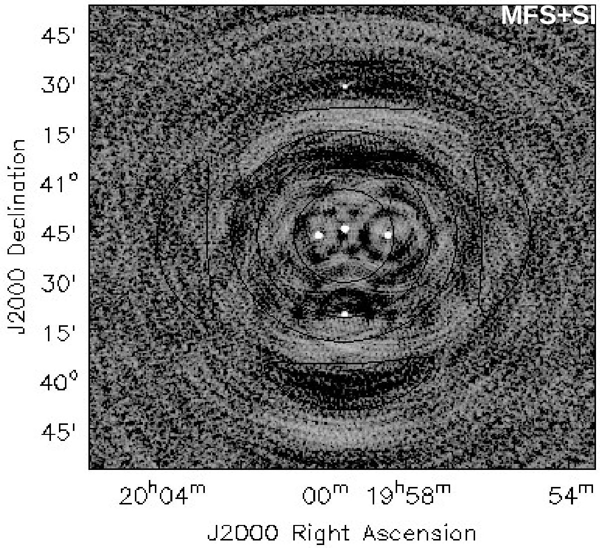
- Computing & Memory footprint:
  - Image reconstruction cost, as done now, scales as
    $N^2_{CF}\ N_{vis}\ N_{taylorTerms}$ NGVLA/EVLA ~ $10^{3\text{-}4}$
  - Memory foot print scales with $N^2_{terms}\ N^2_{Scales}$ x Image size
- Some compute-hotspots deploy well on GPUs/FPGAs
  - Scale from parallel imager (CASA), AWS
    - A 1000-core AWS successful run as a test for a low-order solver

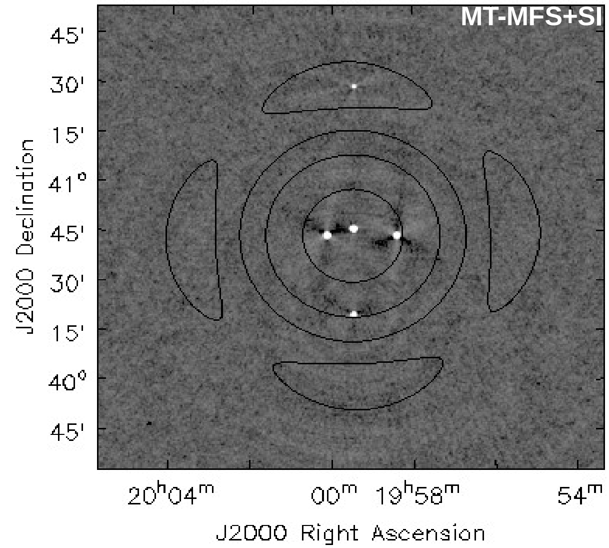


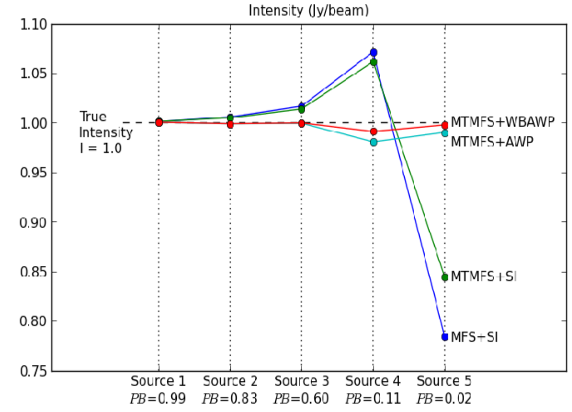
Take-away messages

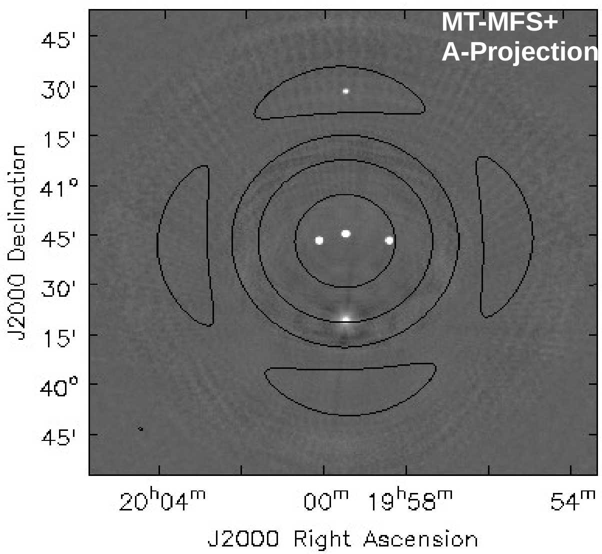
- Develop human resource with multidisciplinary skills
- Will benefit from SKA1/EVLA/ALMA
  - If 10–100GHz range, imaging problem not as bad as for SKA1
- High memory footprint: Review modeling algorithms (a.k.a. “deconvolution” algorithms), storage, display, mining,...
  - Also depends on affordable computing h/w
- Always non-coplanar & wide-band: W- + A-Projection + MTMFS: Software for heterogeneous h/w to mitigate current bottlenecks.
- Wide-field full-pol imaging : Learn from work in progress...
- Scalable algorithms/software
  - Multi-node+massively parallel accelerators or massive cluster + memory ...
- Heuristics for auto-tuning.

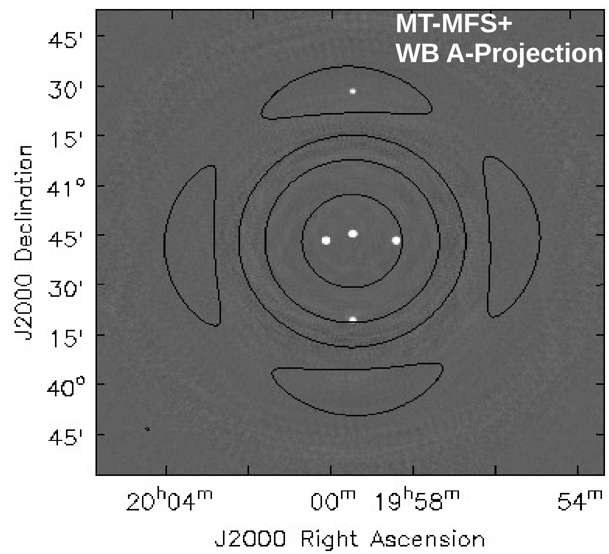
Simulations

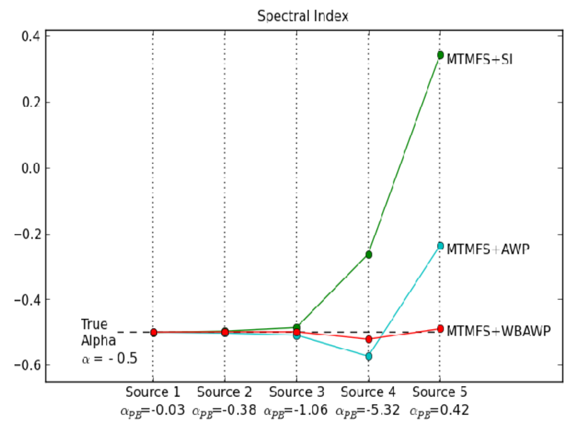












Bhatnagar, Rau & Golap: ApJ, 2013

Pointing SelfCal

- Pointing errors: ~20 – 30 arcsec.

  Residual Pointing error at 10GHz: 2% of PB

  Time variable: Ref. Pointing time-scale: ~30 min.

  Estimated DR limit: ~10000:1