Advanced Algorithms

Aug. 30th 2019

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

- Non algorithmic 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

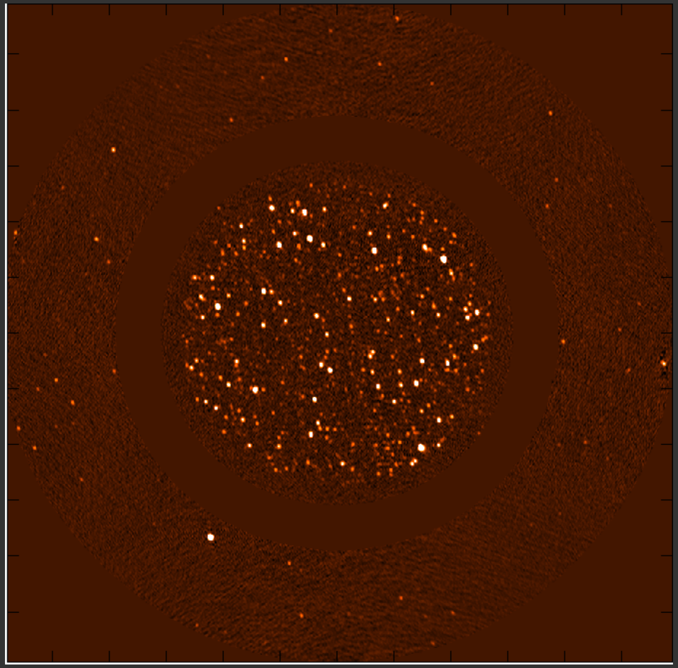
$\Delta\nu$: *Wide band receivers: >60% fractional bandwidth*

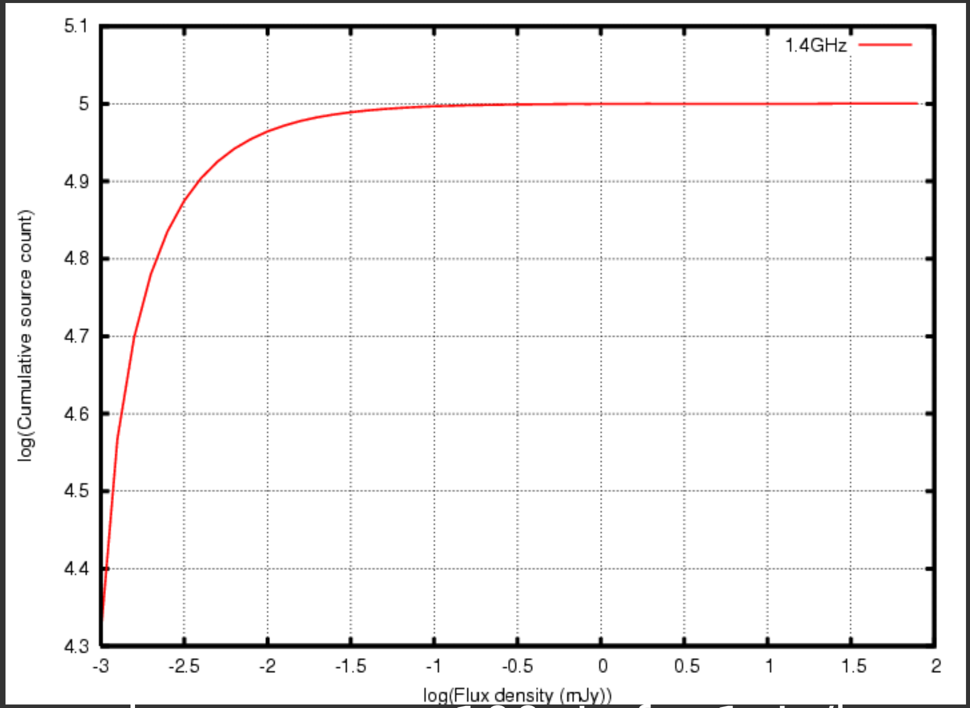
A_{eff}: *More antennas: 30 -- many 100s*

Long baselines: To beat confusion limit

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

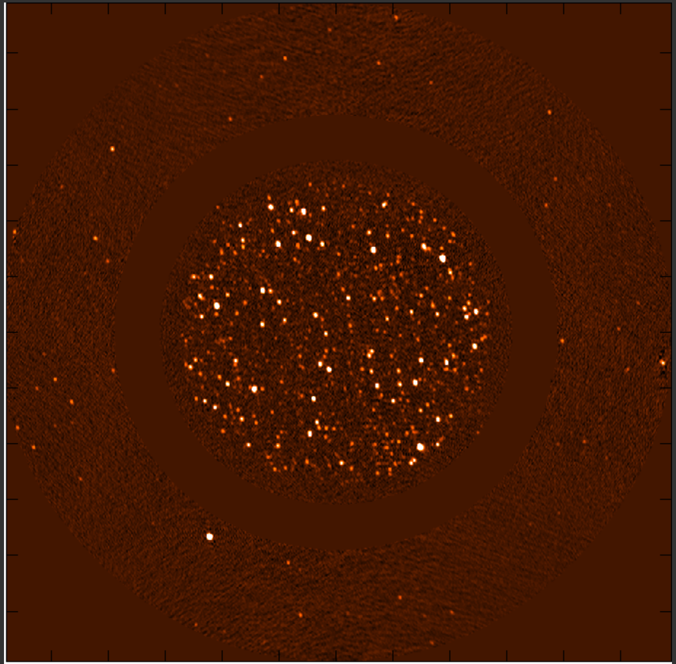
Sky at low frequencies: No. of sources

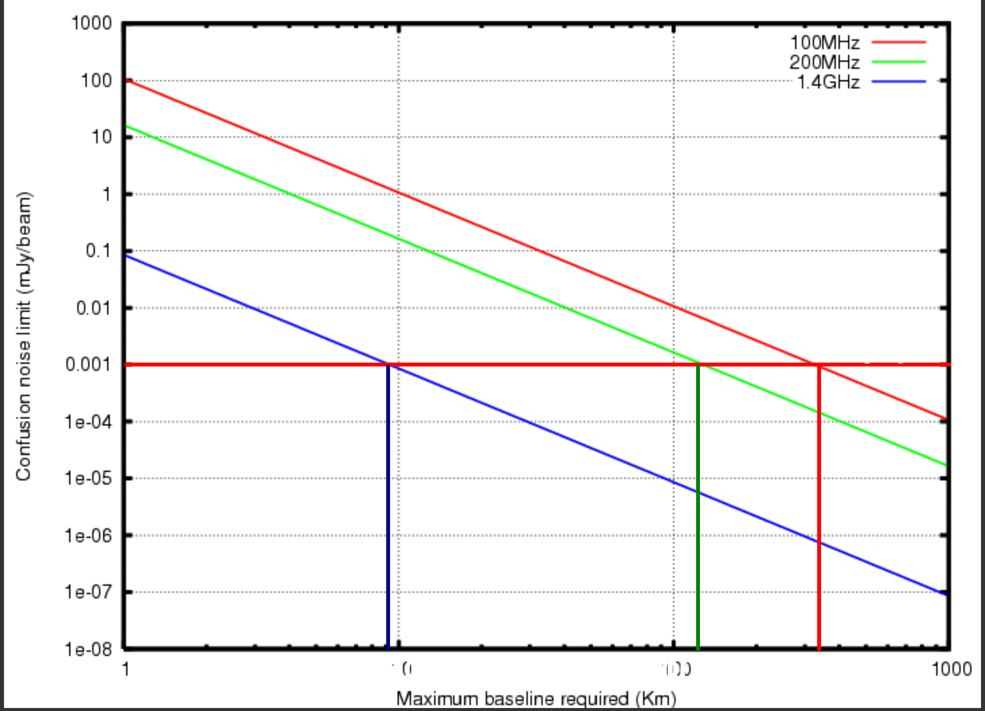




- PSF side-lobe at 1% level → deconvolve sources >100μJy for 1μJy/beam RMS
- $10^{4\text{-}5}$ sources per deg^2 >10μJy @1.4GHz
  - Source size distribution important at resolution < ~2"
- Implications for imaging
  1. Wide-field imaging
  2. HDR imaging: few X 100 mJy – 1 Jy source ~few sq. deg.
  3.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

Sky at low frequencies: Confusion limit

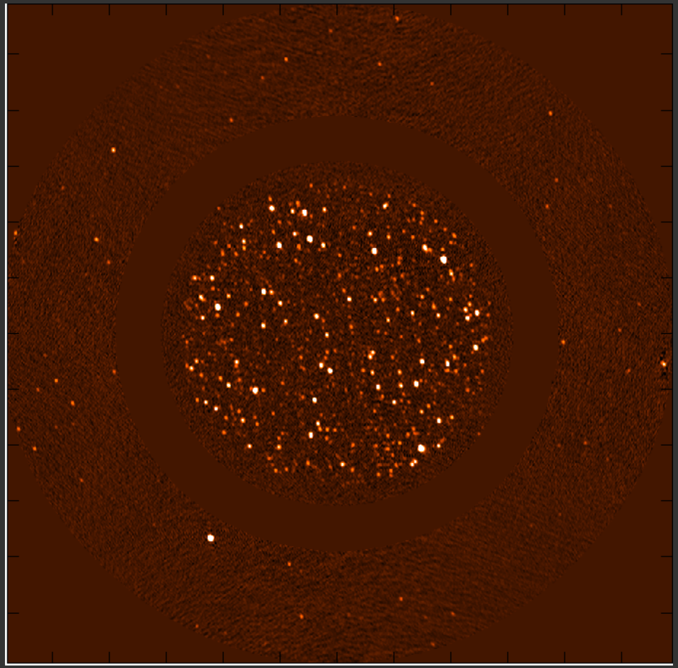


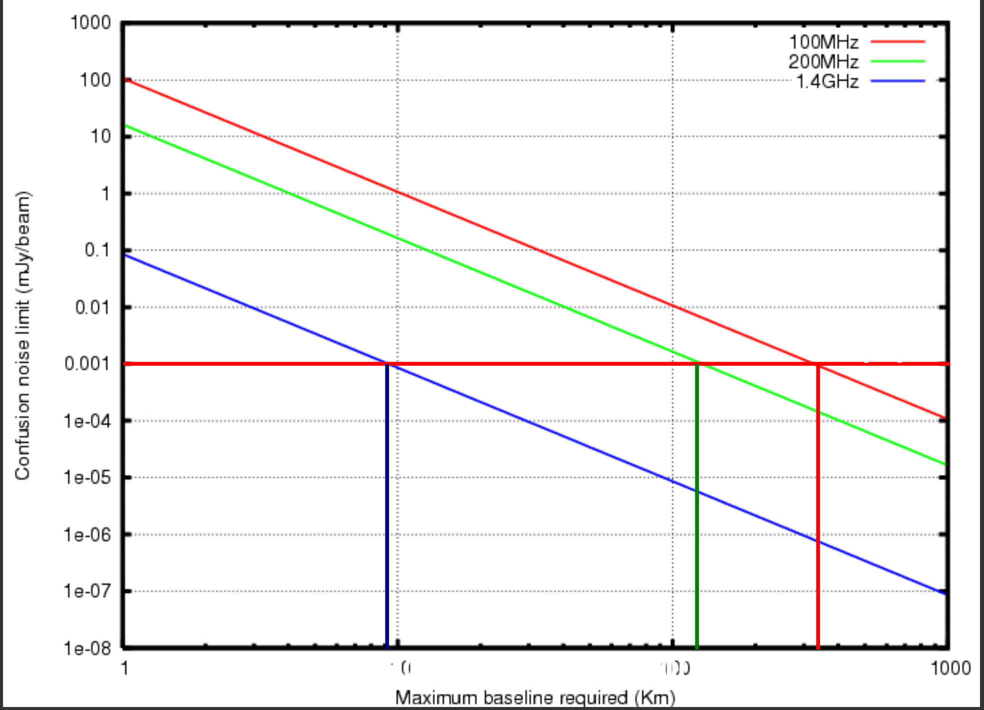


- $\sigma_{\text{confusion}} \propto (\nu^{-2.7}/B^2_{\text{max}})$: B_{max} ~100 Km at 200MHz for $\sigma_{\text{confusion}}$ ~ 1µJy/beam
- Implications for imaging
  1. Long baselines: $B_{\text{max}} > 2\text{-}3$ Km & DR > 10^4
  2. Wide-field effects: W-term, PB effects, ionospheric effects
  3.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

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 fo
    1. Long
    2. Wide- effects
    3. Large

Point source sensitivity 1-sigma 12hr. Synthesis:

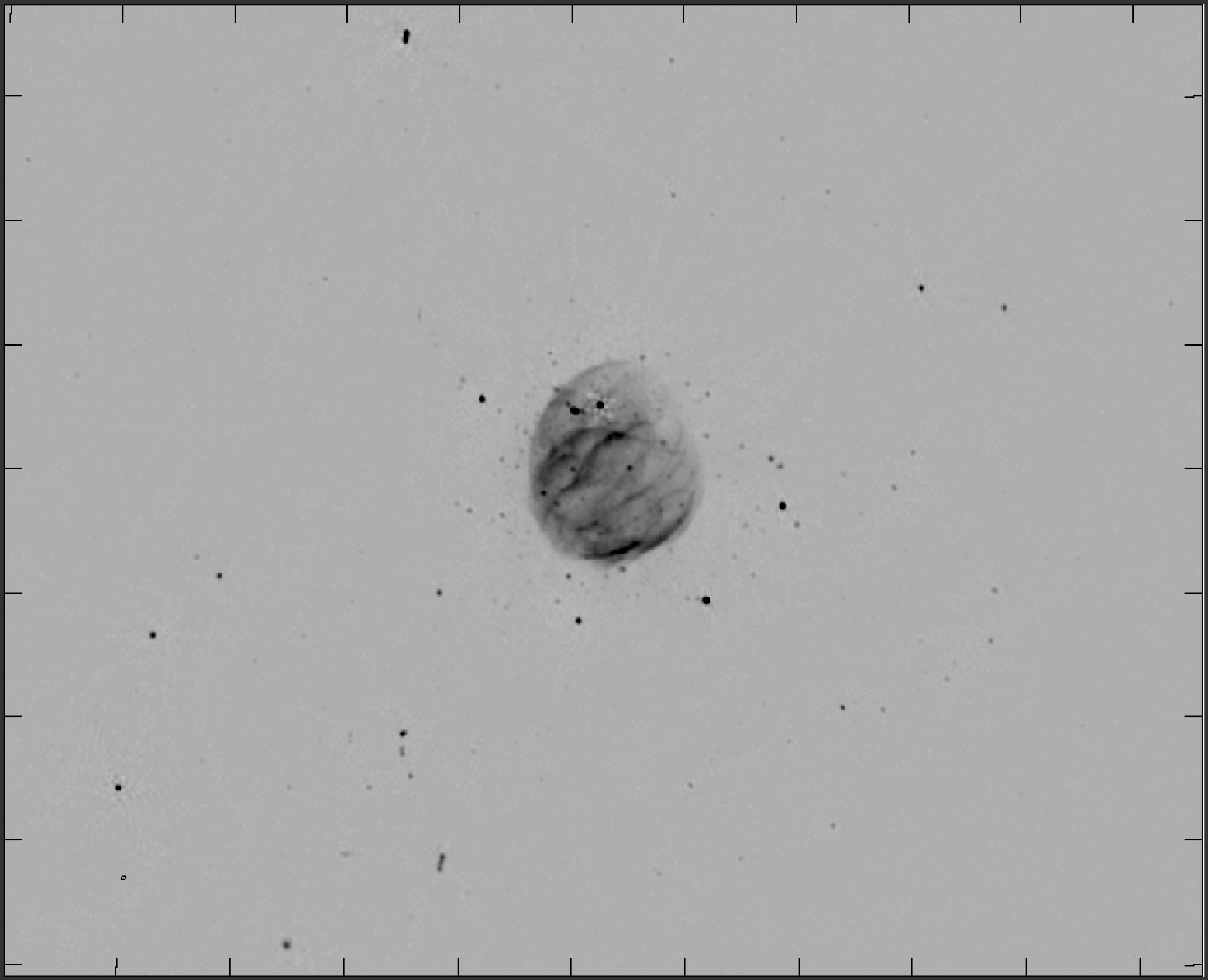
VLA	EVLA	Factor
10uJy	1uJy	10

Data volume:

~1GB	100-1000GB	10^{2-4}

Wide-field, w sing large data volumes is a natural consequence of low frequency and high sensitivity

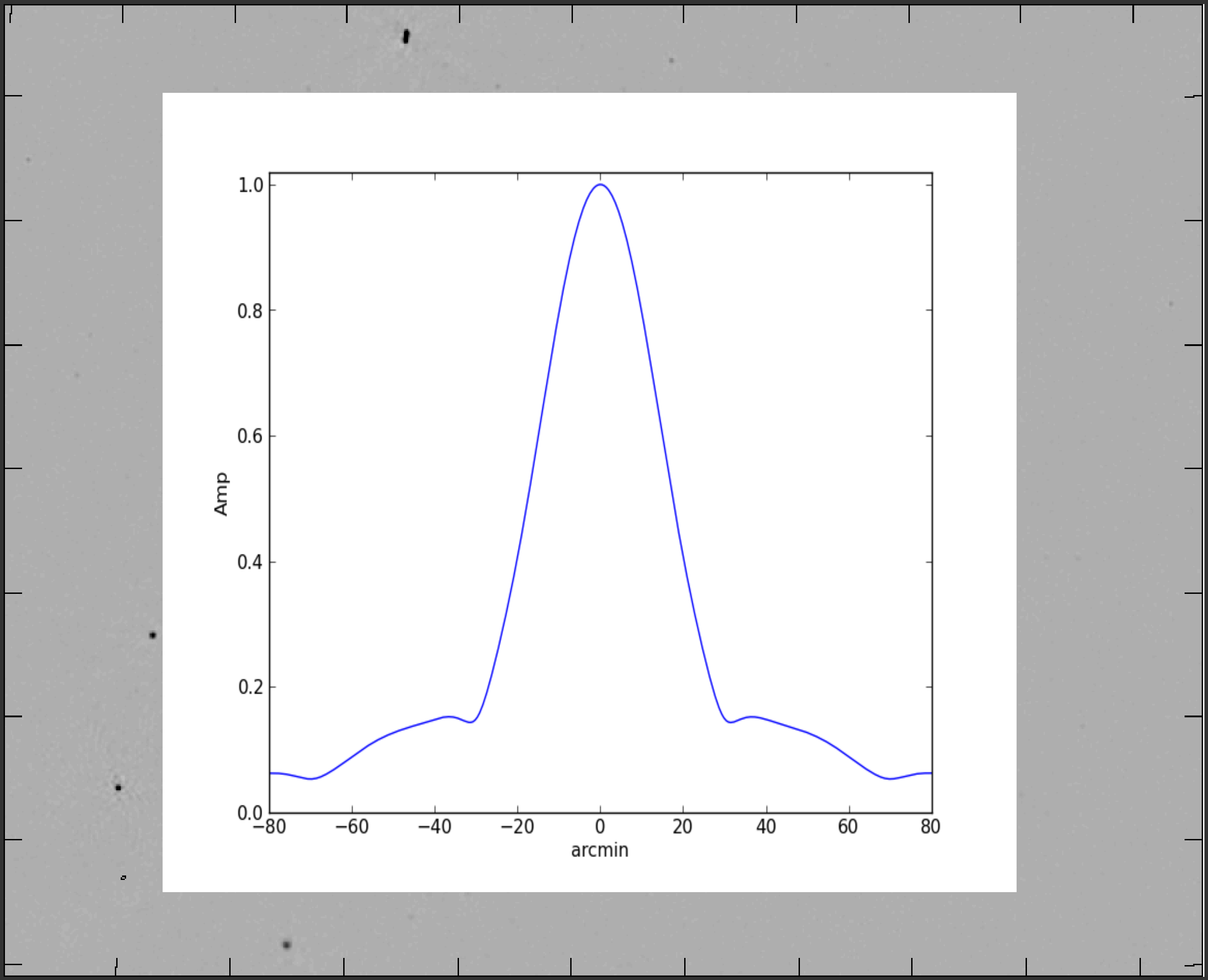
Wide-band implies Wide-field imaging



- 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

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

1. Sky brightness stronger and complex: Multi-Scale deconvolution
2. Ionospheric effects

    Requires DD solvers: An algorithmic & computing challenge in itself

Direction Dependent (DD) Effects

- DI Calibrated ME

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

Data

DI Calibration

DD Term
Instrumental
Ionospheric

Sky

Geometry

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

- Fastest varying term on the RHS determines the averaging scale (time and frequency)
- Imaging equation:

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

Direction Dependent (DD) Effects

- DI Calibrated ME

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

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

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

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

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

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

$$\textit{such that } X_{ij}\; A_{ij} = \mathbf{1}$$

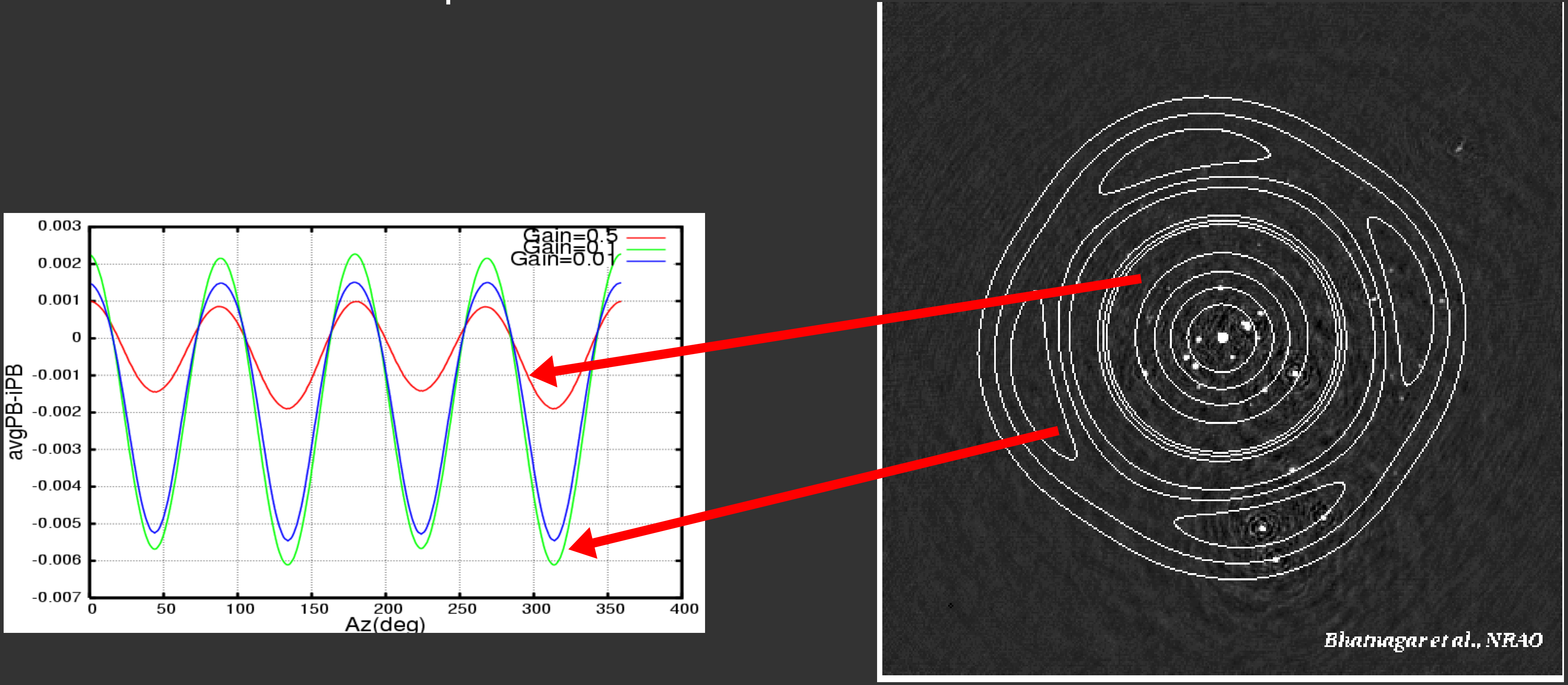
- Then

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

Understand the Physics of the problem; use mathematical techniques to find a solution

Time dependent terms

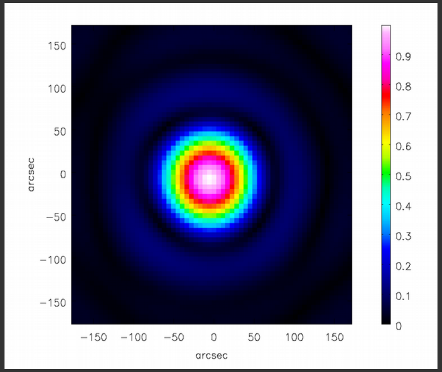
- Antenna PB (*The* $P_{ij}(\mathbf{s},\nu,t)$)
  - Time dependence

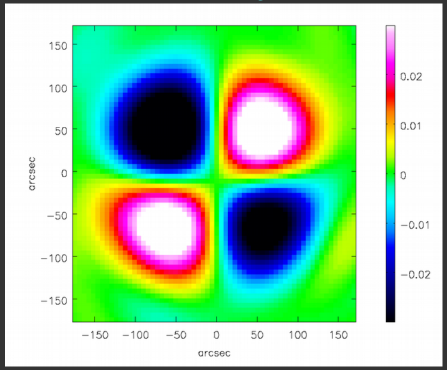


Bhatnagar et al., NRAO

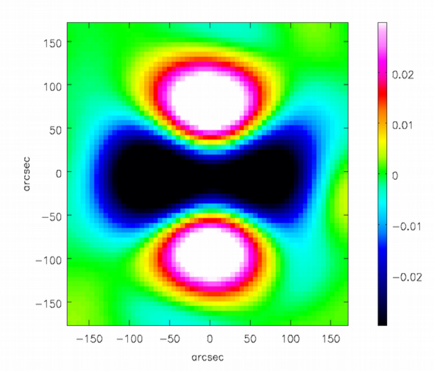
Polarization dependent terms

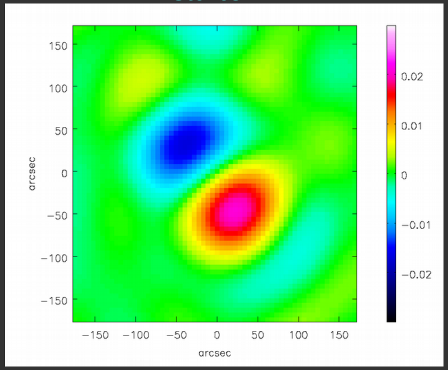
- Antenna PB ($The\ \ P_{ij}(\mathbf{s},\nu,t)$)
  - Polarization dependence


Stokes-I


Stokes-Q

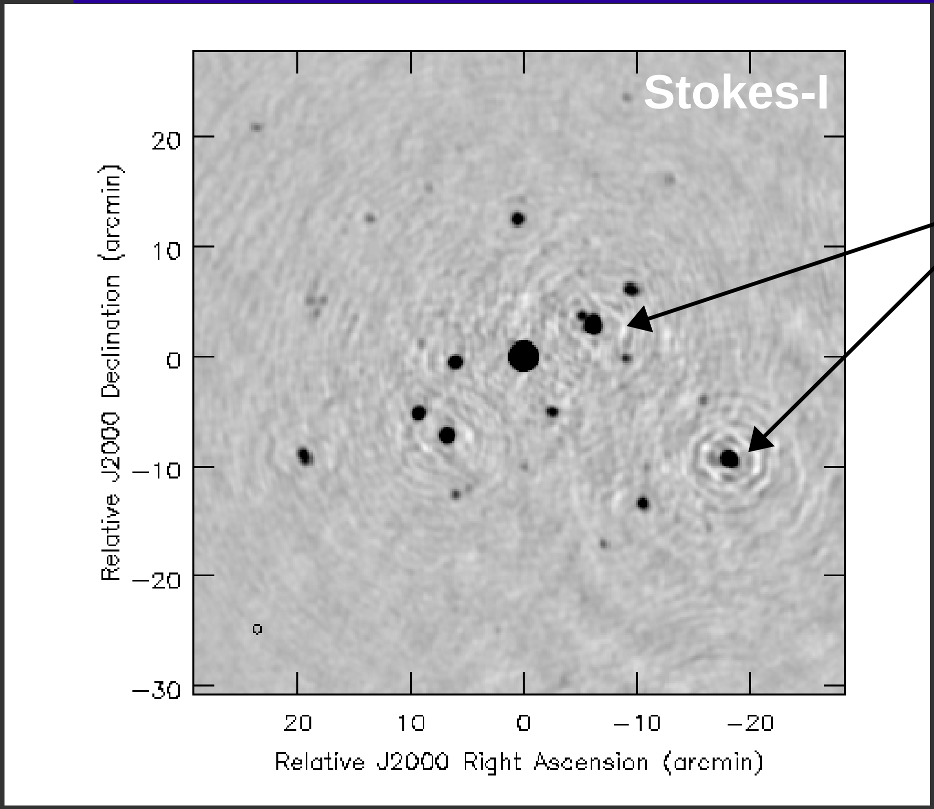
Stokes-Q, -U leakage
~3-5% in the main-lobe
Higher in the first side-lobe


Stokes-U


Stokes-V

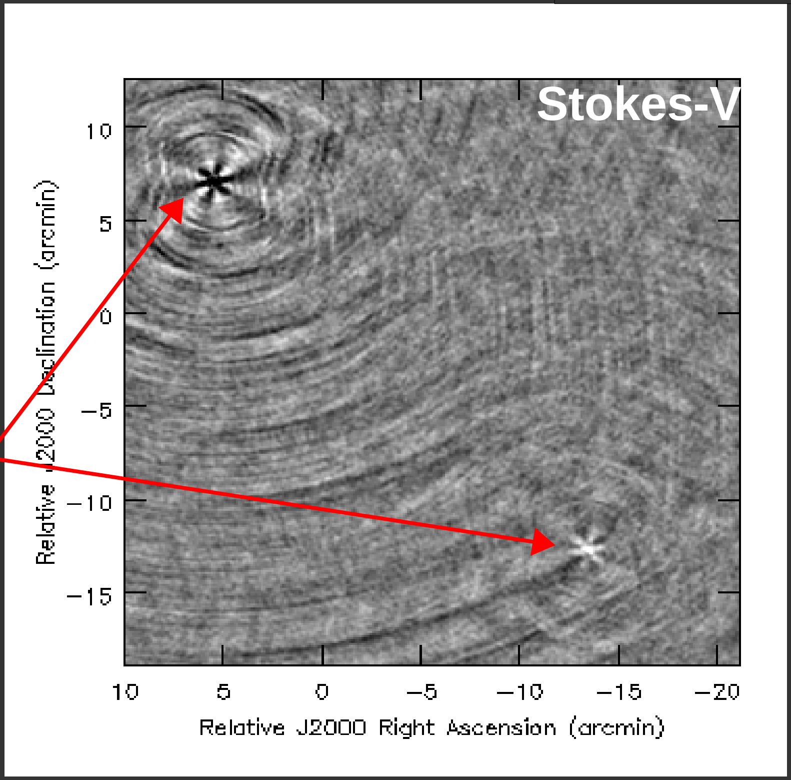
Stokes-V leakage
~3-4% in the main-lobe
Higher in the first side-lobe

Time+Polarization dependence



Errors due to PB
Squint + Rotation + Pointing errors

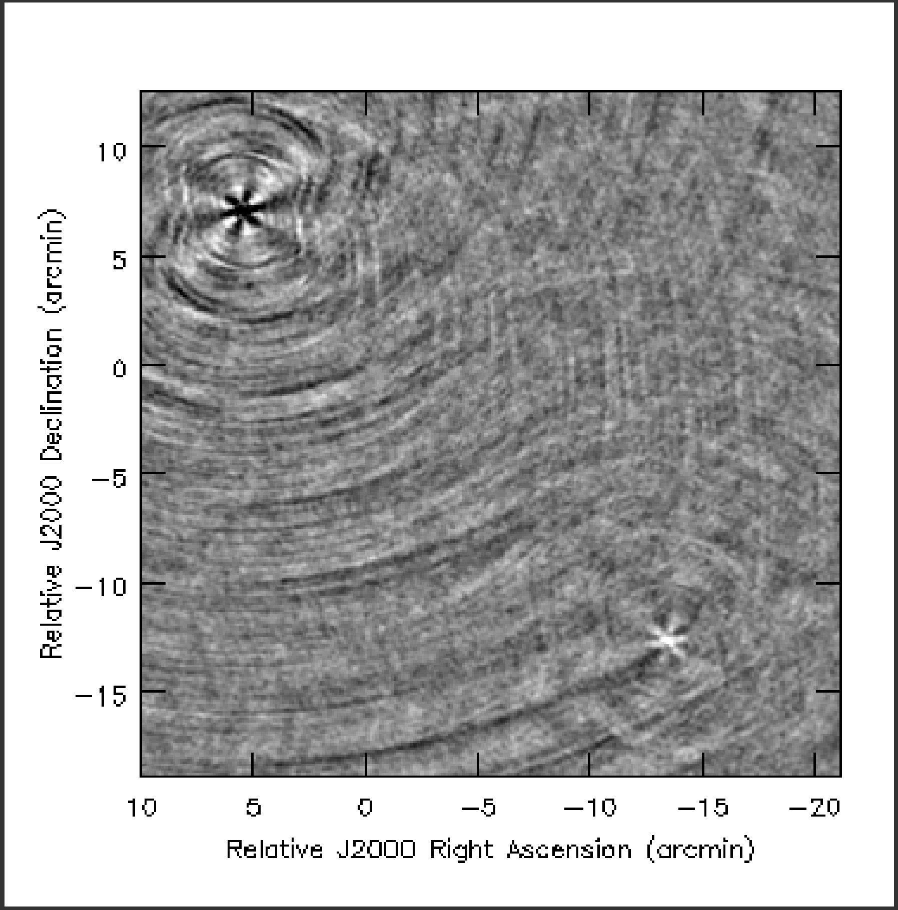
Purely instrumental
Stokes-V artifacts



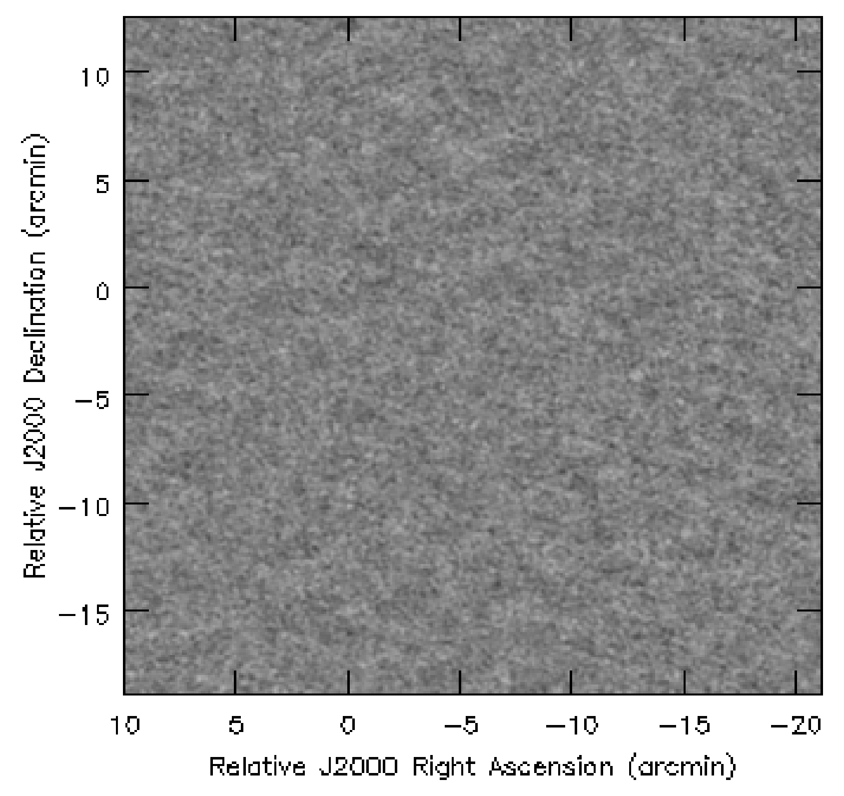
Due to time-variable PB

PB Polarization Effects

Stokes-V Images



A-Projection

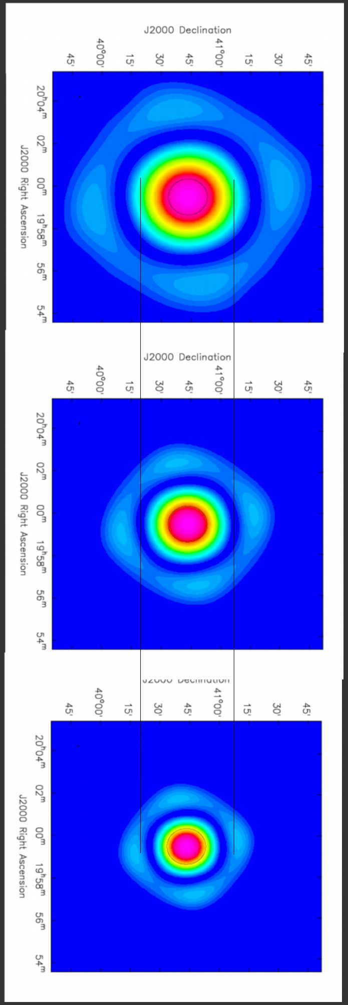


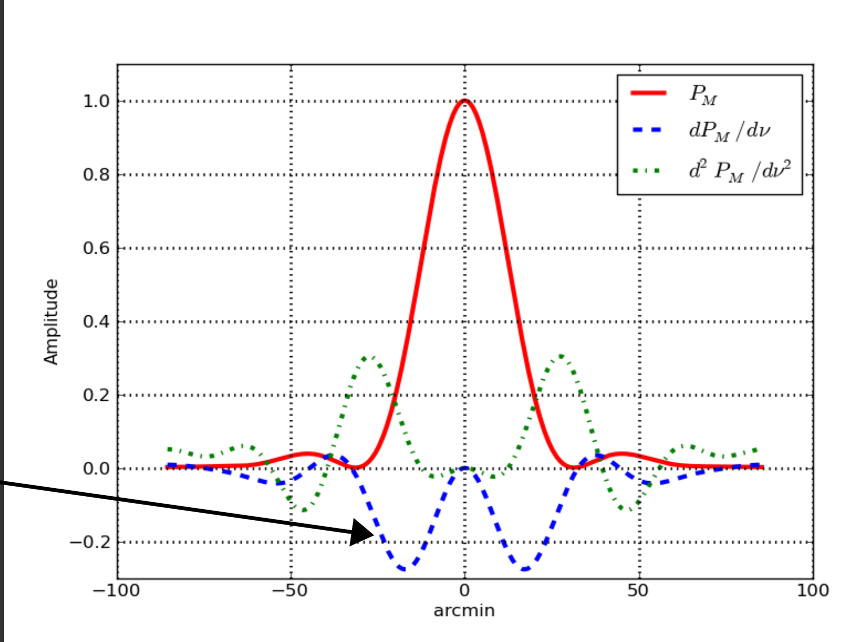
- L-Band VLA imaging
- DR ~ 10^4

Instrumental frequency dependence

- Continuum imaging $I^{continuum} = \int P_{ij}(\mathbf{s},\nu,t)\; I(\mathbf{s},\nu)\; \boxed{d\nu}$

- Antenna PB ($The\;\; P_{ij}(\mathbf{s},\nu,t)$)
  - Frequency dependence

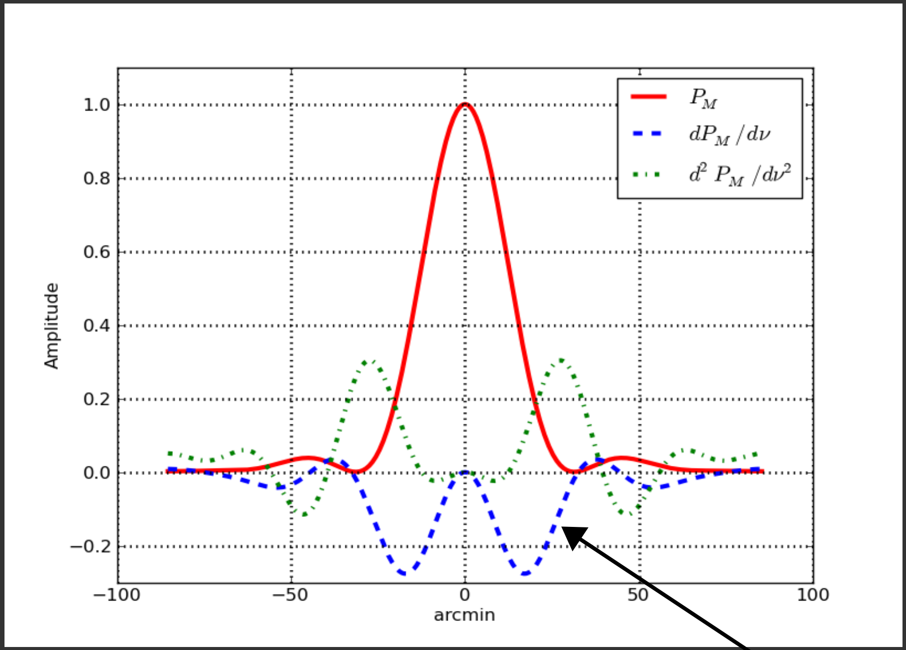


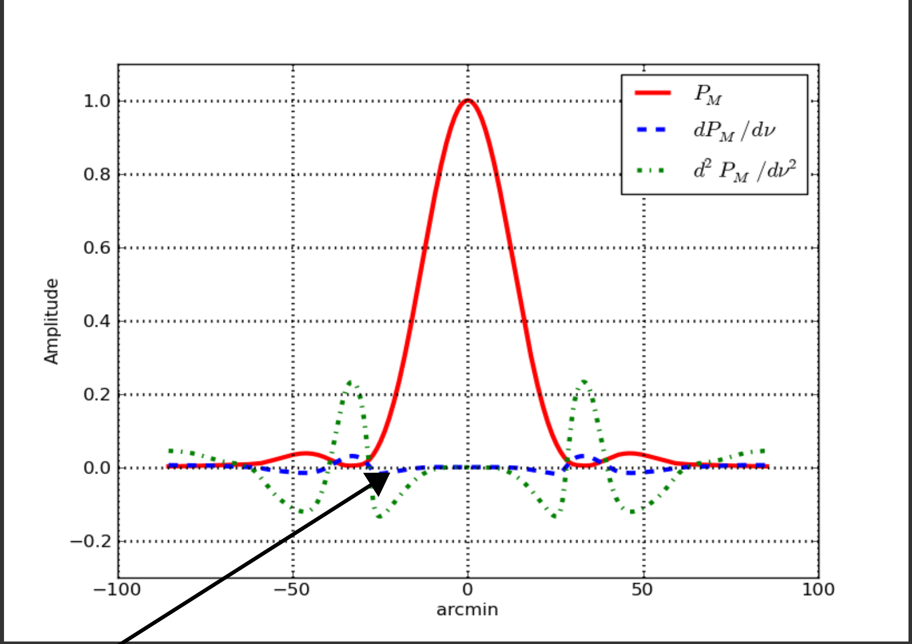


PB Freq. dependence (blue curve)

Wide-Band AW-Projection

- Correct for PB effects + W-term
  - Polarization: Squint + in-beam polarization
  - Time variability: Rotation with Parallactic Angle



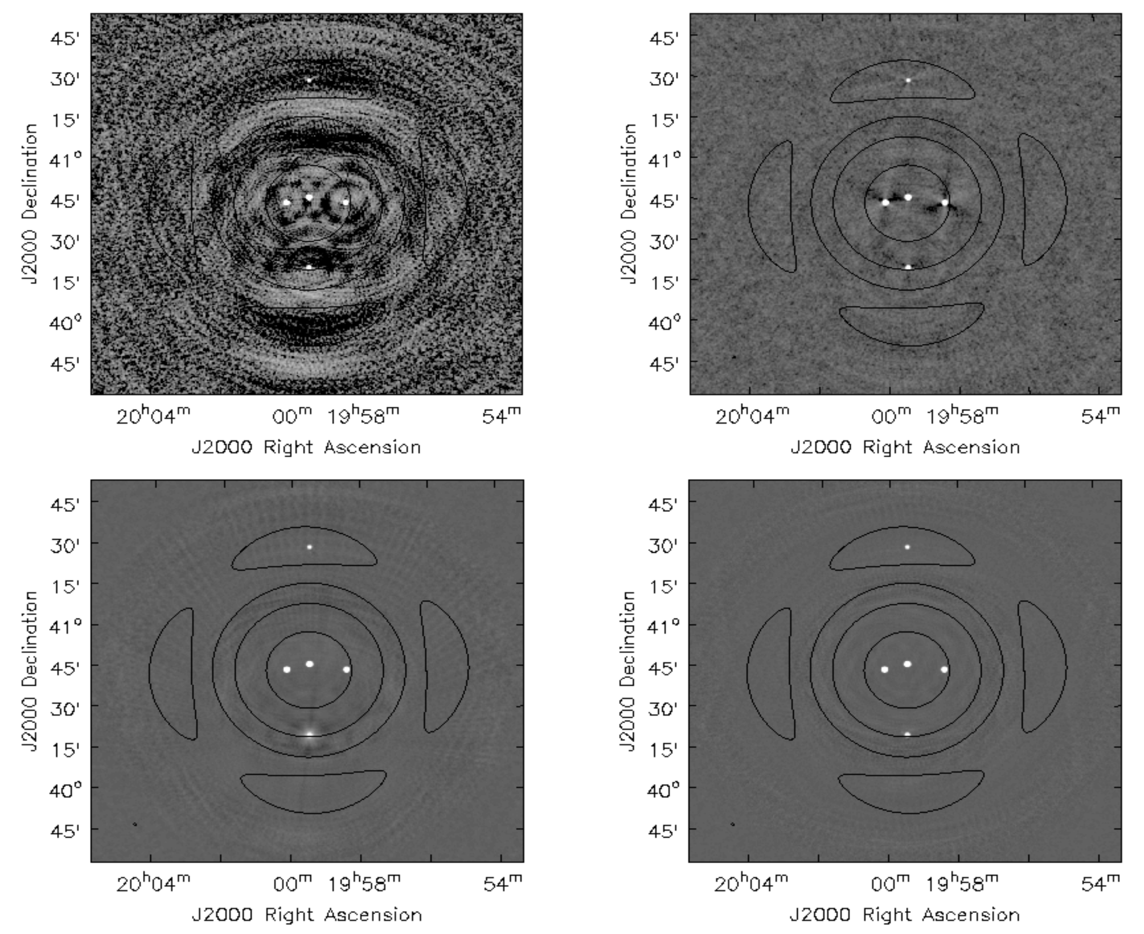


PB Frequency dependence (blue curve)

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

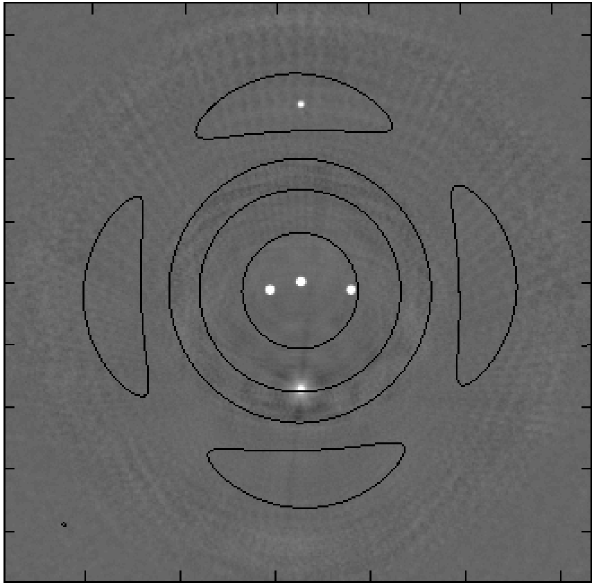
MT-MFS+SI

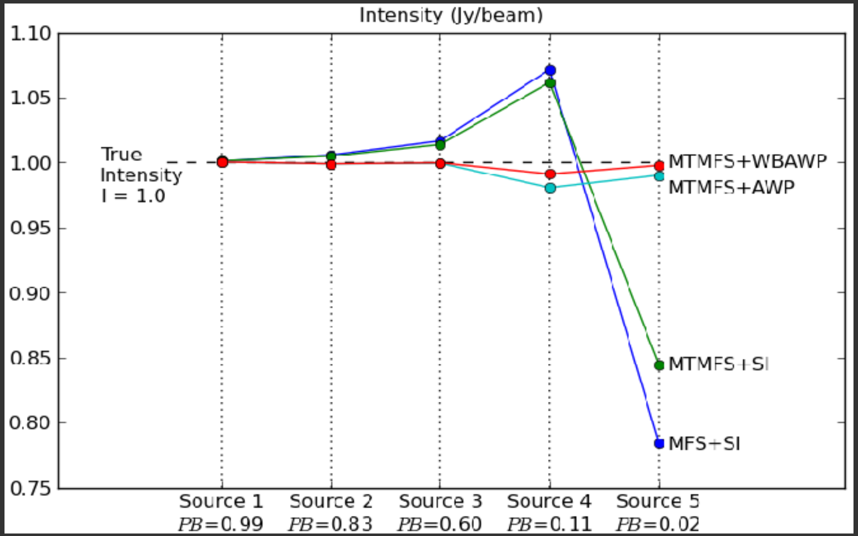
MT-MFS+ A-Projection

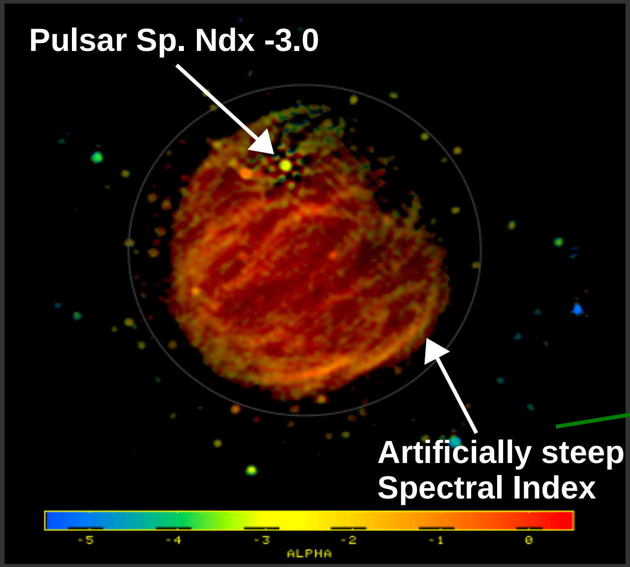
MT-MFS+ WB A-Projection

Ap.J., 2013

Instrumental frequency dependence

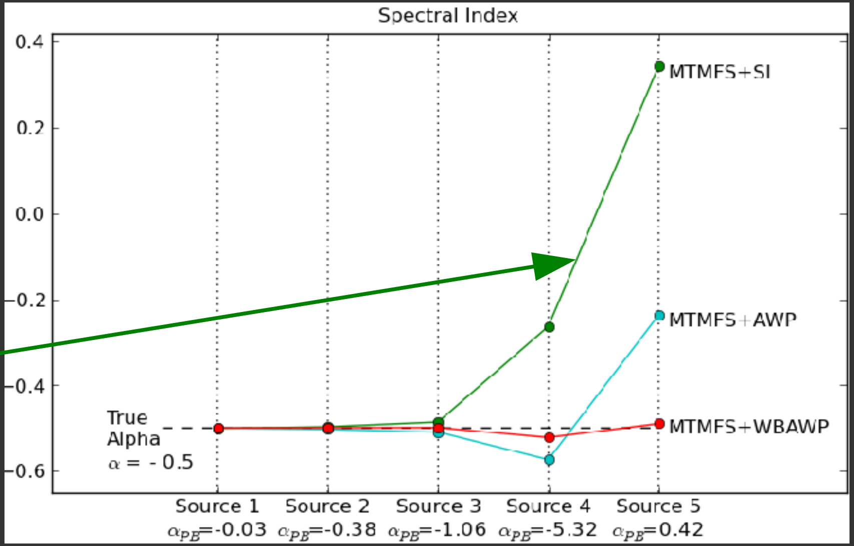






Pulsar Sp. Ndx -3.0

Artificially steep Spectral Index



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

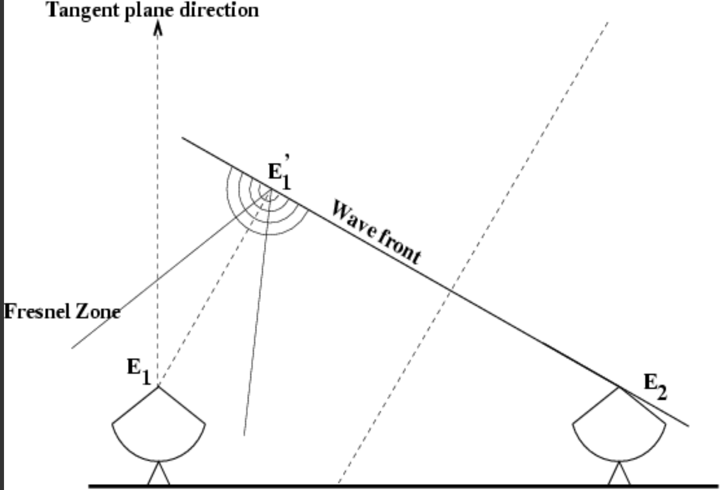
DI Calibration

DD Term
Instrumental
Ionospheric

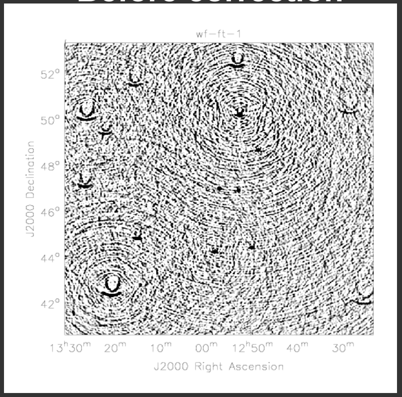
Sky

Geometry

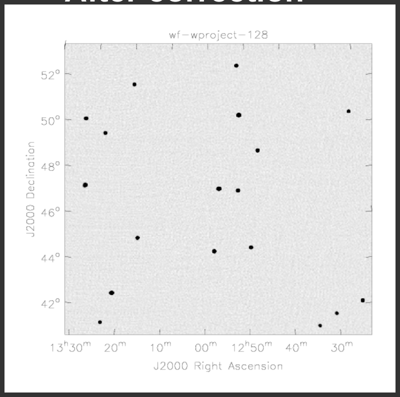
- The geometric term (non co-planar baselines)
  - Transform is no more 2D Fourier Transform



Before correction

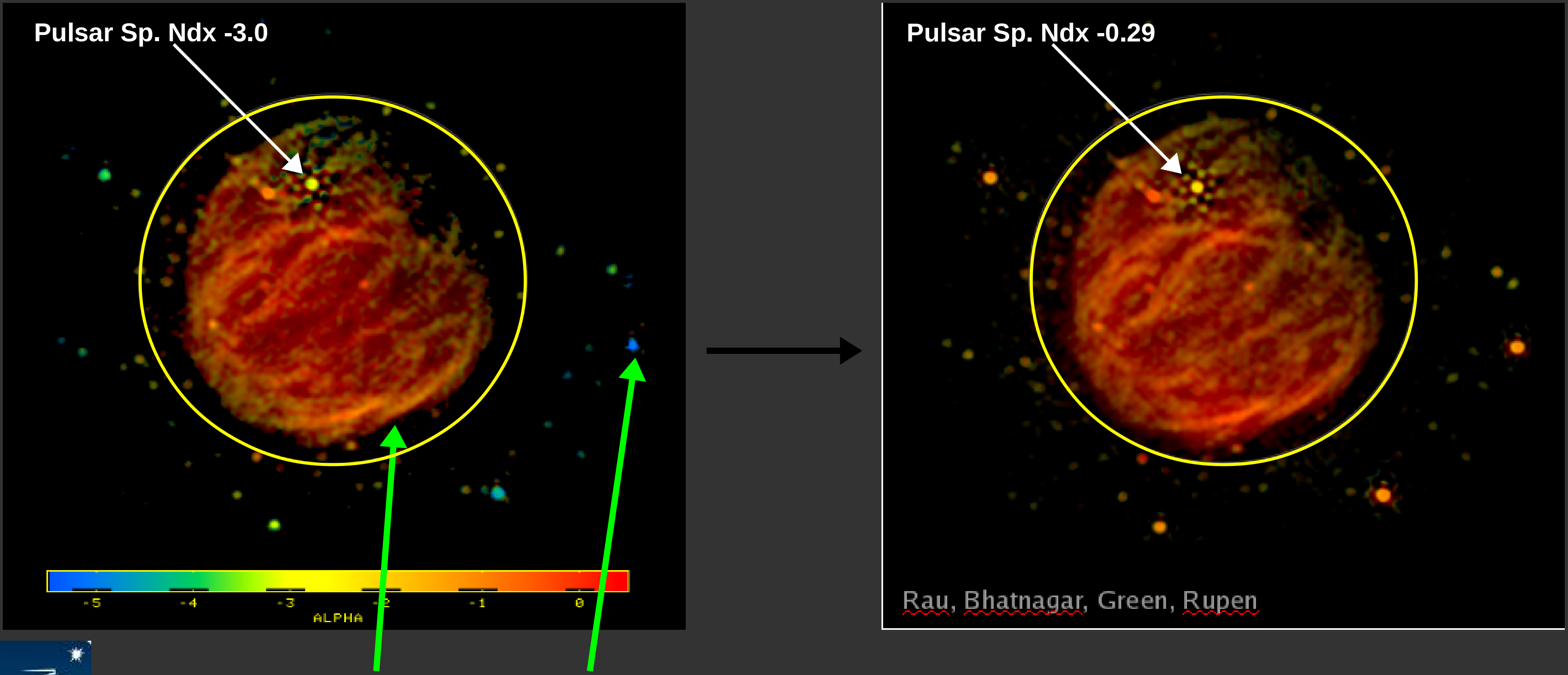


After correction



Wide-Band AW-Projection + MT-MFS

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

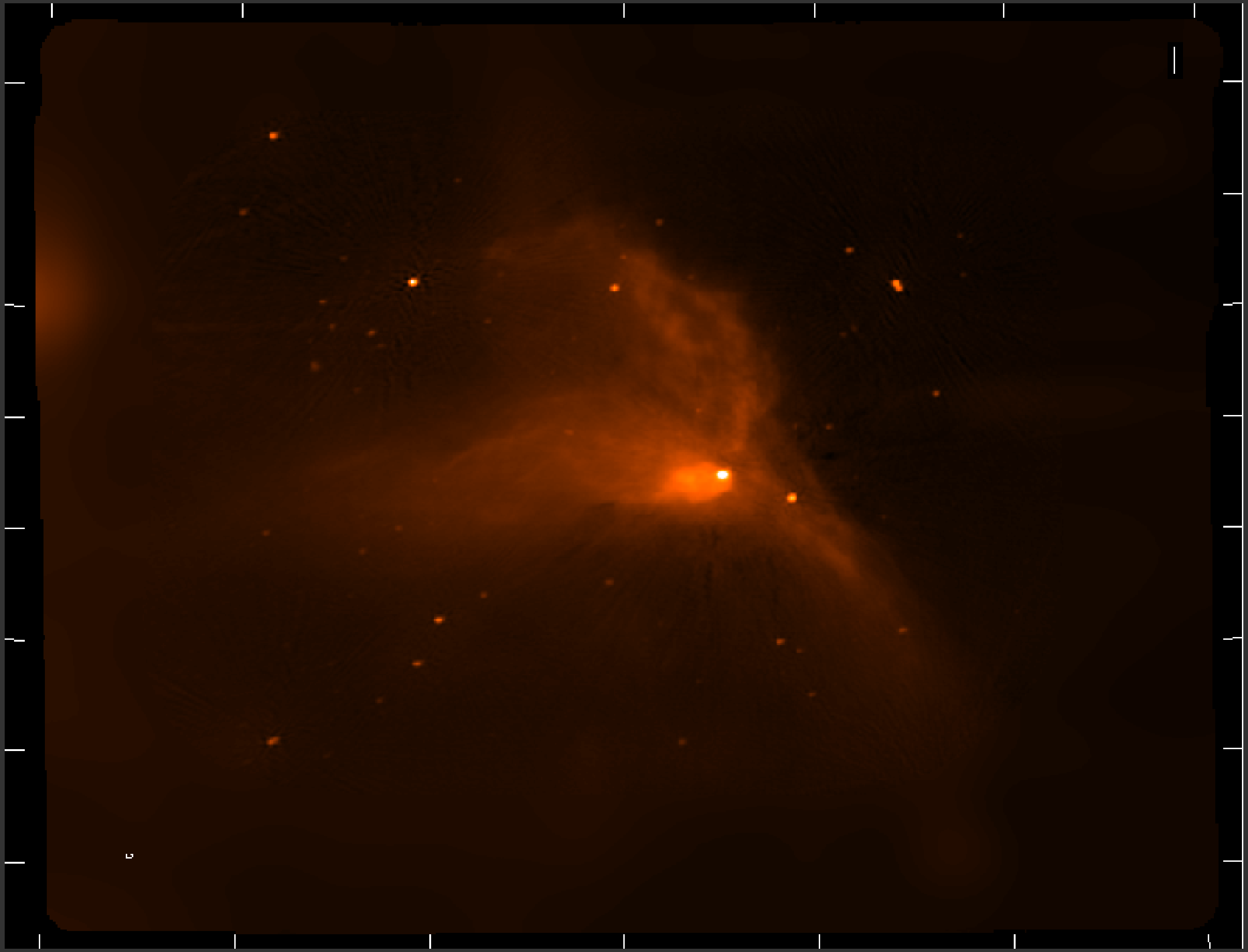


Rau, Bhatnagar, Green, Rupen

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 100-pointing mosaic
- EVLA + GBT Feathering (existing algorithm)
- In progress:
  - Mosaic spectral Index mapping
- Parallel execution / Optimization /
- Numerical tests

Status-1: In production or commission stage

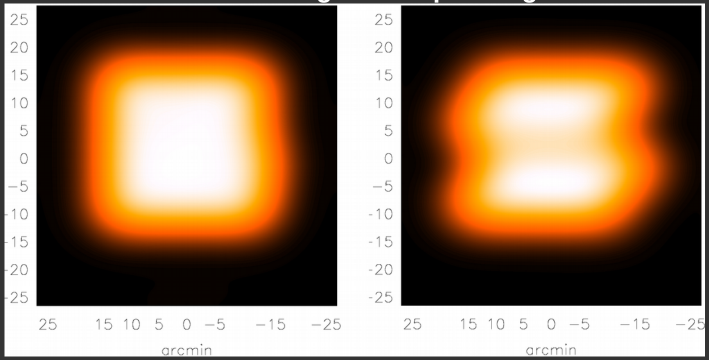
- W-Term correction: Dominant DD term at low frequencies
  - Facted-imaging, W-Projection, W-Stacking
- Extended emission
  - MS-Clean, Asp-Clean, various variants
- Frequency dependence of the sky brightness
  - MS-MFS, MT-MFS
- PB corrections
  - A-Projection: Time and polarization dependence
  - WB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

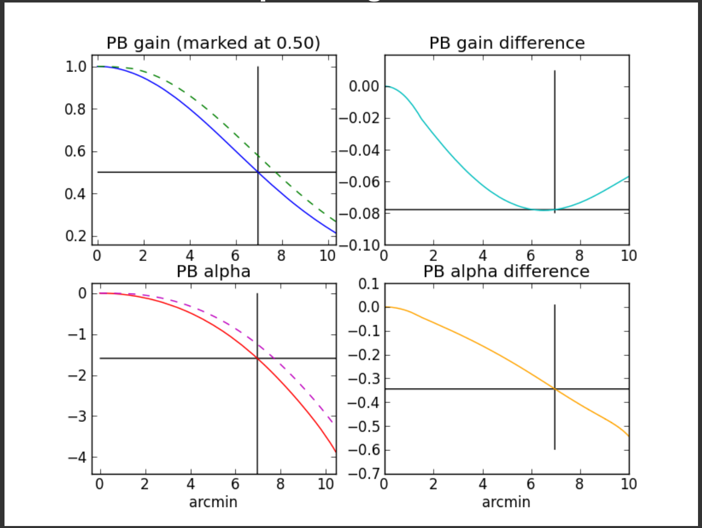
Status-2: In active development

- Imaging with heterogeneous arrays
  - Antenna-to-antenna variations: ALMA, ngVLA (,SKA?)
  - Pointing errors (VLASS 40-pointing mosaic imaging case)

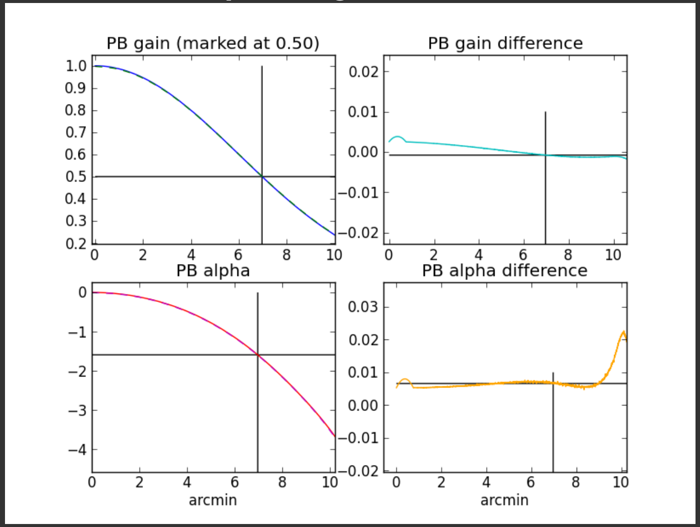
Effect of heterogeneous pointing errors



Before pointing corrections



After pointing corrections



Status-2: In active development

- 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,...

- Ionospheric phase corrections
  - Corrections: Via A-Projection for correction during imaging
  - Ionospheric phase screen solvers: Various “peeling” based solvers
    - » More generic solvers

Computing Cost

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

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

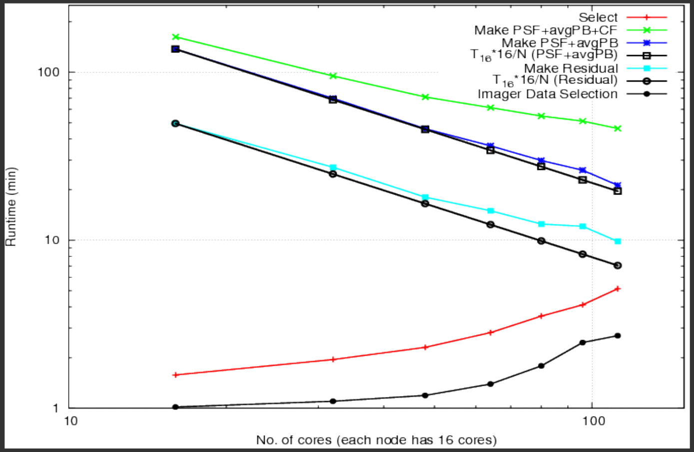
- Computing Scaling
  - Computing costs: $N^2_{\text{support}} \times N_{\text{vis}}$: Dominated by Projection
  - Memory footprint: $N^2_{\text{Scales}} + N^2_{\text{Terms}}$: Dominated by MT-MFS

- Imaging : Pleasantly Parallel (a.k.a “Embarrassingly parallel”)
  - Scatter-Gather Paradigm on the Cluster scale

- Optimal utilization of the computing multi-core CPUs is harder
  - Multiple process per node: Limited by total memory footprint
  - Single multi-threaded process: Algorithmically challenging

Sci. S/W Complexity: High Level Description

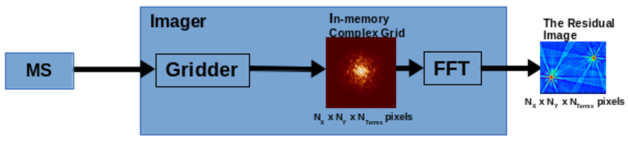
- High Computational Intensity (FLOP per byte)
    - O(10^{2-3}) FLOP per data point
    - Number of data points: O(10^{12-15})

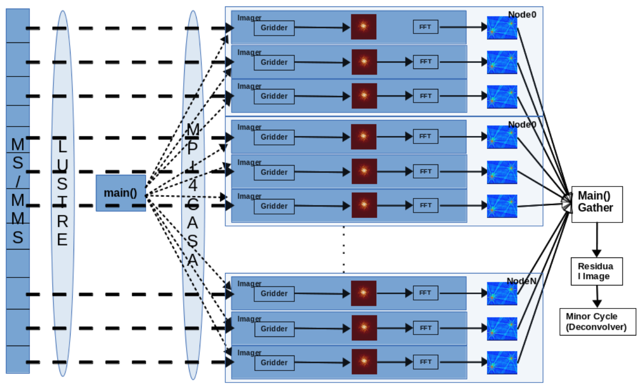


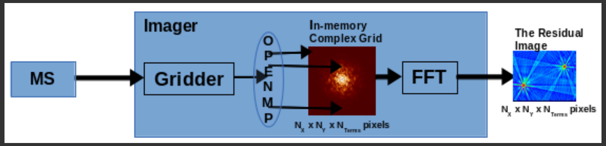
- Imaging is embarrassingly parallel
    - SAMD parallelization architecture measures high efficiency
    - In-coherent gather is OK

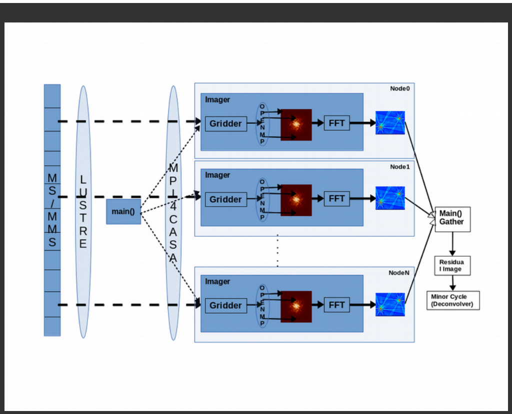
$$I^D=\sum_p FXG_pV_p \quad \text{or} \quad F\sum_p XG_pV_p \quad \text{or} \quad FX\sum_p G_pV_p$$

← Node s/w complexity









Challenges

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

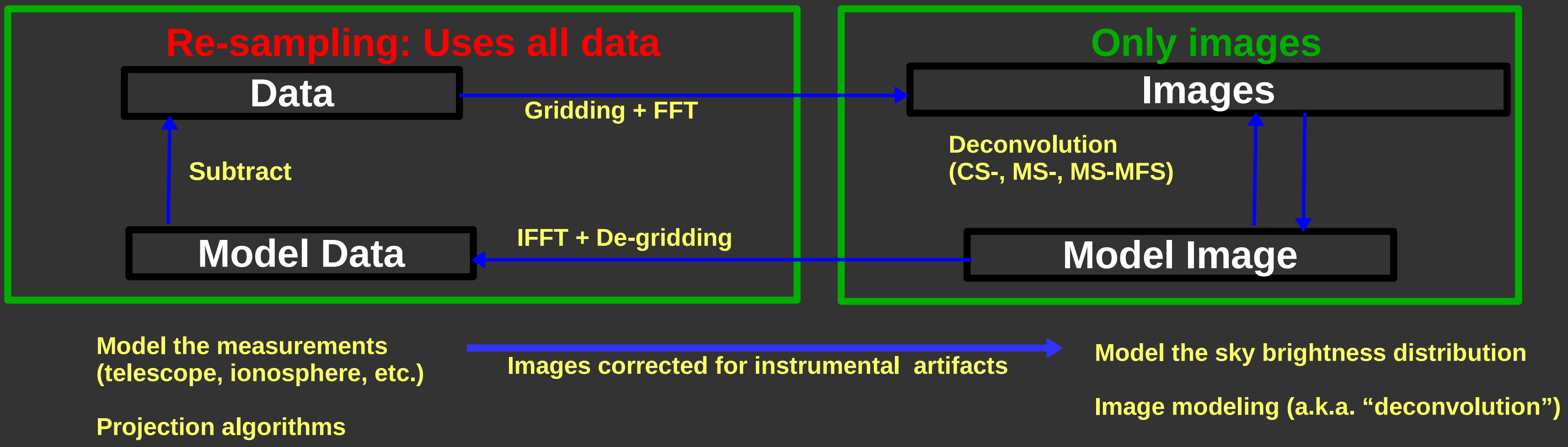
- High Scale-dynamic range imaging:
  - Ratio of max. to min. scale: O(10^5)!
    - Imaging with the EVLA A + B + C + D-array
    - ALMA long baselines + Core
    - ngVLA in general

- 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

Challenges

- Rate of convergence: Crucial for ngVLA-scale problems
  - Optimal algorithms, Optimal utilization
- ngVLA:
  - (Very) high data rates
  - Pipeline processing
  - Computing, algorithmic, hardware, software solutions to reduce computing cost
- SKA sensitivity → wider-field imaging, expose more error terms
  - Instrumental terms: Measure vs Model vs Solve
- We collect enormous amounts of data → more information
  - Are we utilizing the available information optimally?
    - In terms of algorithm design
    - In terms of extracting astrophysical information

Computing Cost



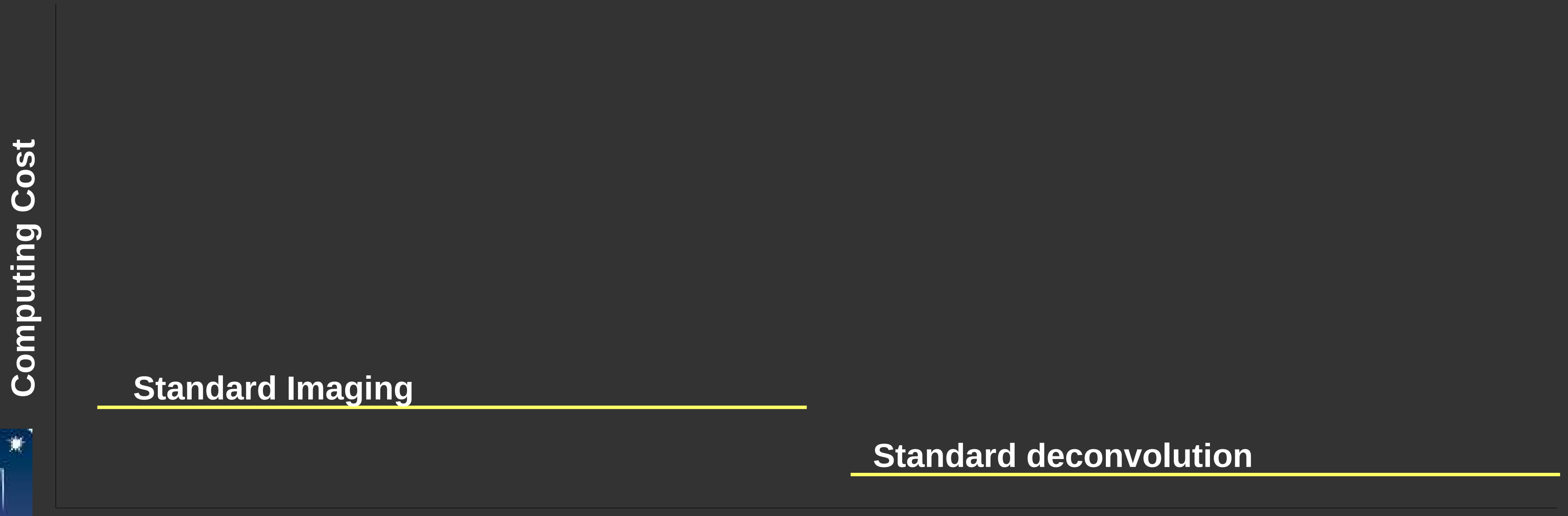
Model the measurements
(telescope, ionosphere, etc.)

Projection algorithms

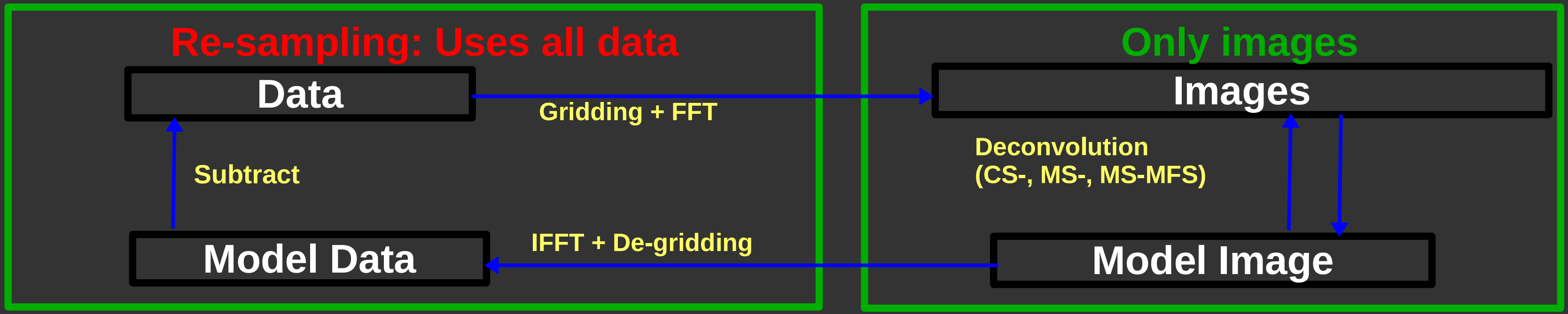
Images corrected for instrumental artifacts

Model the sky brightness distribution

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



Computing Cost



Re-sampling: Uses all data

Data

Gridding + FFT

Subtract

Model Data

IFFT + De-gridding

Only images

Images

Deconvolution
(CS-, MS-, MS-MFS)

Model Image

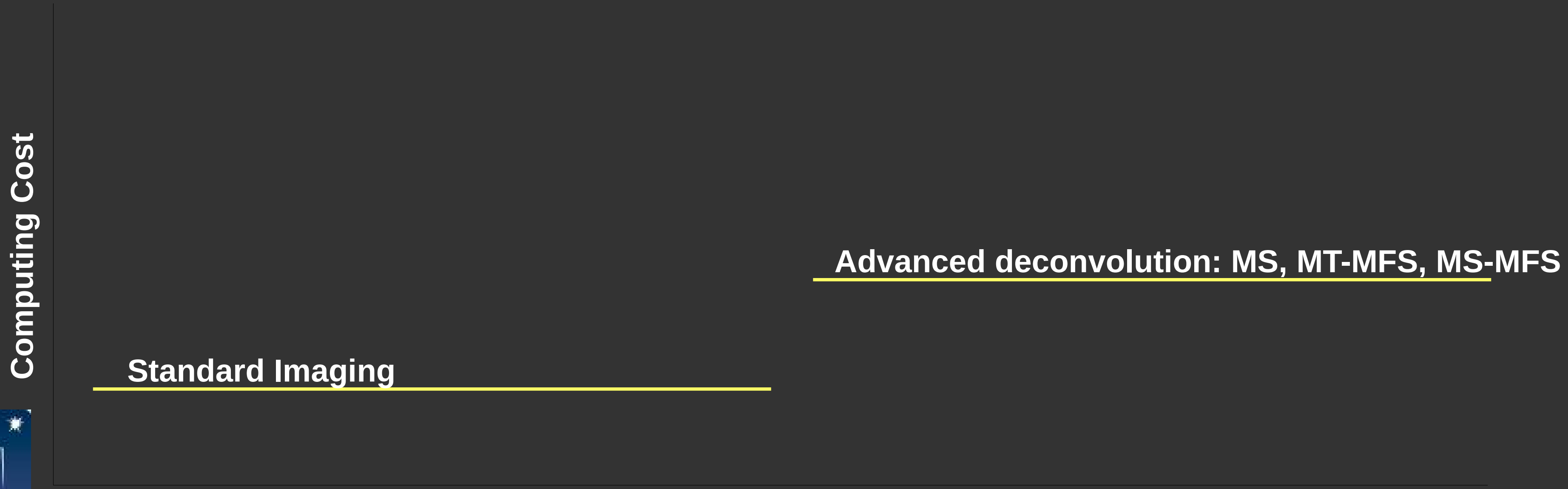
Model the measurements
(telescope, ionosphere, etc.)

Projection algorithms

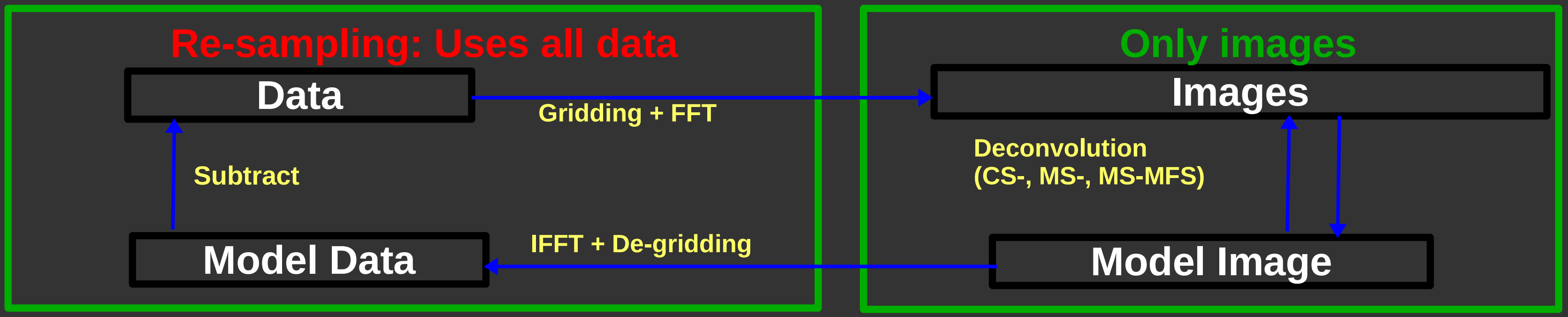
Images corrected for instrumental artifacts

Model the sky brightness distribution

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



Computing Cost



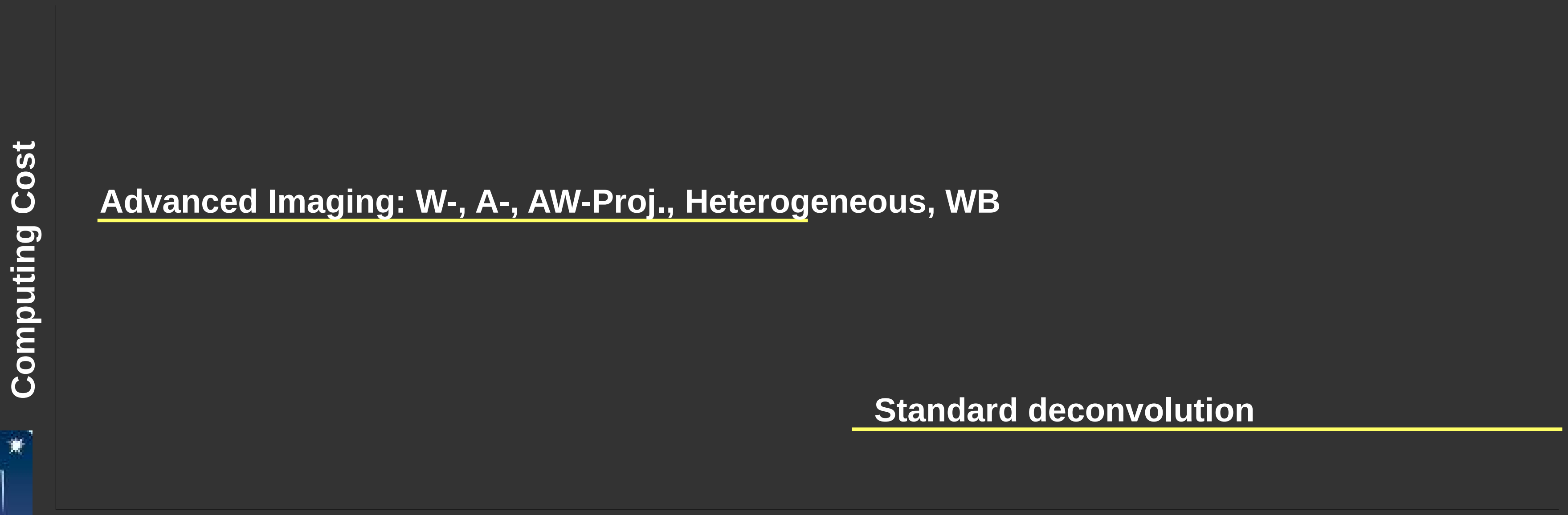
Model the measurements
(telescope, ionosphere, etc.)

Projection algorithms

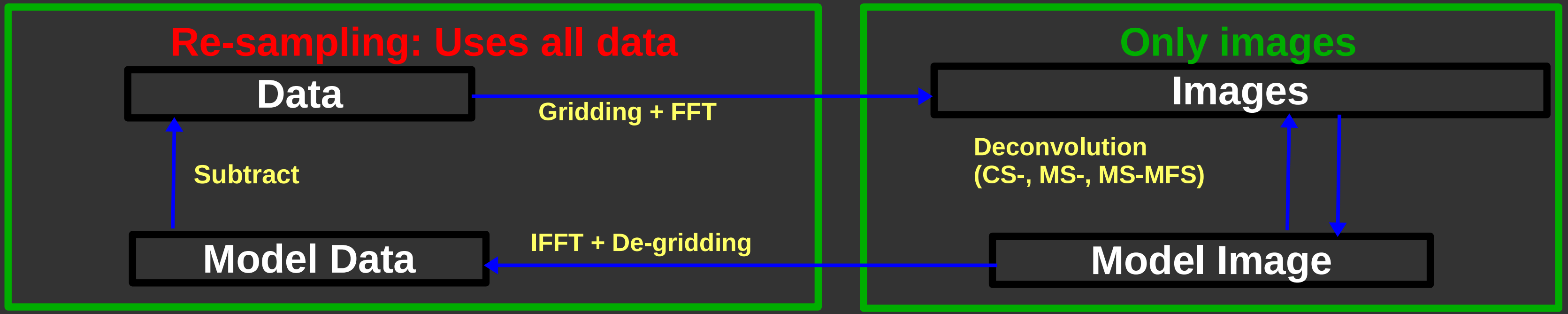
Images corrected for instrumental artifacts

Model the sky brightness distribution

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



Computing Cost



Re-sampling: Uses all data

Only images

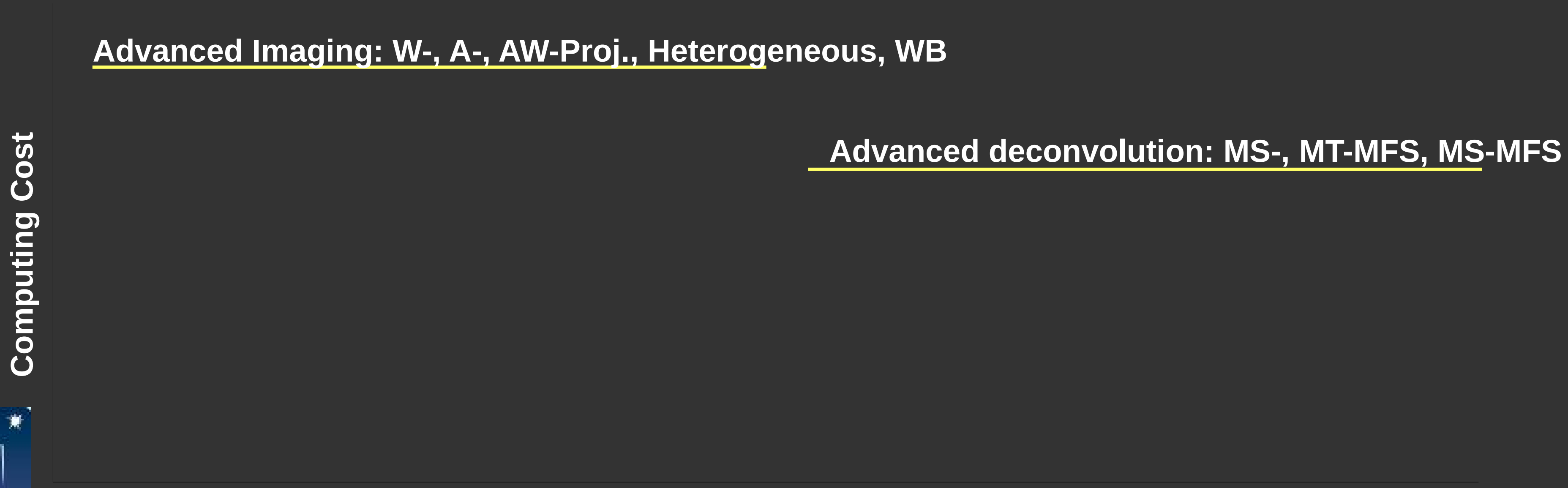
Model the measurements
(telescope, ionosphere, etc.)

Projection algorithms

Images corrected for instrumental artifacts

Model the sky brightness distribution

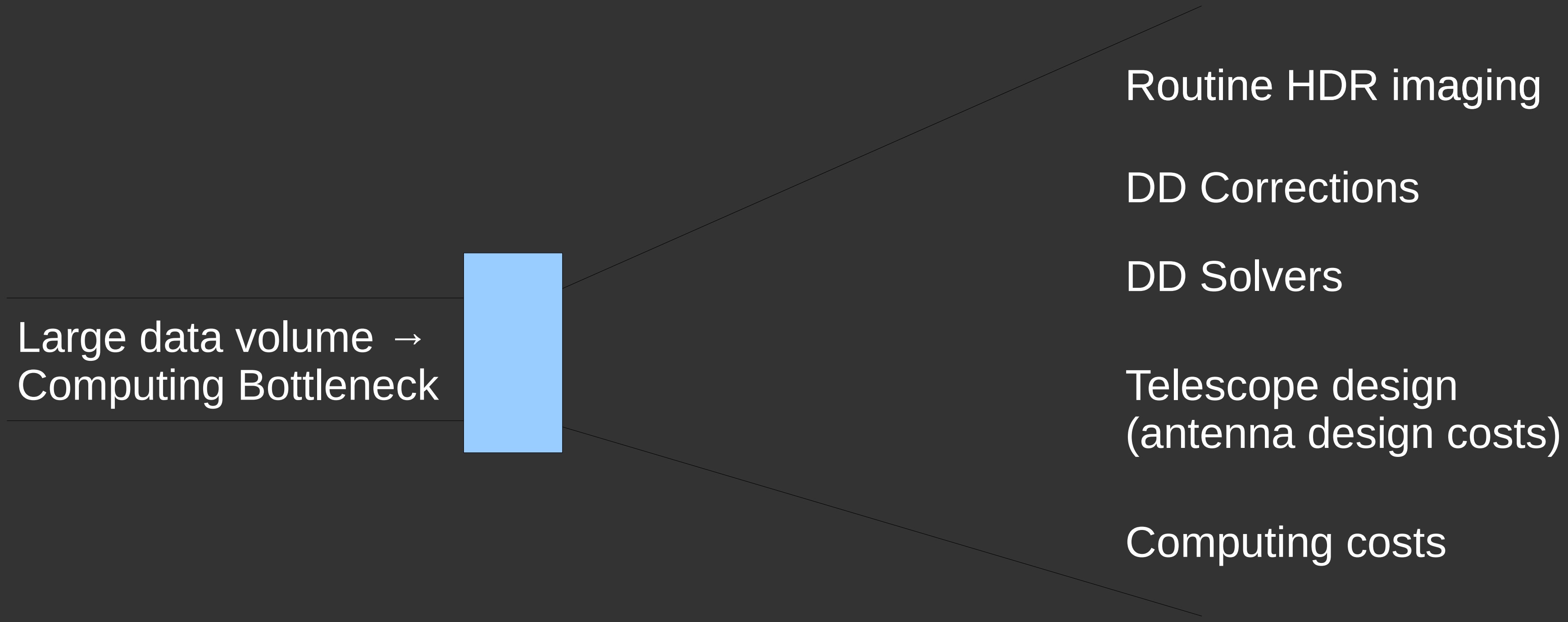
Image modeling (a.k.a. “deconvolution”)



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

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

Challenges



Large data volume → Computing Bottleneck

Routine HDR imaging

DD Corrections

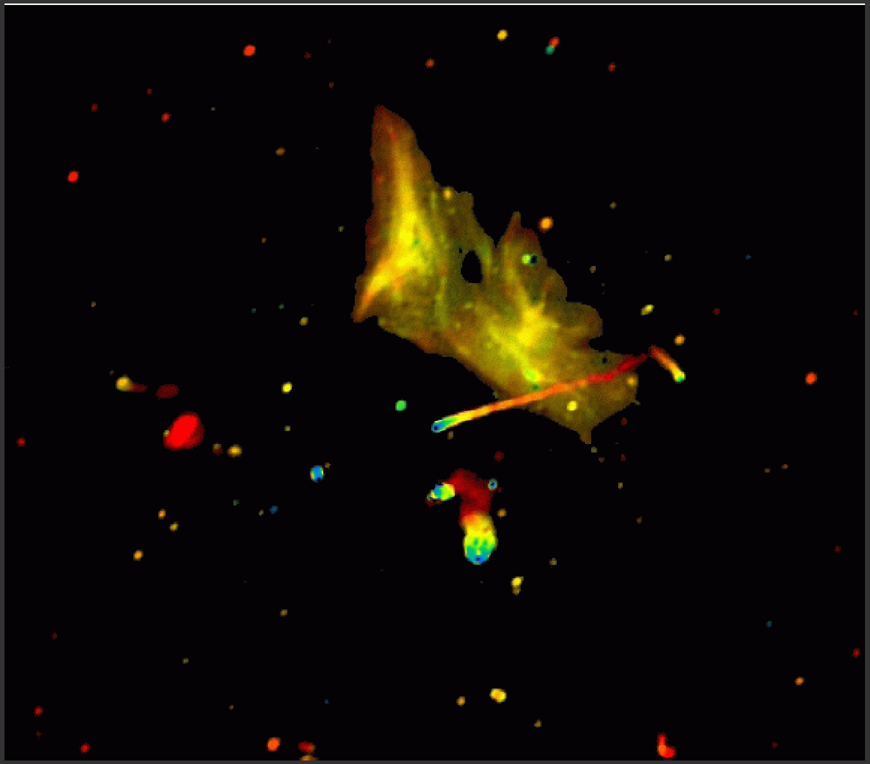
DD Solvers

Telescope design (antenna design costs)

Computing costs

Existence of large computing resources does ensure algorithms will converge!

Imaging with the EVLA @ L-Band



Single pointing, wide-band image

Wide-band 100 pointing mosaic+Single Dish