Advanced Algorithms

March 21st 2023

Sanjay Bhatnagar

Algorithms R&D Group, NRAO

Plan for the talk

- Re-cap of imaging, need for new/advanced algorithms
- Advances in algorithms for the current telescopes
  - Algorithms for wide-field imaging/ Direction-dependent corrections
  - Algorithms for wide-band imaging
- Computing cost issues to understand and keep in mind
  - Computing complexity / computing resource requirements
  - Understand the algorithmic needs of your scientific goal: Not all imaging needs to trigger the most advanced algorithms
  - Use implementations that allow flexibility in algorithmic choices and combinations
- Next-gen instruments and the next-generation challenges
  - SKA
  - ngVLA
    - » 214 antennas, 1000Km baselines
    - » Fractional bandwidth: up to 60%
    - » Range of scales: mas → 10s arcsec

Telescope sensitivity

- Noise limit for imaging with interferometric radio telescopes

$$Noise \propto \frac{T_{sys}}{A_{eff}\sqrt{\Delta\nu\,\Delta T}}$$

10 – 100x improvement in sensitivity over the past decade!

- Sensitivity improvements achieved by

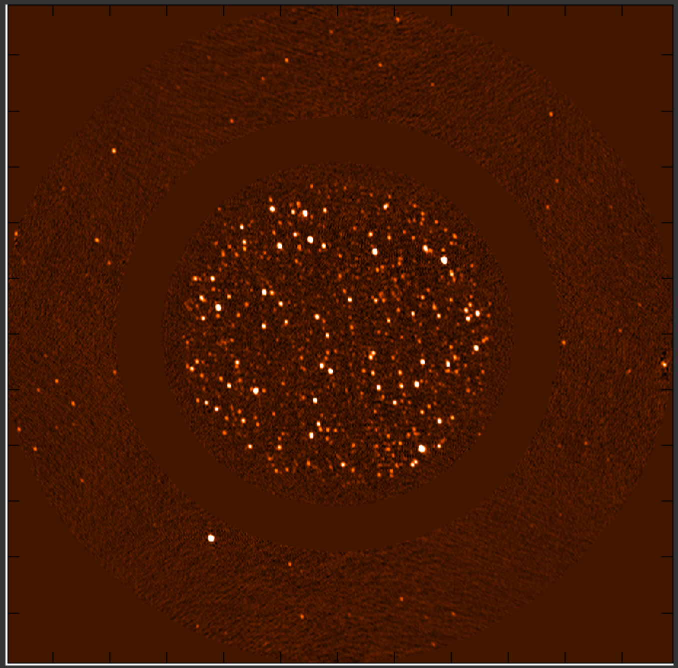
$\Delta\nu$: *Wide band receivers: >60% fractional bandwidth*

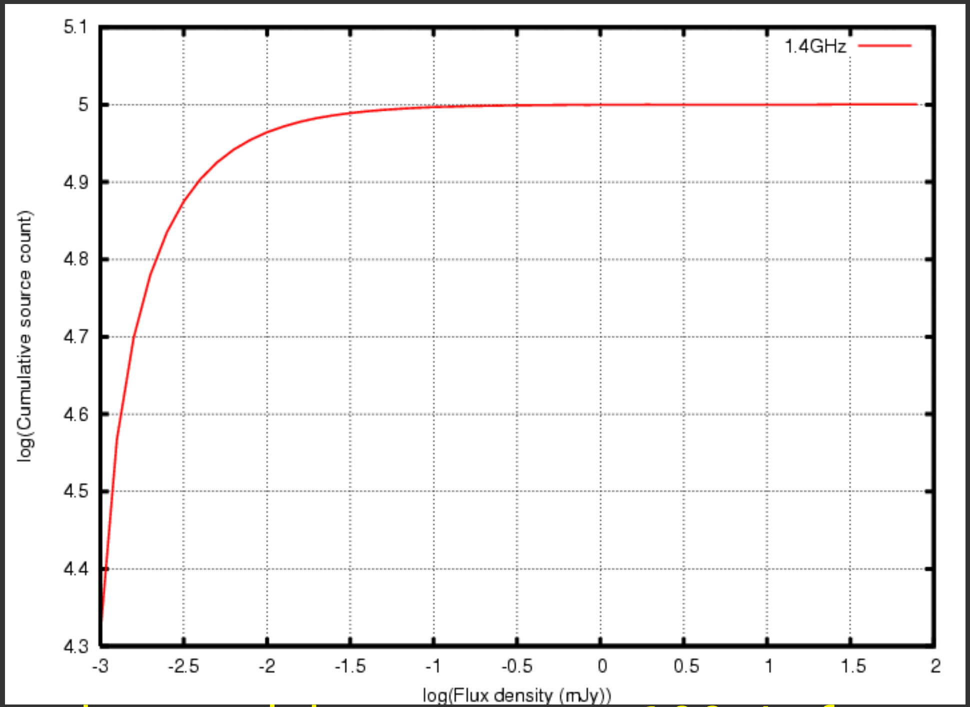
A_{eff}: *More antennas: 30 -- many 100s*

Long baselines: To beat confusion limit

ΔT: *Long integration times: many hours -- months*

Sky at low frequencies: No. of sources

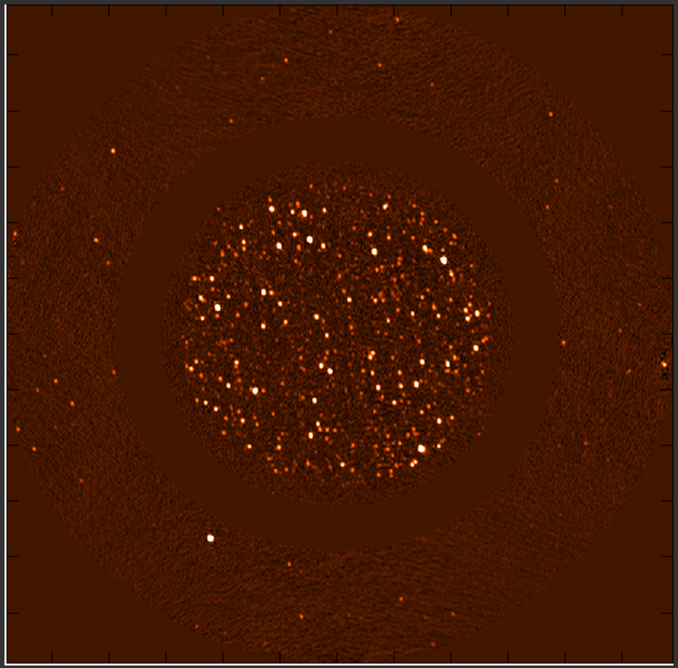


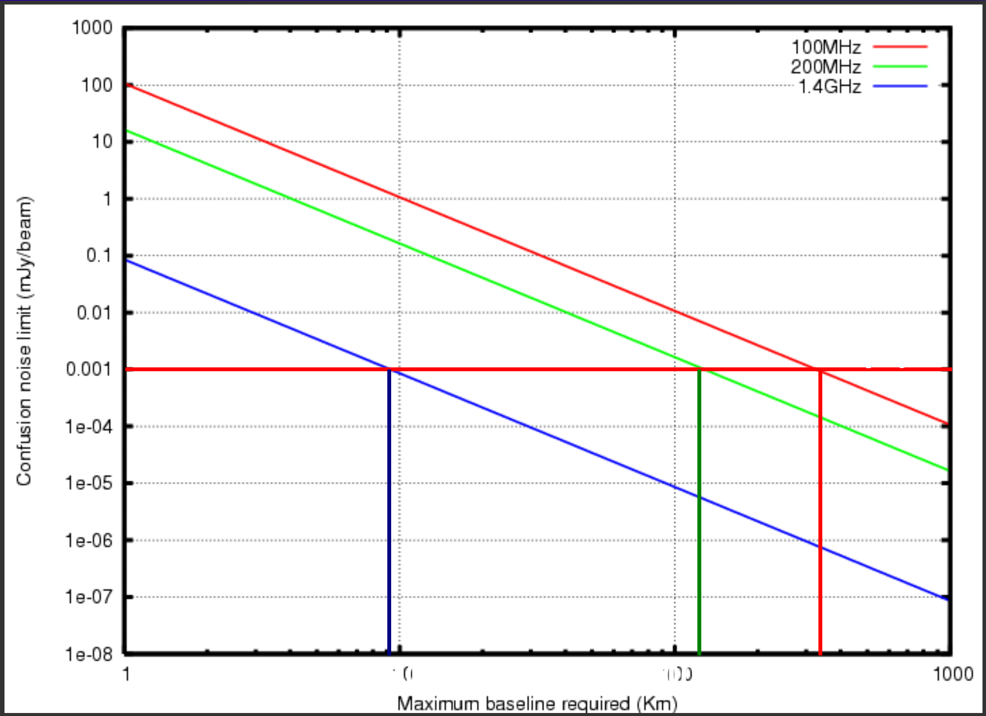


- PSF side-lobe at 1% level requires deconvolving sources >100µJy for 1µJy/beam RMS
- $10^{4\text{-}5}$ sources per deg^2 >10µJy @1.4GHz
  - Source size distribution important at resolution < ~2"
- Implications for imaging
  1. Wide-field imaging
  2. Deconvolution of crowded fields (same problem as deconvolution of extended emission)

     For HDR imaging we need to consider few X 100 mJy – 1 Jy source ~few sq. deg.

Sky at low frequencies: Confusion limit

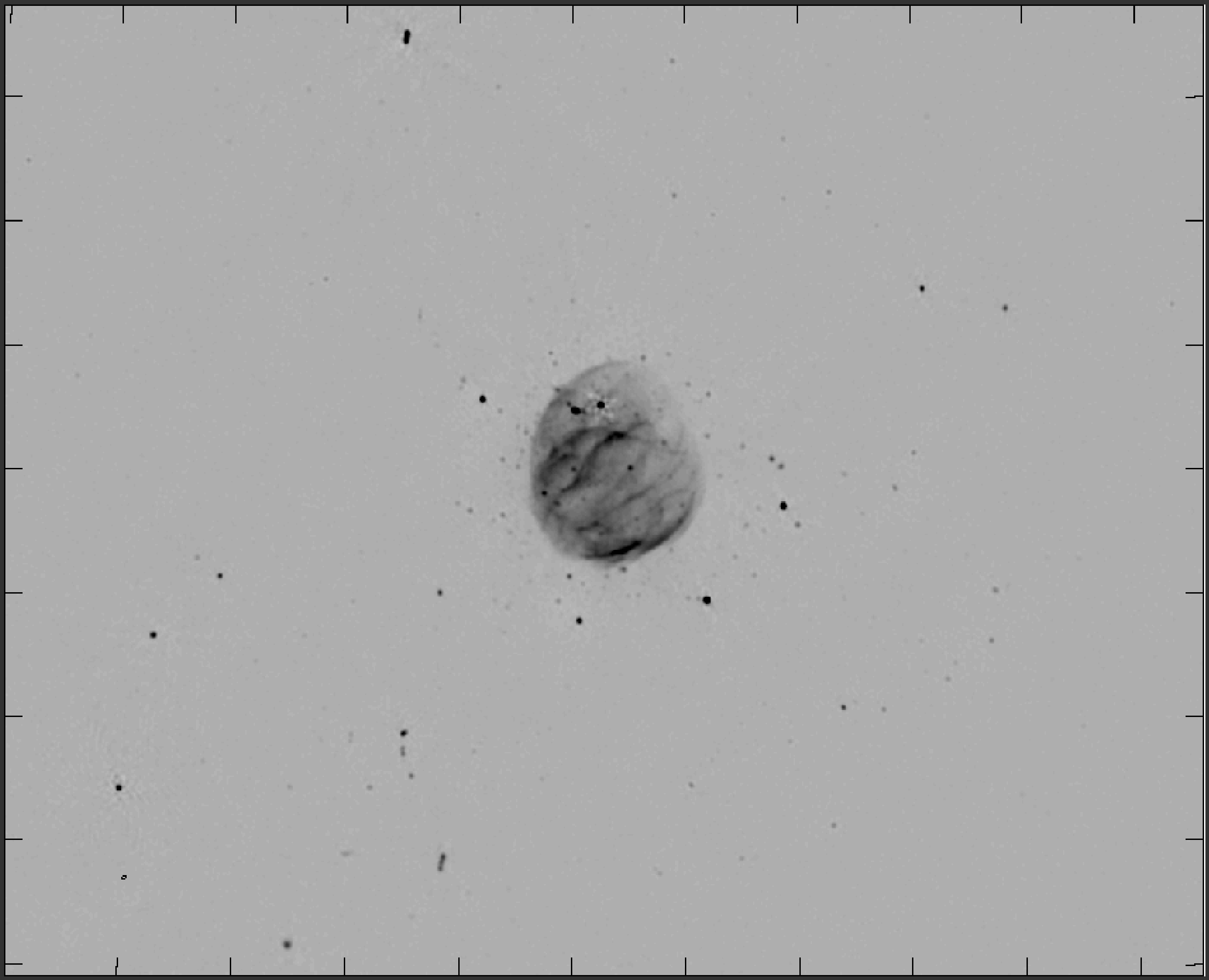




- $\sigma_{confusion} \propto (\nu^{-2.7}/B^2_{max})$: B_{max} ~100 Km at 200MHz for $\sigma_{confusion}$ ~ 1µJy/beam
- Implications for imaging
  1. Long baselines: B_{max} > 2-3 Km & DR > 10^4
  2. Wide-field effects: W-term, PB effects, ionospheric effects
  3. Larger data volume

Wide-field, wide-band, high resolution, HDR imaging using large data volumes is a natural consequence of low frequency and high sensitivity

Wide-band implies Wide-field imaging



- EVLA @L-Band
- BW=600 MHz
  (1.2 – 1.8 GHz)

- Algorithmic Challenge:
  - Time-varying direction-dependent gains
  - Wide-band effects
  - Extended emission with superimposed compact emission
  - Full Stokes + Mosaicking

Imaging challenges

- Challenges in imaging at low frequencies

  1. Wide-field imaging
     - Account for Direction Dependent (DD) effects
     - PB: Time, frequency and poln. dependence
     - W-term
  2. Wide-band imaging
     - All of the above plus...
     - ...frequency dependence of the sky brightness
  3. Sky brightness stronger and complex
     - Multi-Scale deconvolution
  4. Ionospheric effects
     - Requires DD solvers: An algorithmic & computing challenge in itself
  5. Runtime
     - Harvest the vast amount available computing power
     - It is geographically distributed though! ;-)

Direction Dependent (DD) Effects

- DI Calibrated ME

$$V_{ij}^{DI-Cal}(\nu) = \qquad W_{ij} \int P_{ij}(\mathbf{s}, \nu, t) \; I(\mathbf{s}, \nu) \; e^{\iota \mathbf{s}.\mathbf{b}_{ij}} \; d\,\mathbf{s}$$

Data

DI Calibration

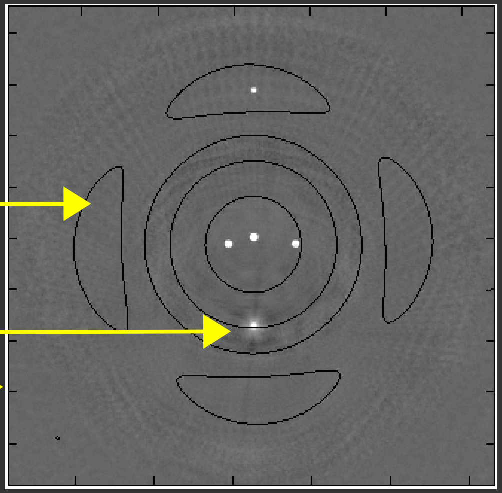
DD Term
Instrumental
Ionospheric

Sky

Geometry

- Removing the effects of the DD terms cannot be separated from imaging

$$I_{continuum}^{Dirty} = \int \int PSF(\nu, t) * \left[PB(\nu, t) \times I^{True} \right] \; d\nu \;\; dt$$



Direction Dependent (DD) Effects

- DI Calibrated ME

$$V_{ij}^{DI-Cal}(\nu) = W_{ij} \int P_{ij}(\mathbf{s}, \nu, t)\; I(\mathbf{s}, \nu)\; e^{\iota\, \mathbf{s}.\mathbf{b}_{ij}}\; d\,\mathbf{s}$$

Data

DI Calibration

DD Term Instrumental Ionospheric

Sky

Geometry

- Standard Imaging assumes:
  - PB is independent of time, frequency and polarization
  - Sky brightness is independent of frequency
  - Geometry is 2D

DD Corrections: Projection Algorithms

- Rewrite $V_{ij}^{DI-Cal}(\nu) = W_{ij} \int P_{ij}(\mathbf{s}, \nu, t)\ I^{True}(\mathbf{s}, \nu)\ e^{\iota\, \mathbf{s}.\mathbf{b}_{ij}}\ d\mathbf{s}$

  as $V_{ij}^{DI-Cal}(\nu) = A_{ij}(\nu, t) * V^{True}(\nu, t)$

- Find an operator X which when applied to the above equation, projects-out the undesirable effects of A?

$$X_{ij}\ V_{ij}^{DI-Cal} = X_{ij}\ A_{ij}\ V^{True}$$

$$such\ that\ X_{ij}\ A_{ij} \approx \mathbf{1}$$

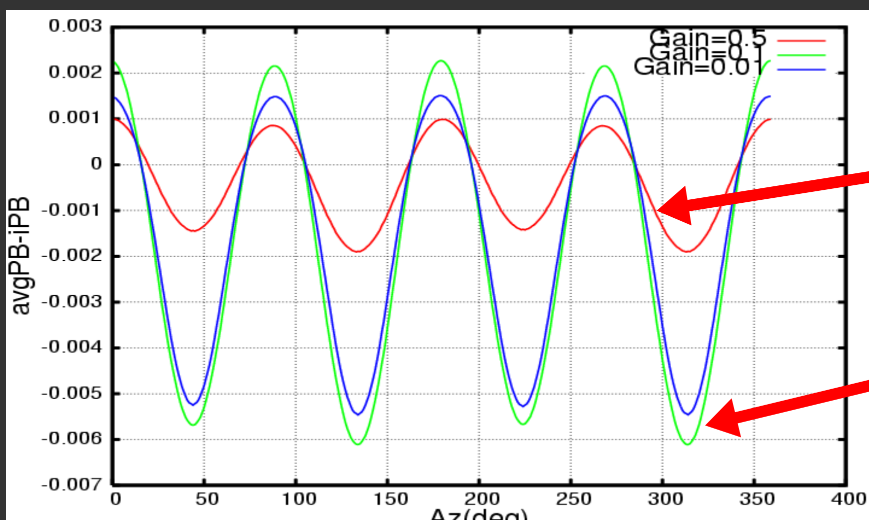
- Then

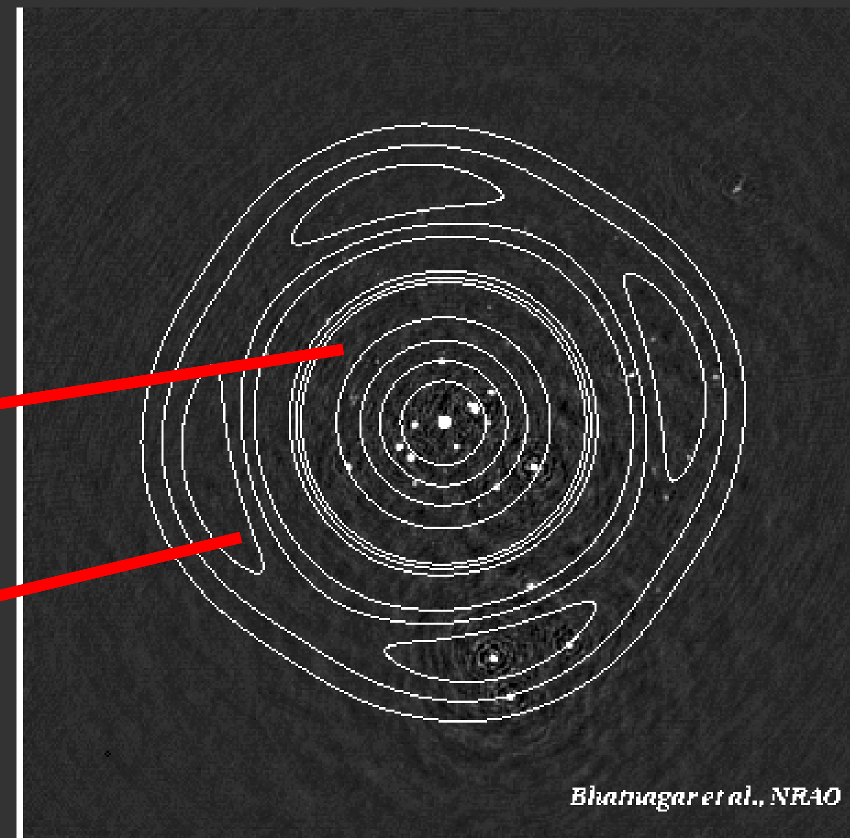
$$F\ X_{ij}\ V_{ij}^{DI-Cal} = F\ V^{True} = I^{True}$$

X encodes the Physics of the problem

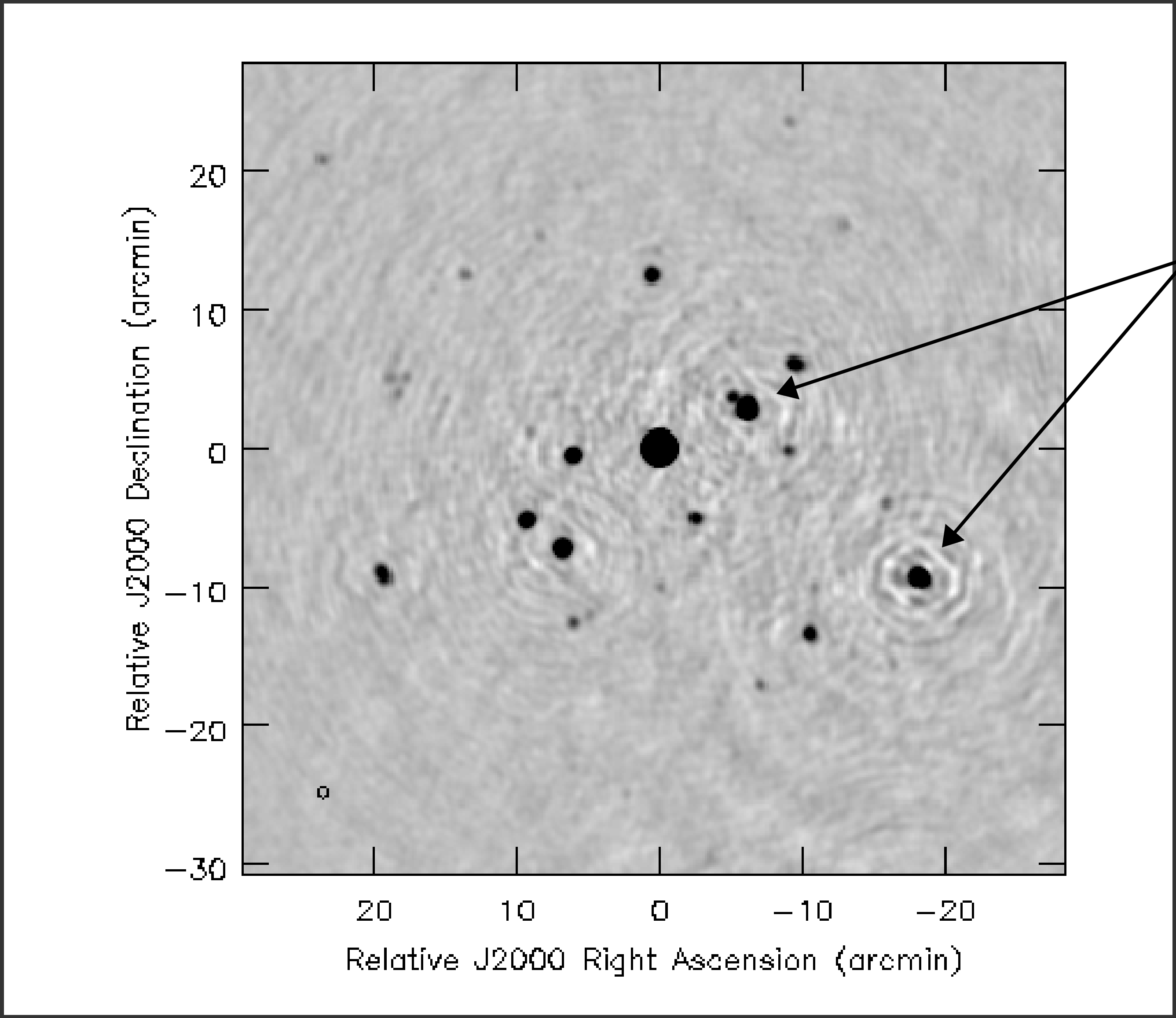
Time dependent terms

- Antenna PB ($The \quad P_{ij}(s,\nu,t)$)
  - Time dependence





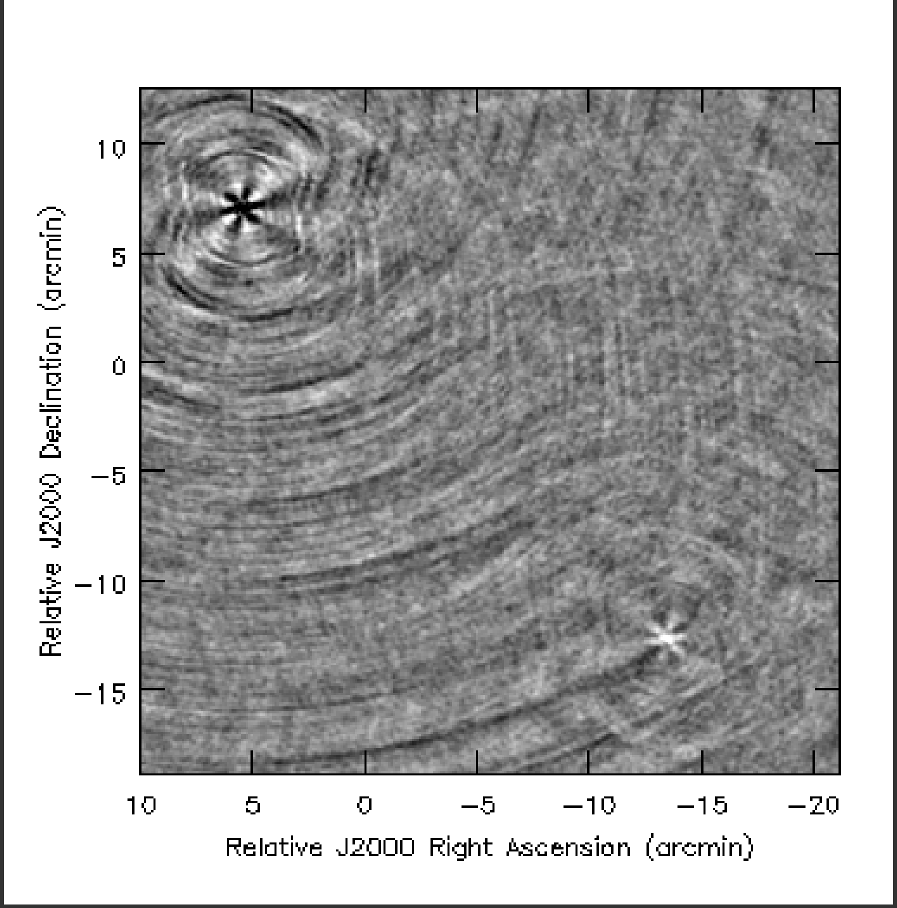
Time+Polarization dependence



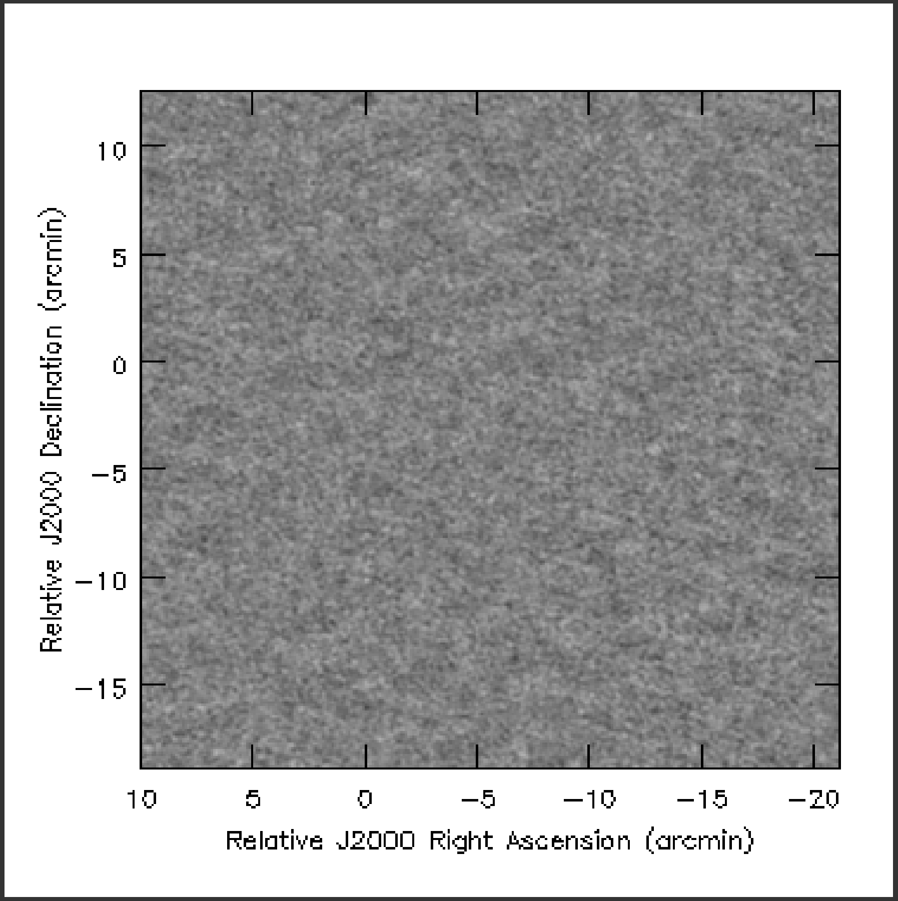
Errors due to PB Squint + Rotation + Pointing errors

PB Polarization Effects

Stokes-V Images



A-Projection →

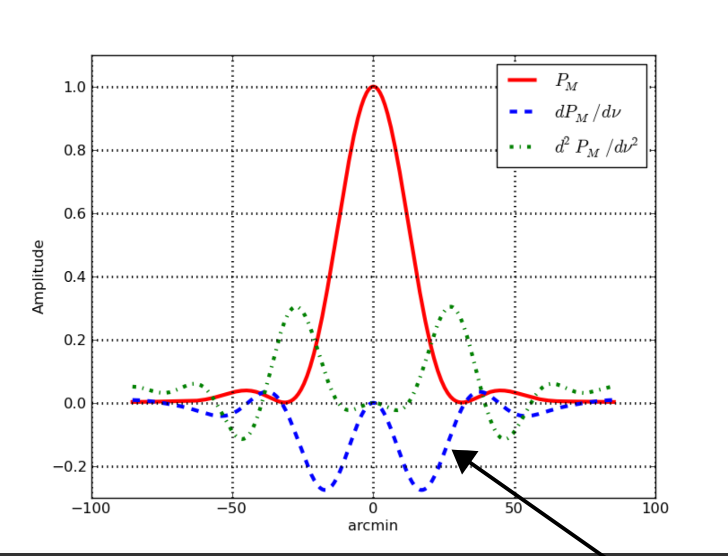


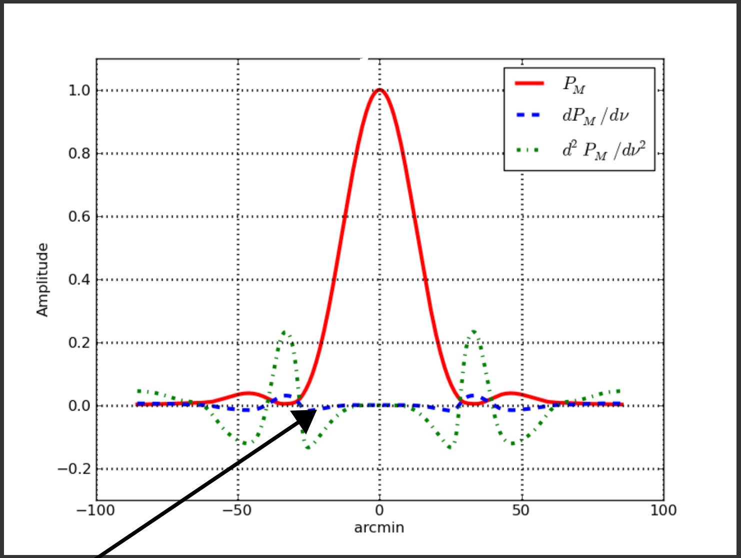
- L-Band VLA imaging
- DR ~ 10^4

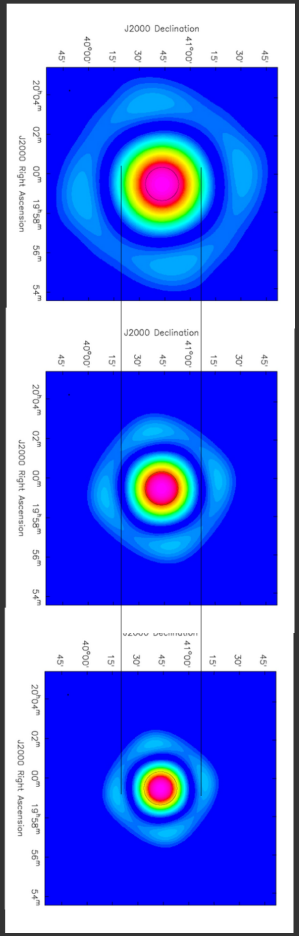
Wide-Band AW-Projection

- Correct for frequency dependence of the PB effects
  - Polarization: Squint + in-beam polarization

$$I^{continuum} = \int P_{ij}(s, \nu, t)\; I(s, \nu)\; d\nu$$







PB Frequency dependence (blue curve)

WB AW-Projection + MT-MFS

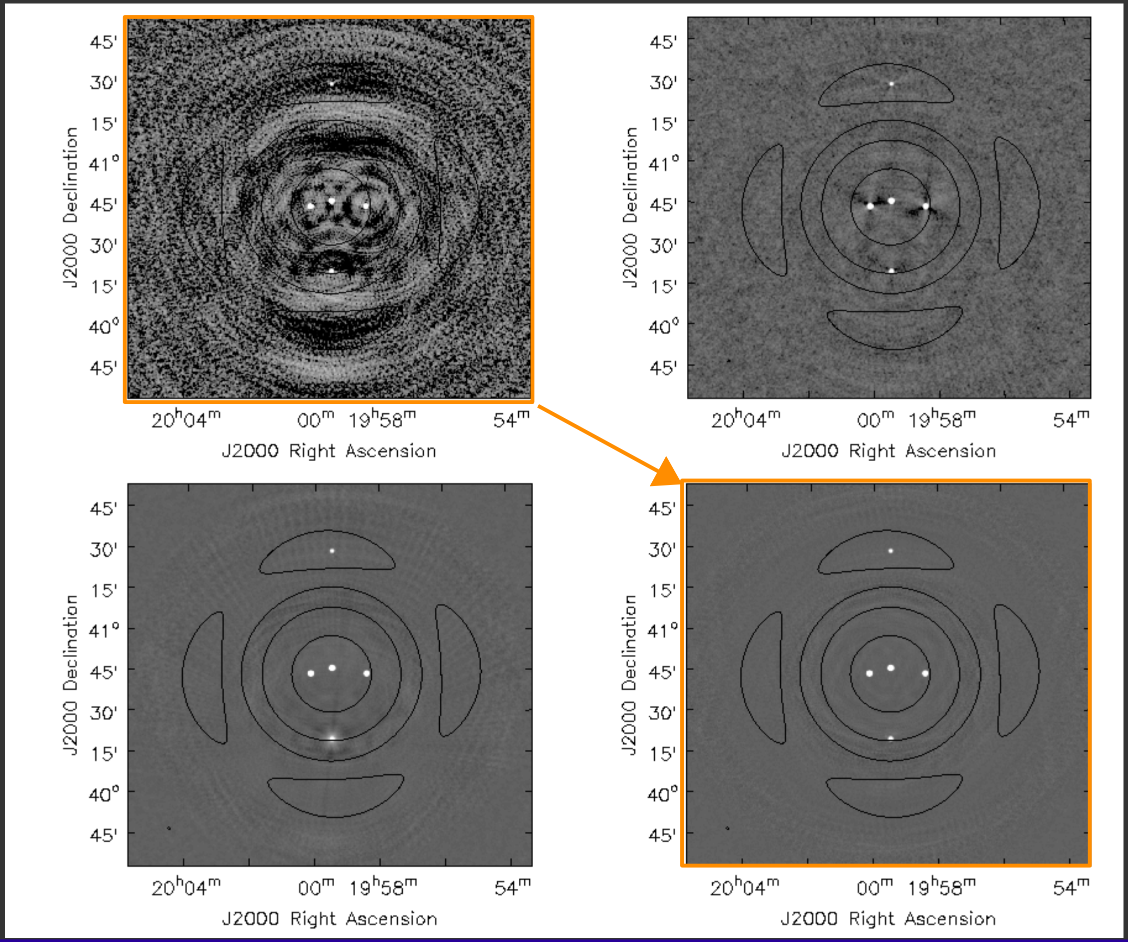
- Simultaneously account for the PB effects and frequency dependence of the sky
  - PB effects corrected by WB A-Projection
  - PB-corrected image used in MT-MFS for model the frequency dependence of the sky brightness

MFS+SI

MT-MFS+SI

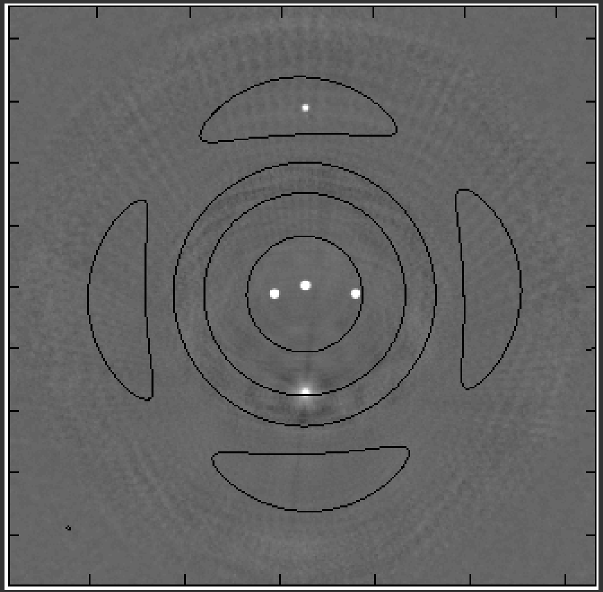
MT-MFS+
A-Projection

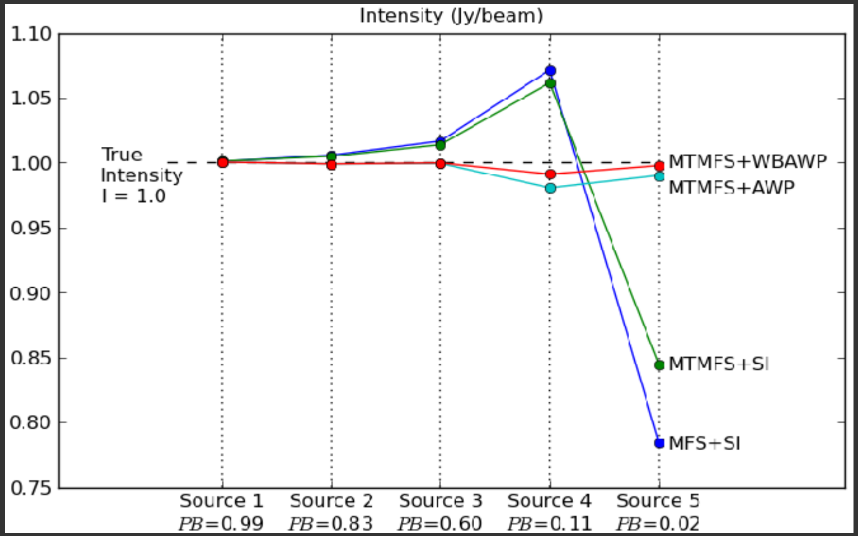
MT-MFS+
WB A-Projection

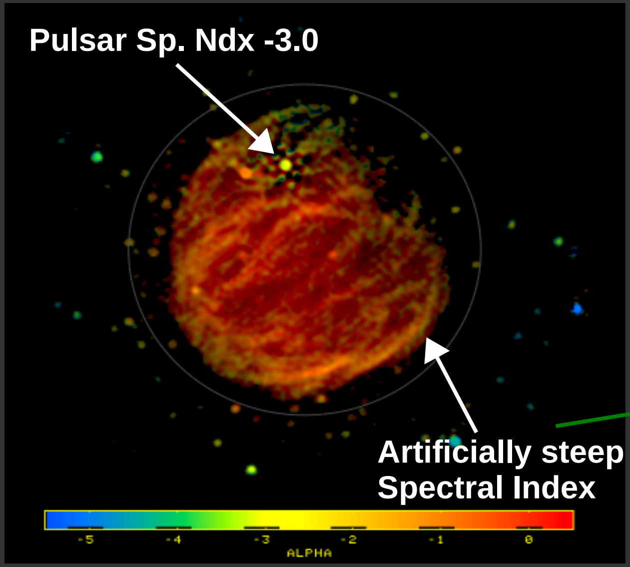


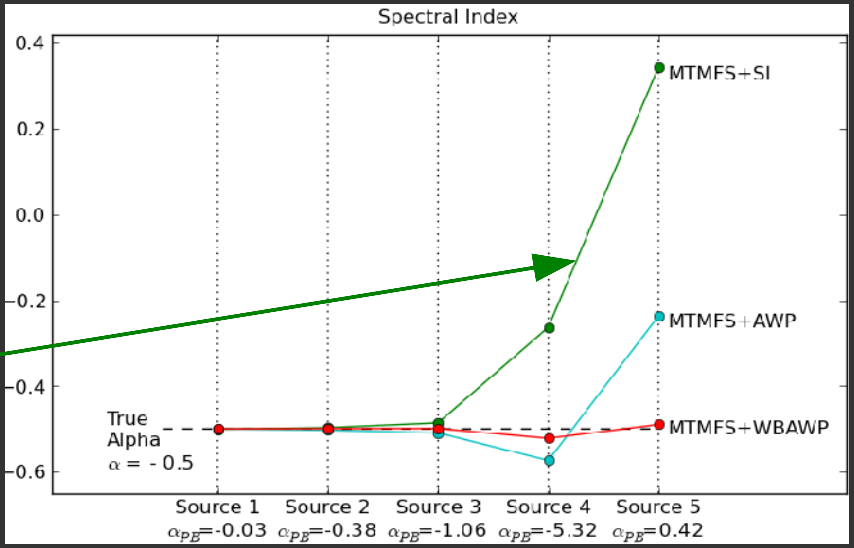
Ap.J., 2013

Instrumental frequency dependence









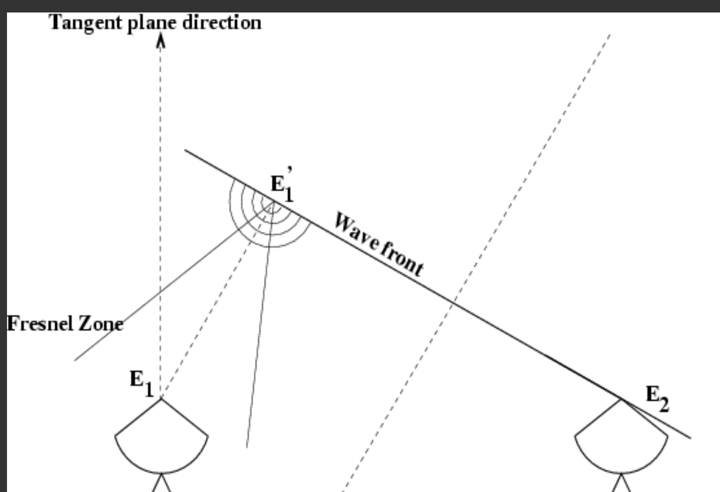
Non co-planar baselines: W-Term

- Imaging

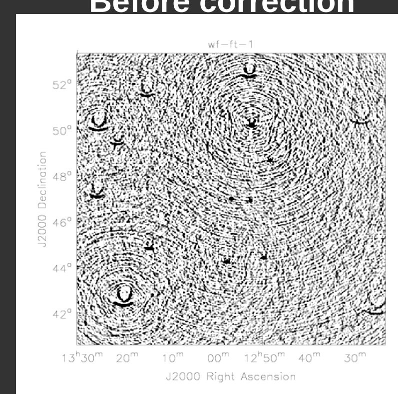
$$V_{ij}^{DI-Cal}(\nu) = W_{ij} \int P_{ij}(\mathbf{s},\nu,t)\; I(\mathbf{s},\nu)\; e^{\iota\, \mathbf{s}.\mathbf{b}_{ij}}\; d\mathbf{s}$$

Data | DI Calibration | DD Term Instrumental Ionospheric | Sky | **Geometry**

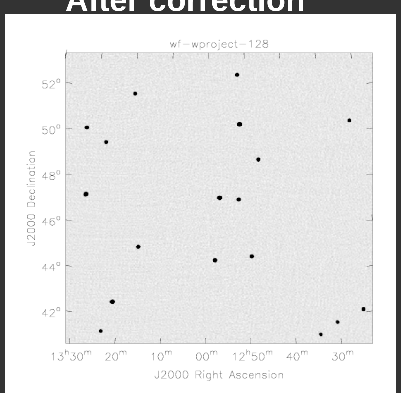
- The geometric term (non co-planar baselines)
  - Transform is no more 2D Fourier Transform



Before correction



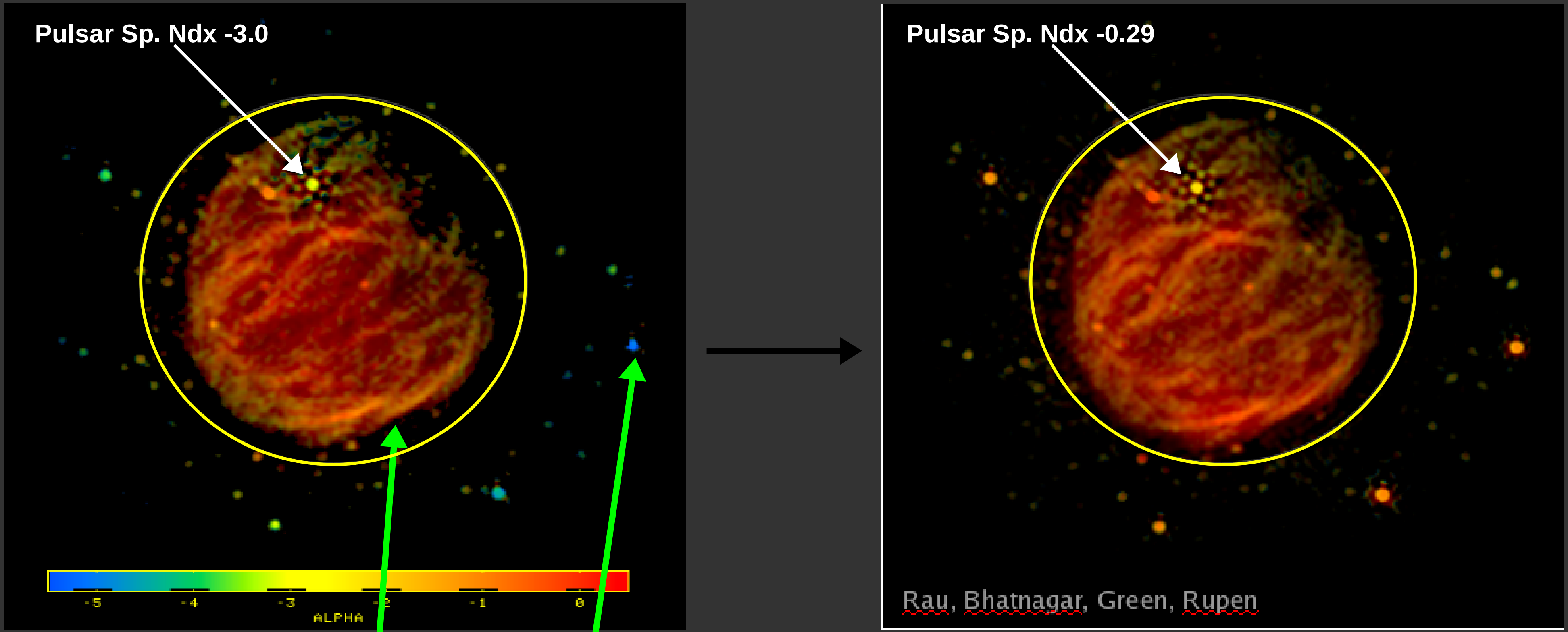
After correction



NRAO

Wide-Band AW-Projection + MT-MFS

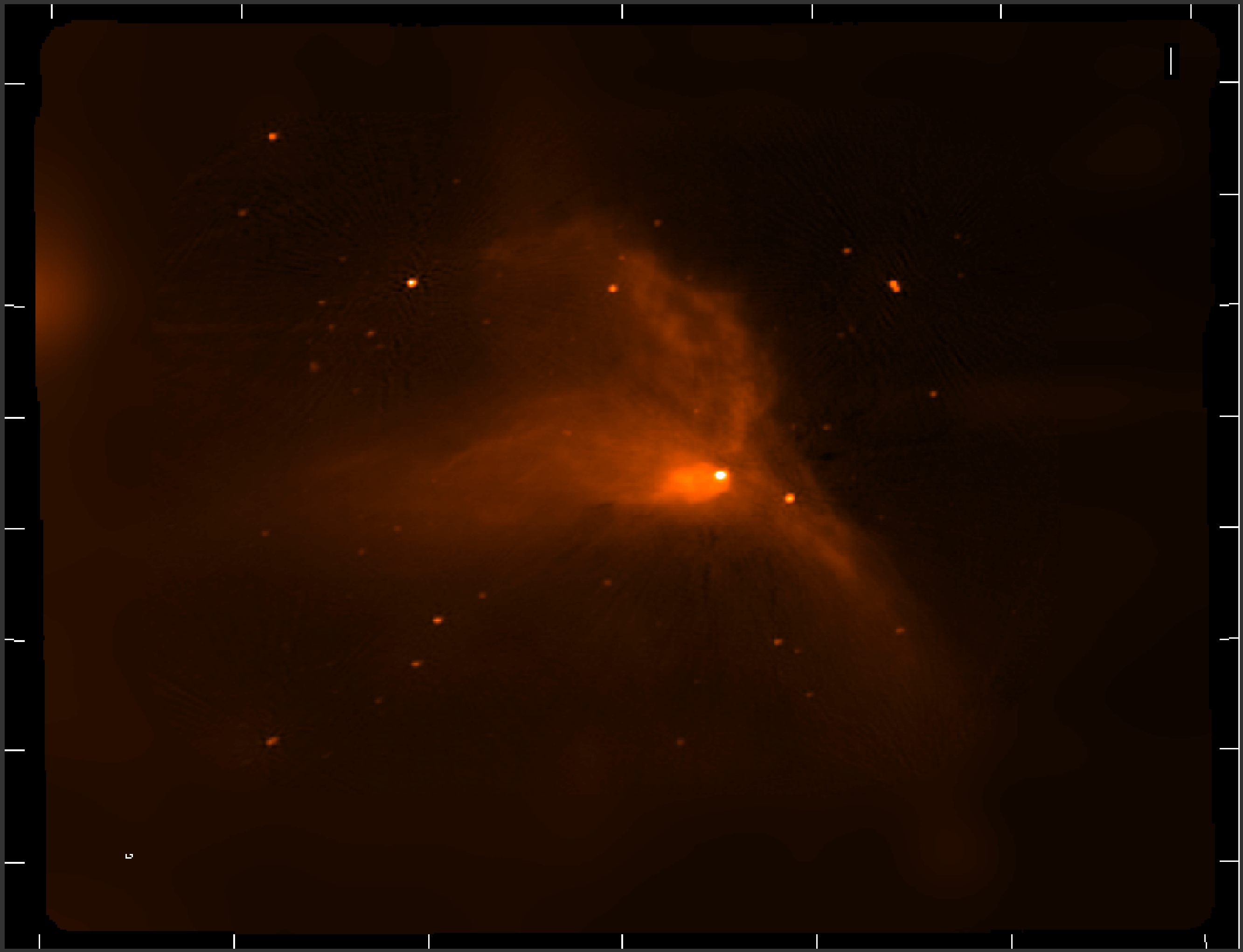
- Intensity weight Spectral Index Map
- Wide-field Spectral Index maps comes out in the wash correctly



Artificially steep (due to PB)

A&A, 2008, ApJ, 2013

Wide-band Mosaic Imaging + SD



- Simultaneous corrections for instrumental effects+ Frequency Dependence of the Sky
- WB AW-Projection + MS-MFS + Mosaic
- Wide-band ~200-pointing mosaic
- EVLA + GBT Feathering (existing algorithm)
- In progress:
  - Mosaic spectral Index mapping
- Parallel execution / Optimization /
- Numerical tests

Status: In production or commission stage

- W-Term correction: Dominant DD term at low frequencies
  - Facted-imaging, W-Projection, W-Stacking
- Extended emission
  - MS-Clean, Asp-Clean, various variants
- Frequency dependence of the sky brightness
  - MT-MFS
- PB corrections
  - A-Projection: Time and polarization dependence
  - WB A-Projection: Also frequency dependence

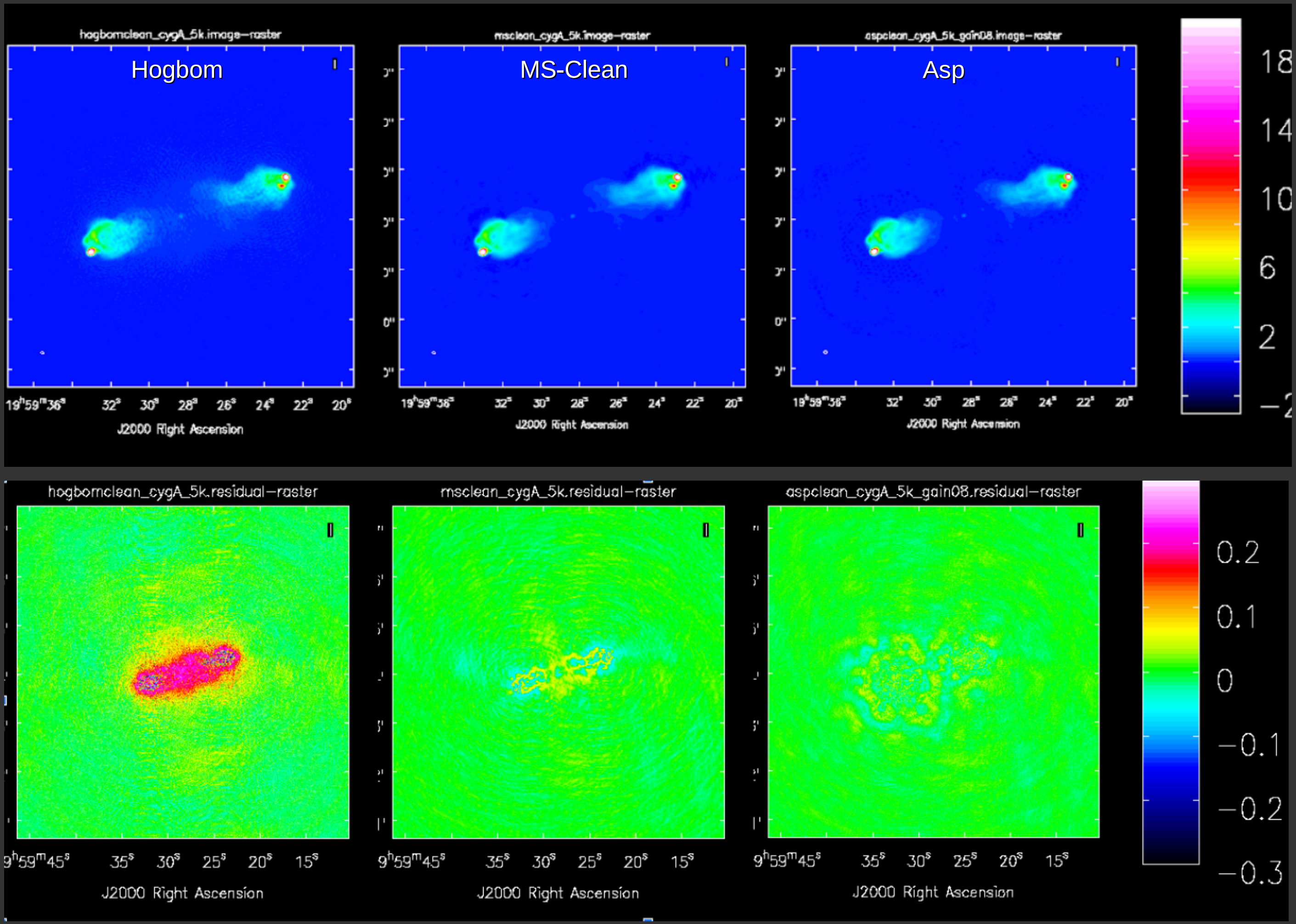
- Recently Commissioned:
  - `W-Term + WB A-Projection + MT-MFS`
    - » Simultaneously account for instrumental and sky terms
  - `Wide-band Mosaic`
    - » All of the above for mosaic imaging (e.g. VLASS)

Status: In active R&D

- Improved scale-sensitive deconvolution algorithms
  - Asp-Clean, WAsp (Wide-band Asp-Clean)
- Full-polarization imaging
  - Extend PB correction to full polarization
  - RM Synthesis at the sensitivity and band-width now available
- Parallelization
  - Many projects takes weeks of computing for imaging
  - Cluster computing: High Performance Computing (HPC), High Throughput Computing (HTC)
  - CPUs, GP-GPUs, FPGAs,...

Asp imaging algorithms

- Adaptive Scale Pixel (Asp): Scale-sensitive image reconstruction of complex emission
  - Asp-Clean: Narrow-band implementation, now in production-CASA
  - WAsp: Wide-band Asp

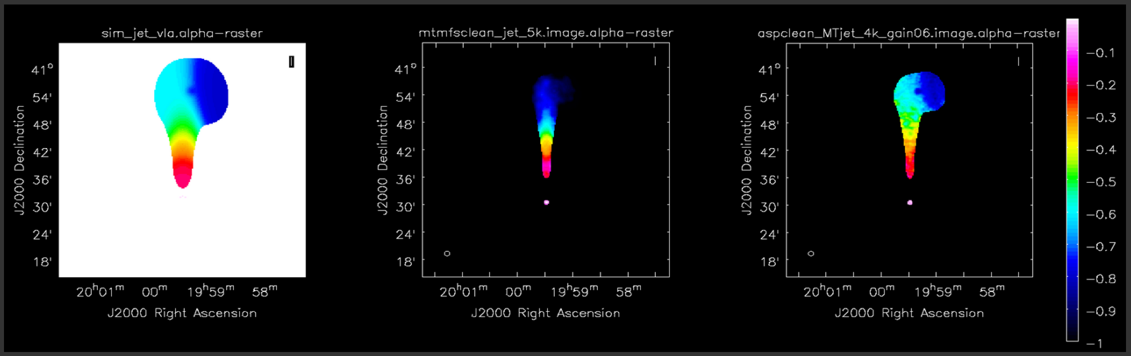


WAsp imaging algorithms

- Adaptive Scale Pixel (Asp): Scale-sensitive image reconstruction of complex emission
  - WAsp: Wide-band Asp

MS-MT-MFS WAsp

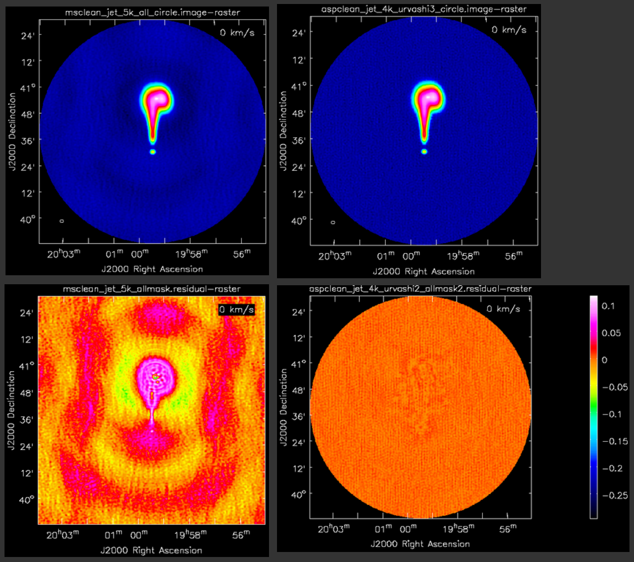
Spectral Index Mapping



MS-MT-MFS WAsp

Stokes-I

Residual



Computing Cost

- Imaging + deconvolution accounts for >90% of the computing cost in a "typical" end-to-end processing

DataArchive → Flagging/Calibration → Imaging-Deconv. → ImageArchive

- Computing Scaling
  - Computing costs: $N^2_{support} \times N_{vis}$: Dominated by Image-making (a.k.a. the "Major cycle")
  - Memory footprint: $N^2_{Scales} + N^2_{Terms}$: Dominated by Deconvolution (a.k.a. "Minor Cycle"

- Large cluster with cheaper CPU/GPU and small memory for the Major cycle

  Faster computers with lots of memory for the Minor Cycle

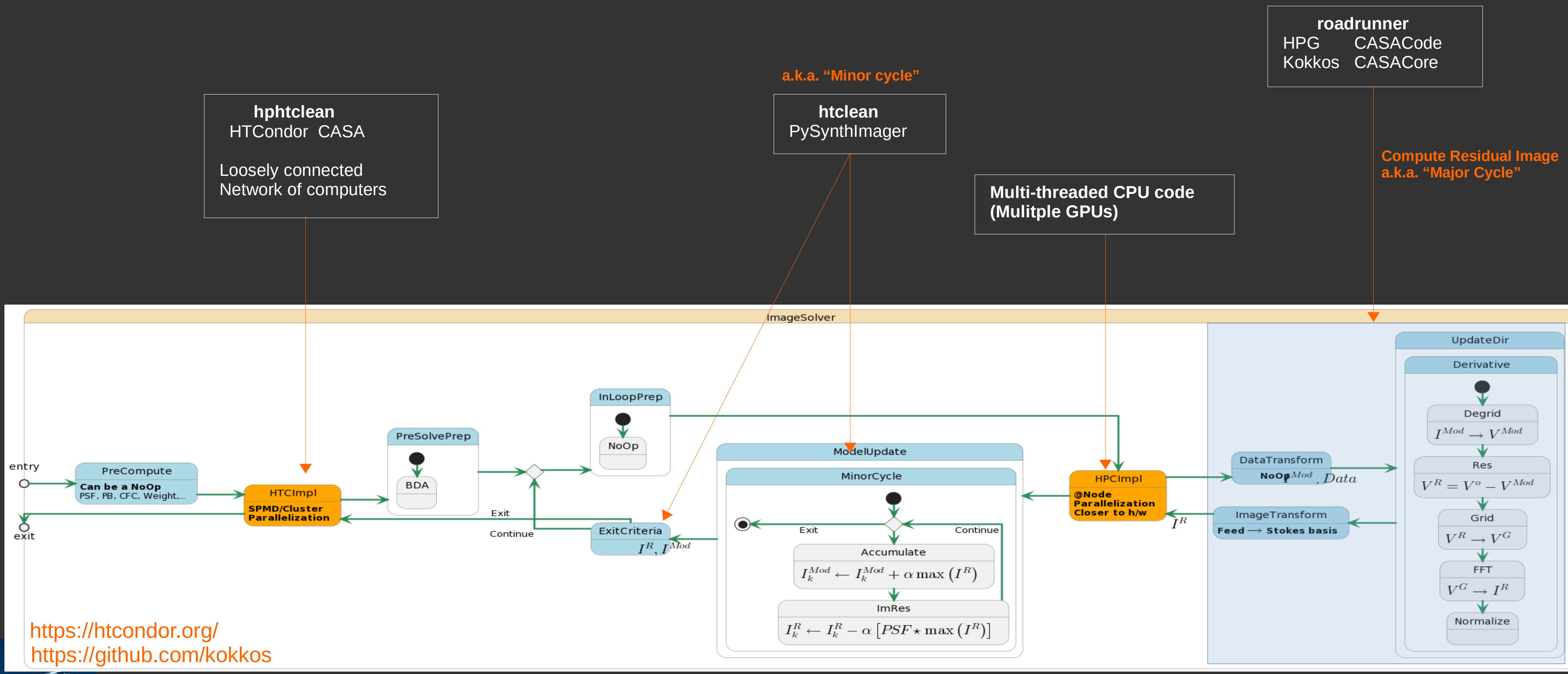
- Imaging : Pleasantly Parallel (a.k.a "Embarrassingly parallel")
  - Scatter-Gather paradigm on the Cluster scale

Scale of computing

- Estimates of the size of computing
- ngVLA
  - 30 – 50 PFLOP/s
  - Currently the largest facility has about 40 PFLOP/s!
  - No. of CPU cores: O(Million)
    - » These have to run 24x7 to keep-up with the data rates!!
- SKA: Latest estimates are similar
- Available computing power:
  - A typical desktop: 0.1 TFLOP/s
  - A typical GPU: 10 TFLOP/s
- Utilization
  - Single-digit percentage is typical
  - 5-7% utilization of the CPU-cluster at NRAO

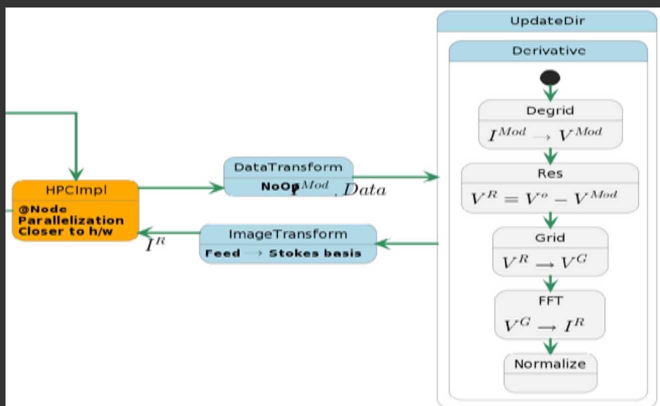
Algorithm Architectural Component

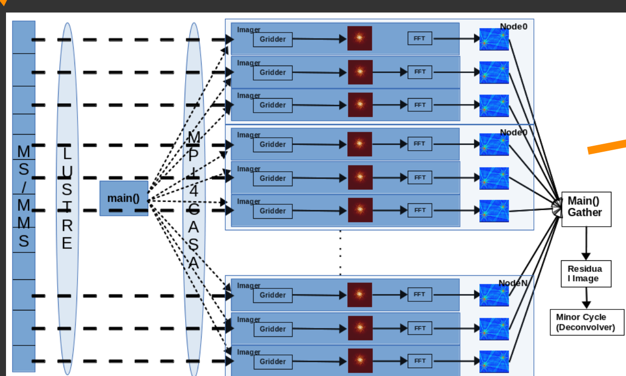
- Both, calibration and imaging as standard χ^2 minimization problem
- Specialization of the framework components deliver various calibration and imaging algorithms
  - Prototypes: CPU single/multi-cores, cluster, variety of GPUs,..., external GPU cluster

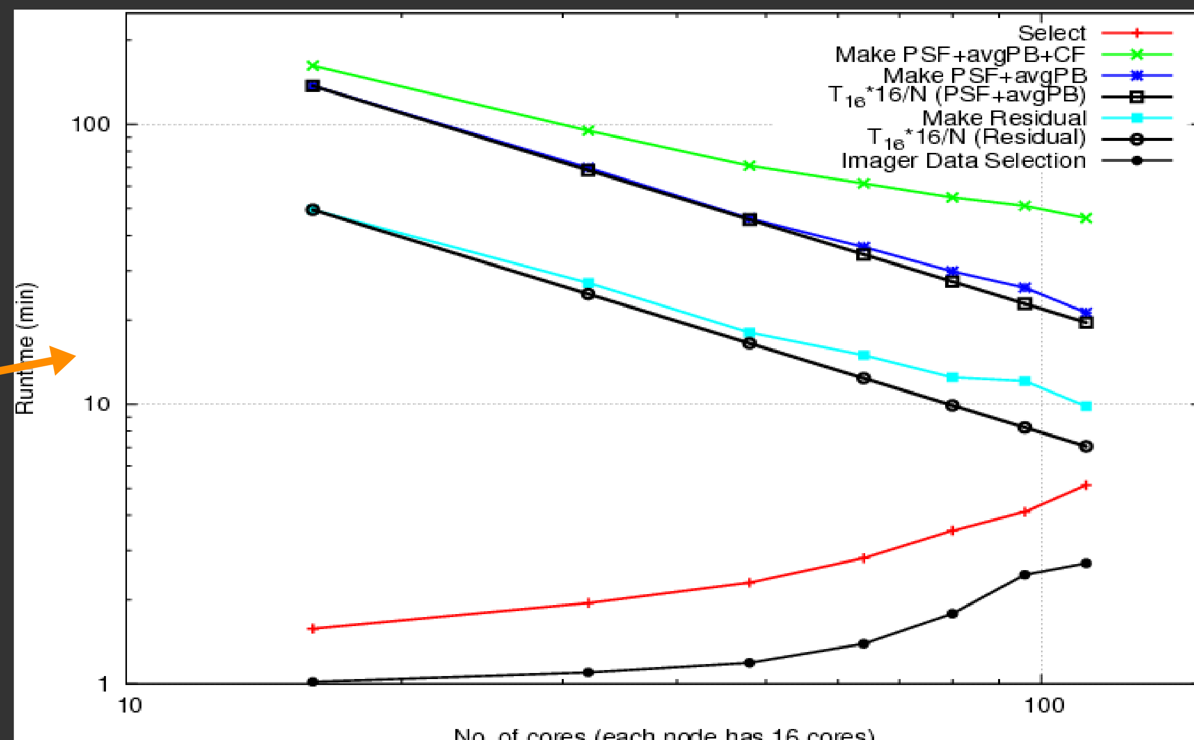


Parallel processing

- High Computational Intensity (FLOP per byte)
  - O(10^{2-3}) FLOP per data point. Number of data points: O(10^{12-15})



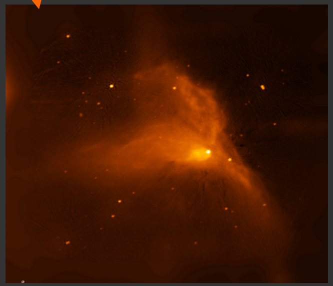


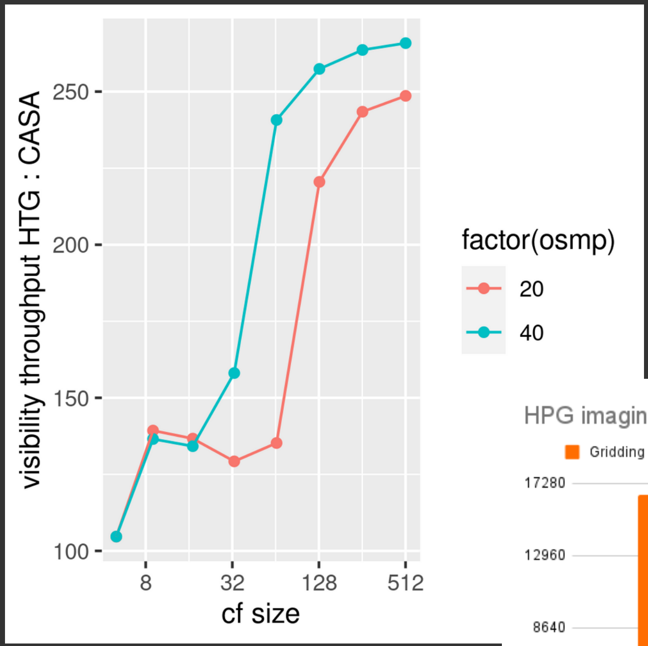


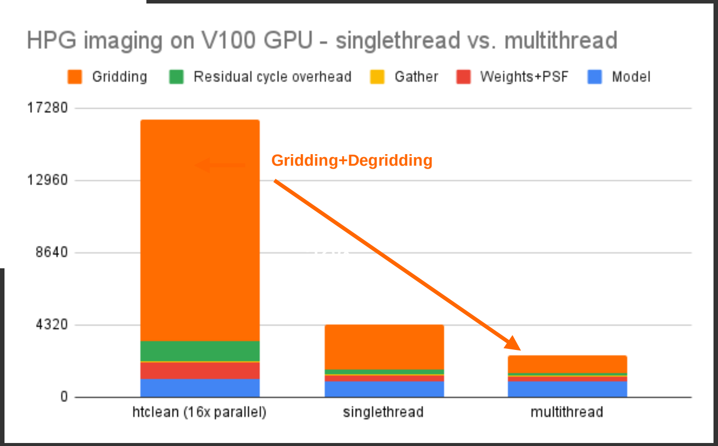
ngVLA will need O(Million)-way parallization!

High Performance Gridder

- A gridder on a GPU connected to a CPU host (NGVLA Memo #05, #07)
  - Measured speed-up: 100 – 200x compared to a single CPU core
  - 200-pointing wide-band mosaic: 7-10 days (parallel tclean) vs 2.5hr (HPG)



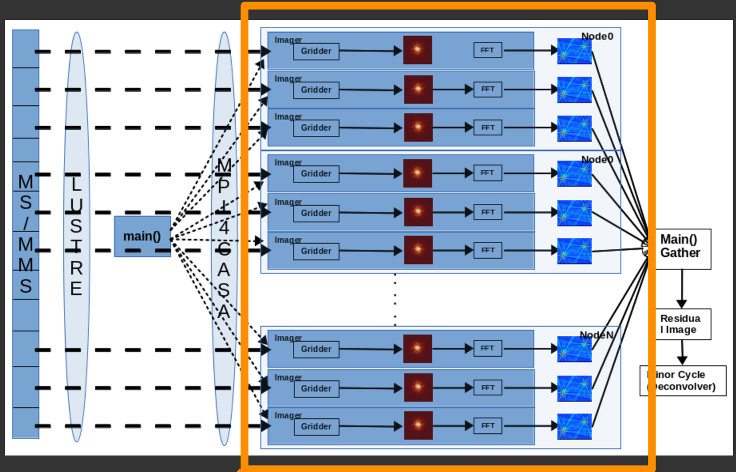




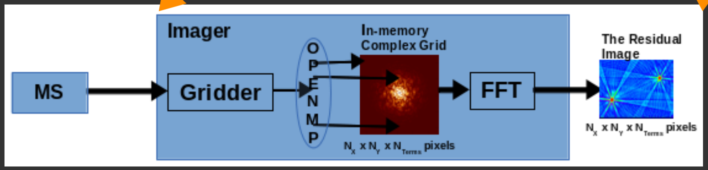
Multi-core CPU

Single CPU-core +GPU

Multi CPU-cores +GPU



runtime complexity reduction

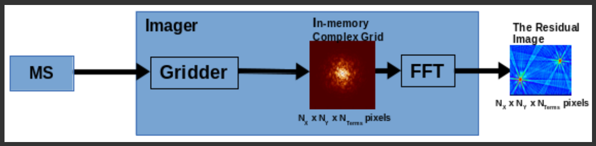


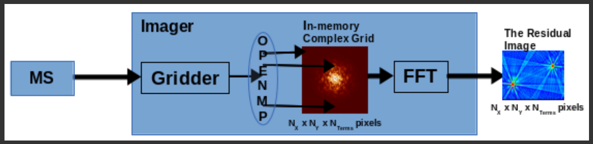
Scientific S/W Complexity

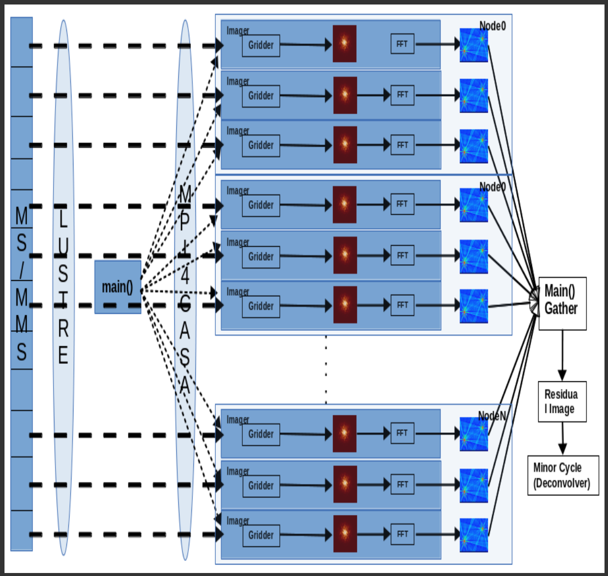
- Imaging is embarrassingly parallel
  - SAMD parallelization architecture measures high efficiency
  - In-coherent gather is OK

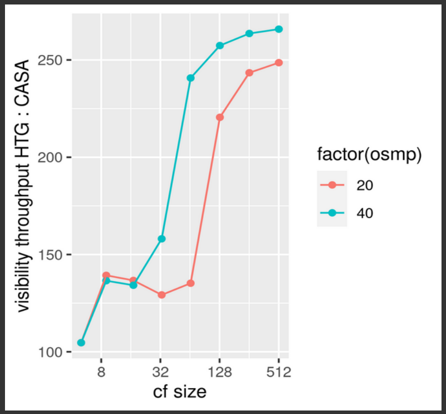
$$I^D=\sum_p FXG_pV_p \quad \text{or} \quad F\sum_p XG_pV_p \quad \text{or} \quad FX\sum_p G_pV_p$$

Node s/w complexity

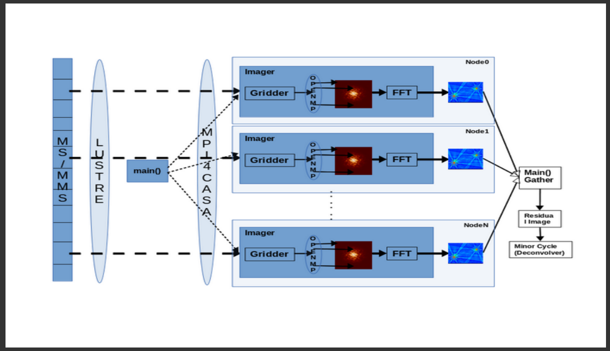








100-200x fewer processes



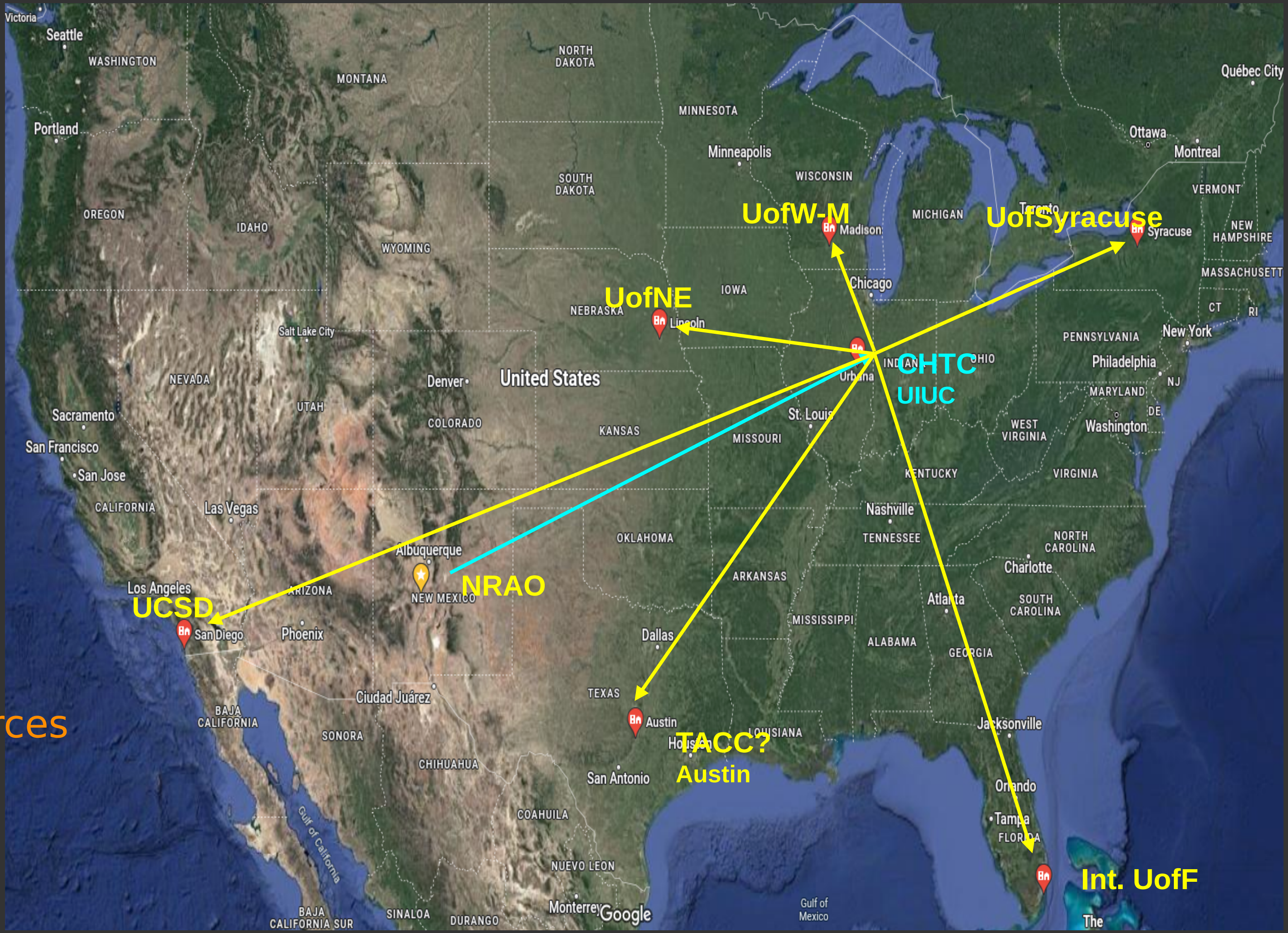
Complexity reduction

O(Million) **$O(10^3)$**

High Throughput Computing Prototype

- <u>Computing at a national scale</u>
- Breakup into smaller problems
- Trigger from NRAO, deploy nationally
- Execute as-n-when resources become available.

Using the HTCondor scheduler software
https://htcondor.org/



- <u>Work in progress</u>
- Get more resources
- Improve reliability, usability of h/w and s/w
- Use international resources

Challenges

- Algorithms
  - Wide-band RM Synthesis
  - DD Solvers: Ionospheric screen
  - Efficient multi-scale algorithms for both imaging & deconvolution

- Computing
  - Optimal use of available computing resources
  - Use of (massively) parallel hardware
    - Multi-core CPUs, GP-GPUs
  - Memory footprint
  - Data I/O: SKA-, ngVLA-class problem
    - Algorithms are fundamentally iterative

Computing Cost

Re-sampling: Uses all data

The Major Cycle

Gridding + Degridding + FFT

Only images

The Minor Cycle

Deconvolution (Hogbom,, MS-Clean, MT-MFS, Asp-Clean,...)

Model the measurements (telescope, ionosphere, etc.)

Projection algorithms

Images corrected for instrumental artifacts

Model the sky brightness distribution

Image modeling (a.k.a. “deconvolution”)

Computing Cost

Standard Imaging

Standard deconvolution

Computing Cost

Re-sampling: Uses all data

The Major Cycle

Gridding + Degridding + FFT

Only images

The Minor Cycle

Deconvolution (Hogbom,, MS-Clean, MT-MFS, Asp-Clean,...)

Model the measurements (telescope, ionosphere, etc.)

Projection algorithms

Images corrected for instrumental artifacts

Model the sky brightness distribution

Image modeling (a.k.a. "deconvolution")

Computing Cost

Advanced deconvolution: MS, MT-MFS, MS-MFS

Standard Imaging

Computing Cost

Re-sampling: Uses all data

The Major Cycle

Gridding + Degridding + FFT

Only images

The Minor Cycle

Deconvolution
(Hogbom,, MS-Clean, MT-MFS,
Asp-Clean,...)

Model the measurements
(telescope, ionosphere, etc.)

Projection algorithms

Images corrected for instrumental artifacts

Model the sky brightness distribution

Image modeling (a.k.a. “deconvolution”)

Computing Cost

Advanced Imaging: W-, A-, AW-Proj., Heterogeneous, WB

Standard deconvolution

Computing Cost

Re-sampling: Uses all data

The Major Cycle

Gridding + Degridding + FFT

Only images

The Minor Cycle

Deconvolution
(Hogbom,, MS-Clean, MT-MFS,
Asp-Clean,...)

Model the measurements
(telescope, ionosphere, etc.)

Projection algorithms

Images corrected for instrumental artifacts

Model the sky brightness distribution

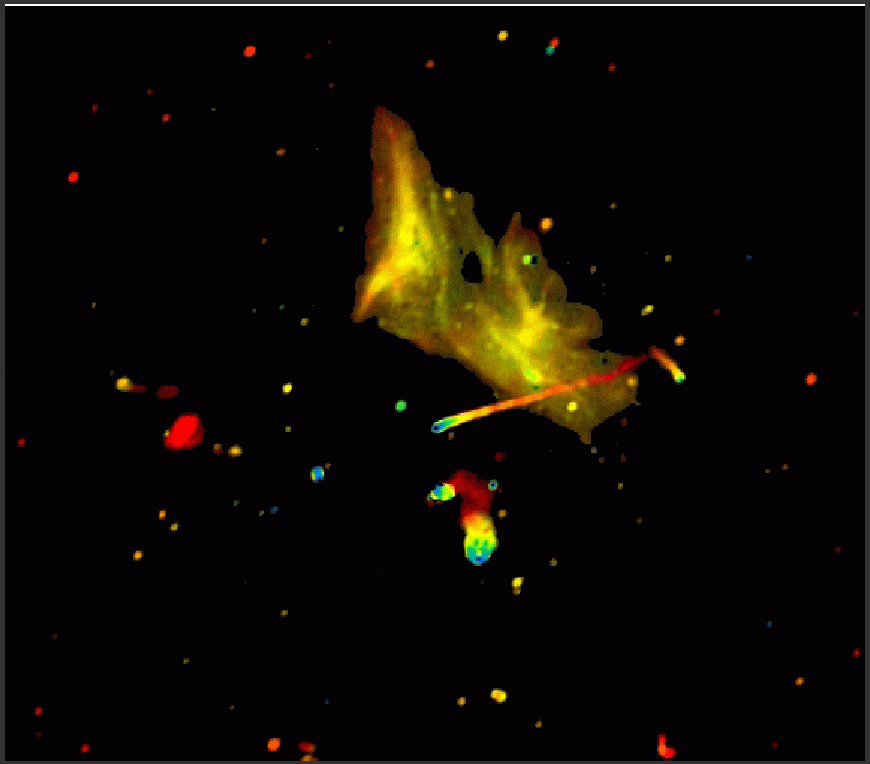
Image modeling (a.k.a. "deconvolution")

Computing Cost

Advanced Imaging: W-, A-, AW-Proj., Heterogeneous, WB

Advanced deconvolution: MS-, MT-MFS, MS-MFS

Imaging with the EVLA @ L-Band



Single pointing, wide-band image

Wide-band ~200 pointing mosaic+Single Dish