Wide-band Wide-field Imaging

Colloquium, Socorro, Feb. 11th 2011

S. Bhatnagar

K. Golap, U. Rau, J. Robnett

NRAO

Algorithms R&D Group activities

- R&D for new post-processing algorithms required for wide-band wide-field full-polarization imaging...in a reasonable computing time.

- Various activities, not all of which I will go in detail today:
  - Wide-field imaging and calibration [active]
  - Wide-band imaging and calibration [active]
  - High Performance Computing [active]
  - RM-Synthesis [active]
    [AIPS Task FARS; Kogan,Greisen,Owen]

  - Wide-band mosaicking [active/on wait]
  - Automatic RFI removal [active/R&D+testing]
  - Improvements in Scale Sensitive Image reconstruction [R&D/planning]
  - Wide-band on-axis calibration, DD-Calibration [Advanced R&D]

Interferometric Imaging

- Interferometric telescopes are indirect imaging devices
  - Observations are in the Fourier domain: The Coherence Function
- van-Cittert Zernike Theorem: Coherence Function is 2D Fourier transform of the Sky Brightness distribution

$$V^{Obs}(u_{ij}, v_{ij}) = S(u_{ij}, v_{ij}) \int I(l, m) \; e^{\iota [u_{ij} l + v_{ij} m]} dl \; dm$$

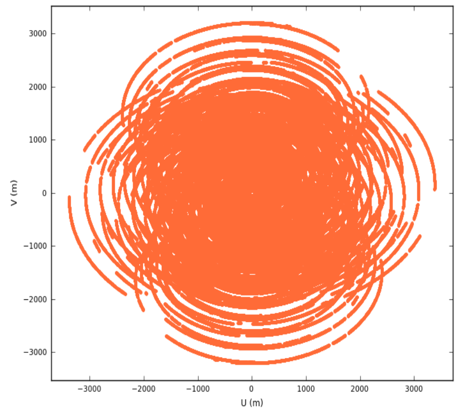
$$= S(u_{ij}, v_{ij}) \cdot V^{Sky}(u_{ij}, v_{ij})$$

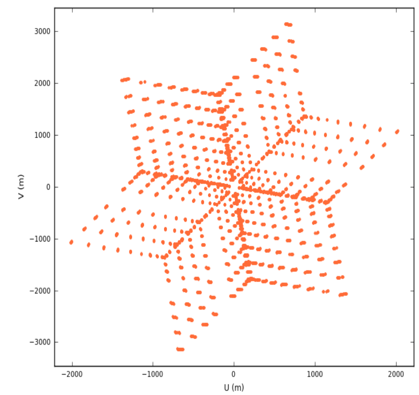
  - Sampling function(S) encodes the incomplete sampling of the data domain
  - (u_{ij}, v_{ij}) are implicitly a function of time

  - Aperture Synthesis:
    - Integration in time → Leads to wide-field issues
    - Integration in frequency → Leads to wide-band issues

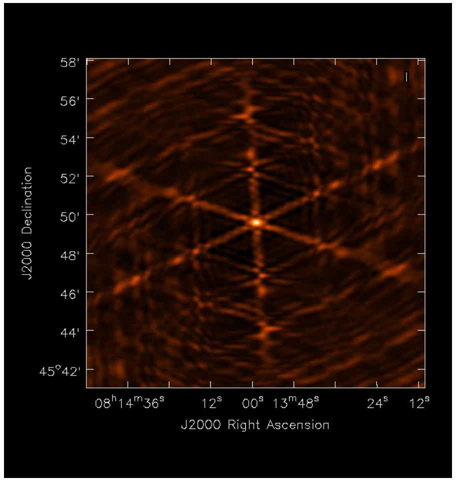
Interferometric Imaging

- $V_{ij}^{Obs} = \sum_t S_{ij}(t) \cdot V_{ij}^{Sky}(t)$





- $I^{Obs} = FT\left[\sum_t S_{ij}(t) \cdot V_{ij}^{Sky}(t)\right] = \sum_t \left[PSF(t) * I^{Sky}\right]$



- Deconvolution algorithms assume
  - I^{sky} is time-invariant
  - I^{sky} is frequency-invariant

High sensitivity imaging

- Sensitivity $\propto \frac{N_{ant}(\eta A_{ant})\sqrt{(N_t \tau)(N_{chan} \Delta \nu)}}{T_{sys}}$
- Data volume $\propto \mathbf{N^2_{ant} \, N_{channels} \, N_t}$

- Higher sensitivity is achieved using larger bandwidths (e.g. EVLA) or larger collecting area (e.g. ALMA) or both (SKA PF).

- Higher sensitivity → Wide-field issues
  - Sources farther out also affect imaging performance
- Long integration in time
  - Need to account for time-variability, farther out
- Long integration in time over wide bandwidths
  - Account for time & frequency dependence of the instrument
  - Account for frequency dependence of the sky

Synthesis Imaging Measurement Eq.

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

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

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

- Today's discussion will use
  - *$M_{ij}^{S}(s,\nu,t)$ to represent antenna Primary Beams (PB)*
  - *$I(s,\nu)$ to represent frequency dependent extended sky-emission*

Image Domain: $I^{Obs} = \sum_t \sum_\nu PSF(\nu,t) * \left[PB(s,\nu,t)\, I^{Sky}(\nu)\right]$

Data Domain: $V_{ij}^{Obs}(\nu) = S_{ij}(t)\left[A_{ij}(\nu,t) * V(\nu)\right]$

$A_{ij}(\nu,t)$ *is correlation of Antenna Aperture Illumination patterns*

Deconvolution and Calibration: Theory

- Calibration and image deconvolution operations can be described as function optimization

$$V^{Obs} = M\, A\, M^S\, I^{True} + N$$

- Image deconvolution (CLEAN, MEM,...) estimates model parameters for the sky-emission

$$\chi^2 = |M^{-1} V^o - A\, I^M|^2 \quad where\ I^M = \sum_k P_k\,; P_k \text{ is the Pixel Model}$$

- Calibration ("antsol", "self-cal")

$$\chi^2 = |V^o - M\, A\, I^M|^2$$

- Corrections for DI terms (M) can be done independent of imaging

  Corrections for DD terms can only be done *during imaging*

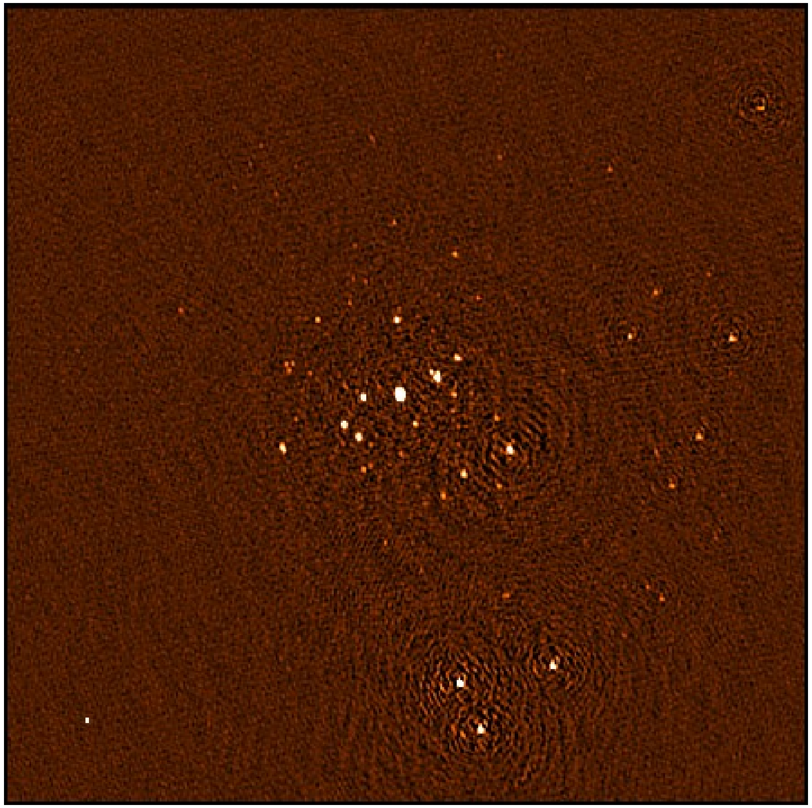
- Accounting for DD terms *fundamentally* couples calibration and imaging.

Advances in Calibration and Imaging Techniques in Radio Astronomy, Rau et al., Proc. IEEE, Vol. 97, No. 8, Aug.2009, 1472

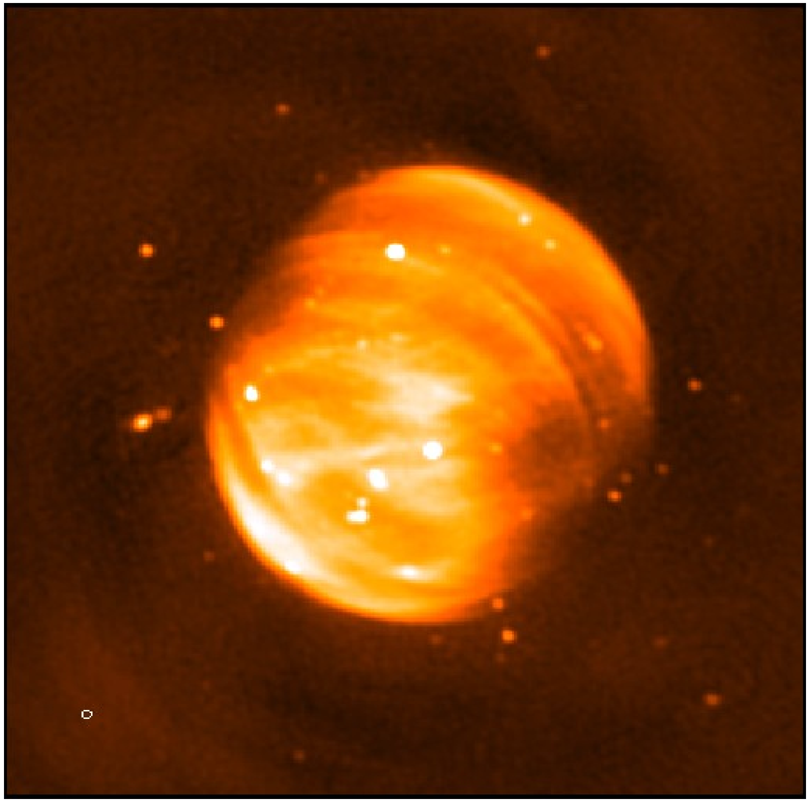
Wide-field wide-band imaging issues

- Wide-field Imaging: Antenna Primary Beams vary in time and direction
  - Residual errors due to conventional imaging techniques are significant
- Wide-band Imaging
  - Antenna Primary Beams & Sky emission vary with frequency
  - Both affects are directionally dependent
- Data volume increase by $10^{2\text{-}3}$x
  - Computing and I/O load increase
  - Deployment on HPC platforms
- Direction dependent calibration
  - Instrumental gains vary across the FoV

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

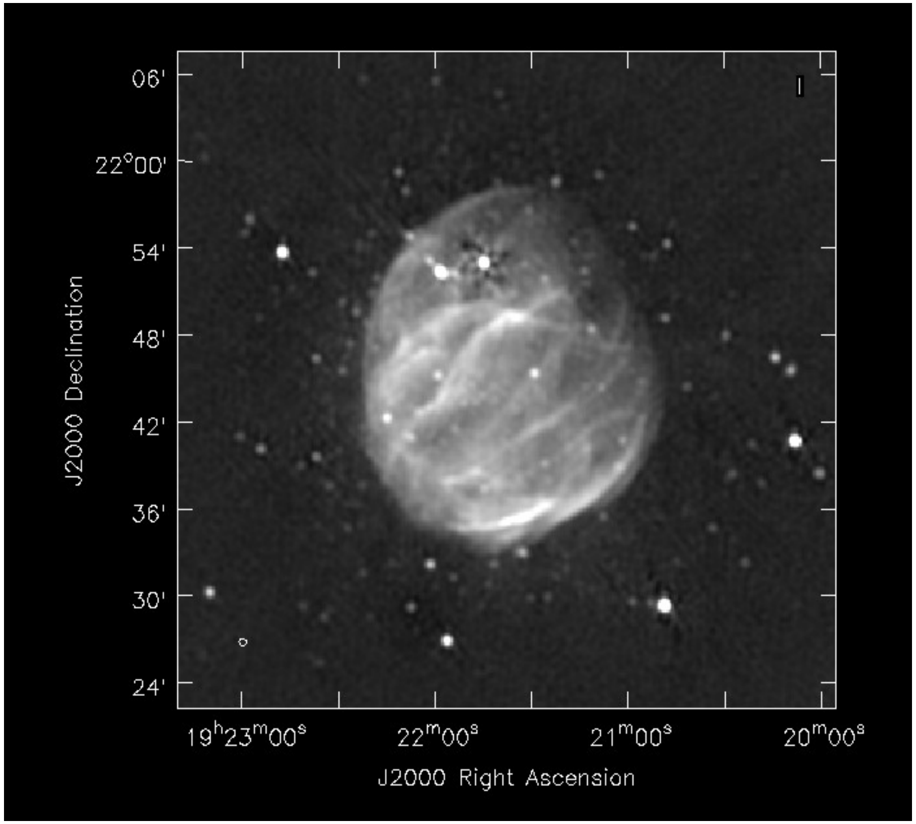
- Two approaches
  - **Faceting:** Partition the data & apply DI techniques per facet
    - Use DFT, multiple passes through the data
    - Difficult to generalize for DD correction/calibration
    - Higher algorithmic and software complexity
  - **Global/Projection methods:** Include DD terms in the Measurement Equation
    - FFT, single pass through the data
    - Parametrization in the natural domain
    - Lower complexity
- Noise per antenna based DoF:

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

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

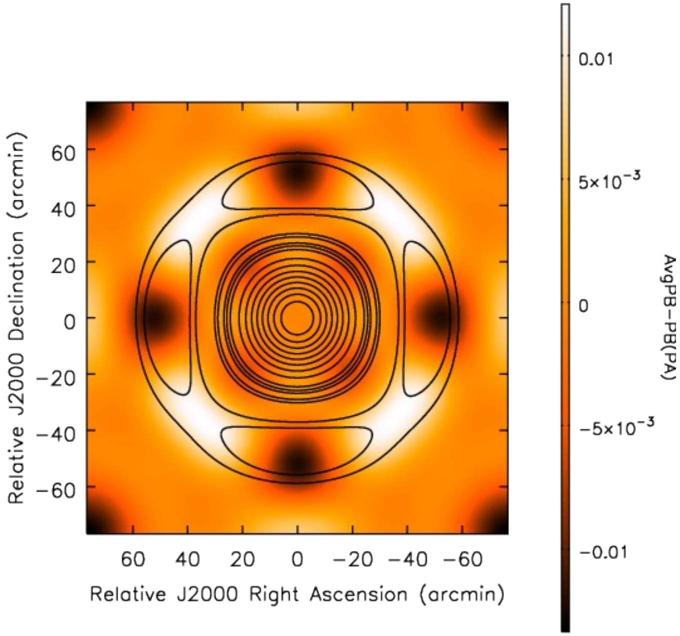
High sensitivity imaging

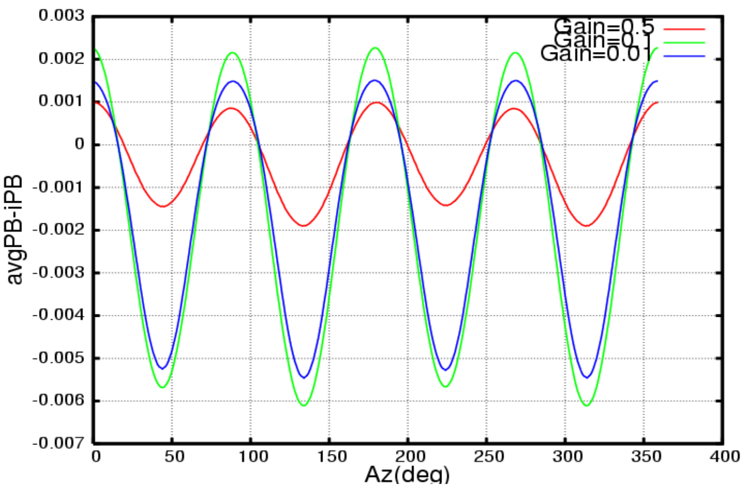
- Higher sensitivity is achieved using larger bandwidths (e.g. EVLA) or larger collecting area (e.g. ALMA) or both.
  - Sources farther out also affect imaging performance



High sensitivity imaging

- Time variability due to antenna Primary Beams increase away from the pointing center
  - Due to PB rotation asymmetry, rotation with PA and pointing errors

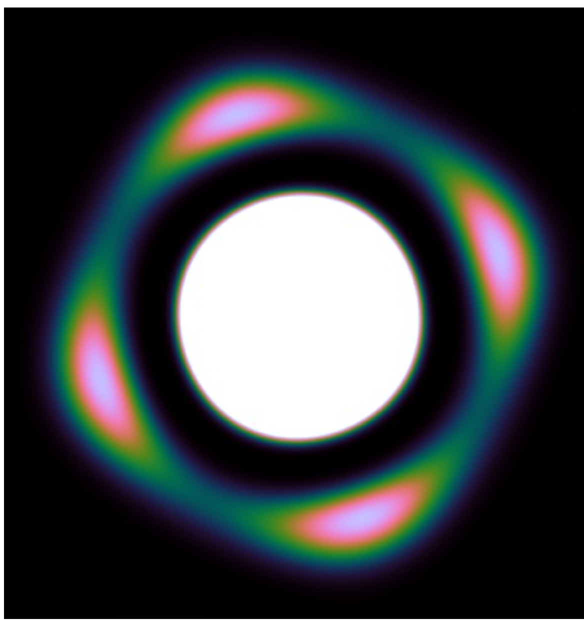


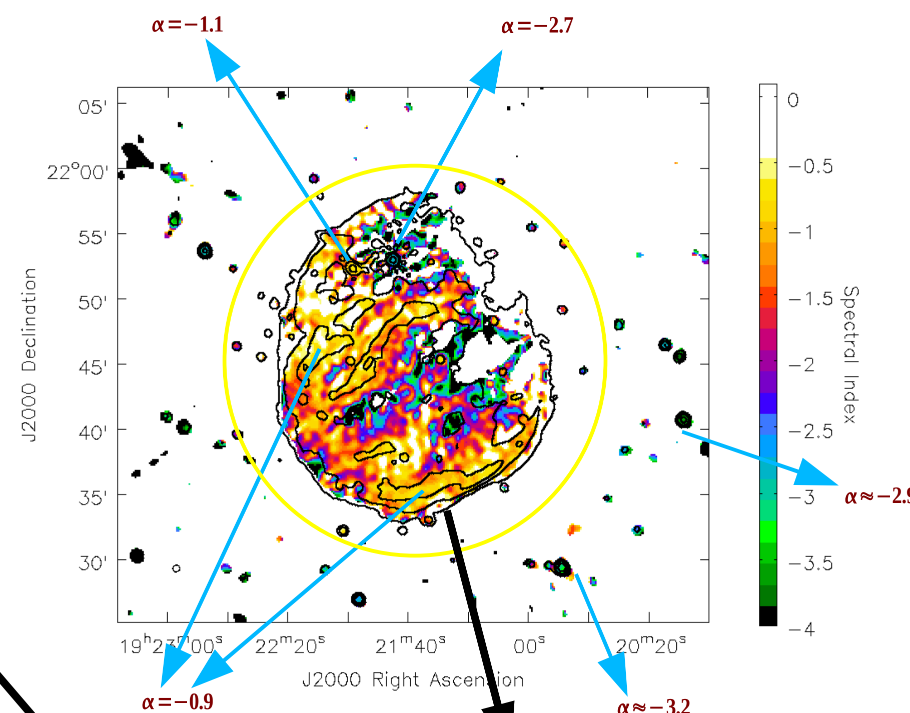


Difference between rotationally symmetric PB model and more realistic PB

High sensitivity imaging

- Time variability of the PB increases away from the center
- Frequency dependence increases with fractional bandwidth

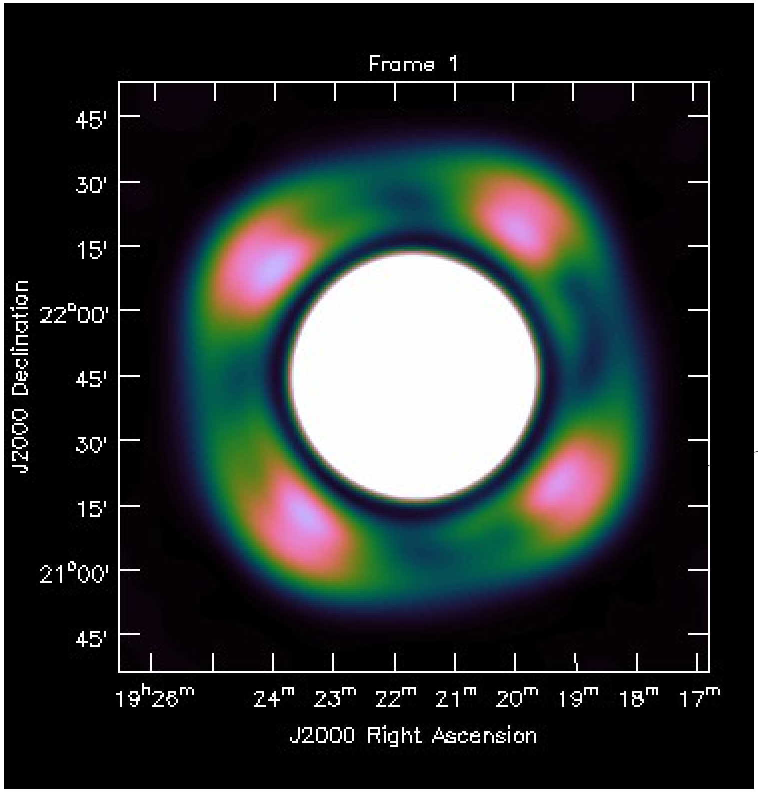




$$I^{Obs} = \sum_t \sum_\nu PSF(t) * [PB(s) I^{Sky}(s)(\nu/\nu_o)^{\alpha(s,\nu)}]$$

High sensitivity imaging

- To the first order, scaling of the PB with frequency

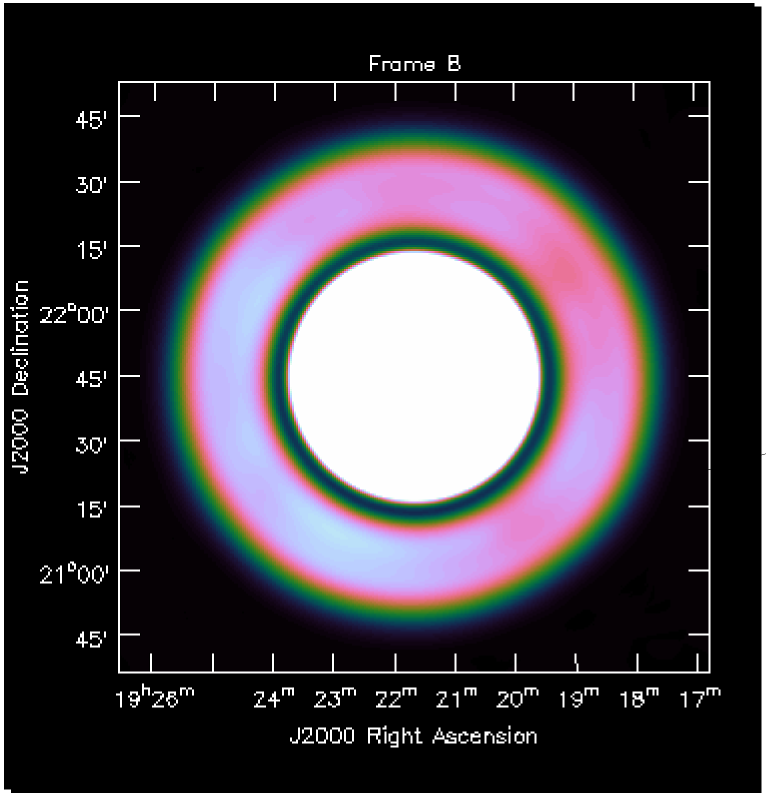


$$I^{Obs} = \sum_t \sum_\nu PSF(t) * \left[PB(s, \nu, t) I^{Sky}(\nu) \right]$$

$$\sum_\nu PB(s, \nu, t)$$

High sensitivity imaging

- Image corresponds to the sum of all the data.
  - Only average of antenna-based quantities are available in the image domain



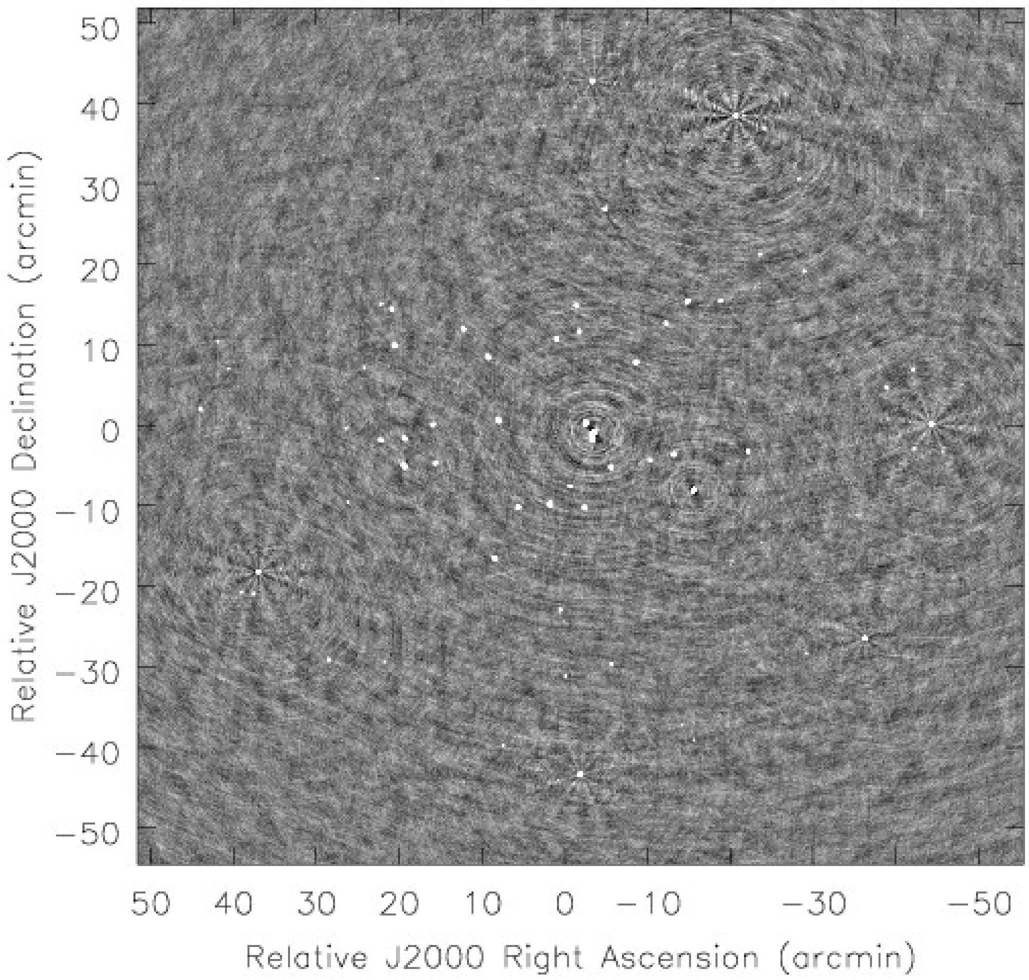
$$I^{Obs} = \sum_t \sum_\nu PSF(t) * \left[PB(s,\nu,t) I^{Sky}(\nu) \right]$$

$$\sum_t \sum_\nu PB(s,\nu,t)$$

- Image domain corrections for time, frequency and antenna dependence is hard

- Projection methods apply corrections in the Natural Domain
  - A-Projection for PB-corrections
  - W-Projection for W-term correction

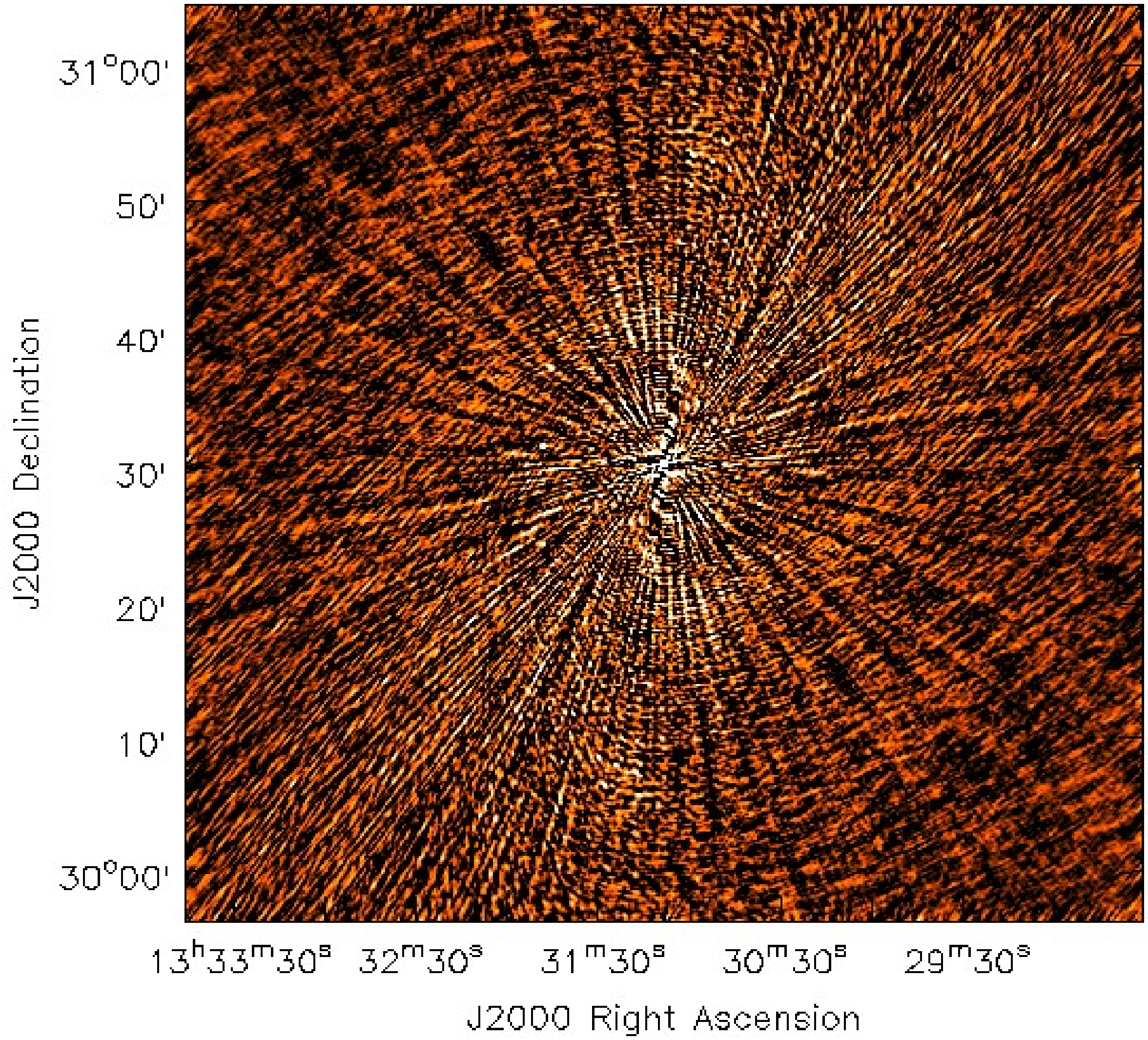
Implications for imaging: Wide-field effects



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

- Errors are due to time-varying Primary Beam
- Errors are directionally dependent
- Imaging performances of the telescope is limited by these errors (and not the thermal noise)

Implications for imaging: Wide-band effects



- 3C286 field
  I=14.4Jy @1.4GHz
  Sp.Ndx=-0.47
  BW = 1.1 GHz

- Conventional imaging
  - Frequency oblivious Image model

•DR = 1600-13000

Implications for imaging: Computing

- To keep time and band width smearing errors below thermal limit for wide FoV, needs finer sampling in time and frequency.
- Data volume $\propto$ $\mathbf{N^2_{ant}\ N_{channels}\ N_t}$
  - $\mathbf{N_{channels}}$ = 1-10GHz/KHz-MHz and $\mathbf{N_t}$ = 10hr/(1-10sec)
  - $\mathbf{N_{ant}}$ **=** 27 (EVLA), ~50 (ALMA), Cast of thousands (SKA)
  - 100-1000x increase in the number of samples to achieve the required sensitivities
    - Algorithm efficiency remains a critical parameter
- Algorithms for wide-field and wide-band effects require more floating point operations (FLOP)
  - Inherent information content in the data is higher
- Need computing platforms with (much) higher I/O rates and FLOPS (FLOP per sec) capacity.
  - ...and larger RAM (possibly)

Parametrization of the ME

Lower the number of parameters in the model that leaves noise-like residuals, higher is the information extracted.

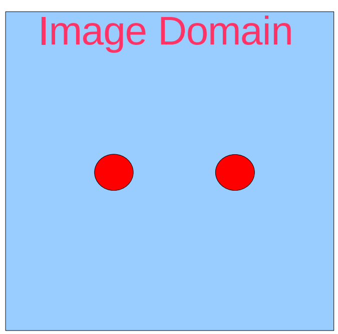
– Papers on Information Theory (possibly by Donoho, 2000)

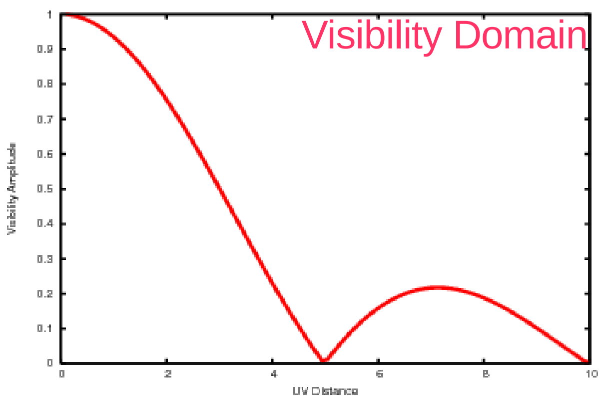
- Models in the Natural Domain of the information one seeks minimizes the number of parameters
- **Image domain:** Natural Domain for sky-emission
  - Structure
  - Frequency and polarization dependence
- **Visibility Domain:** Natural Domain for instrumental effects
  - PB effects
  - Electronics gains, etc.

Parametrized model for sky emission

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

- The function ***I(s)*** represent sky emission
  - Information it represents is inherently in the sky domain
    - – Parametrize structure: Asp-Clean, MS-Clean
    - – Parametrize frequency dependence: MS-MFS

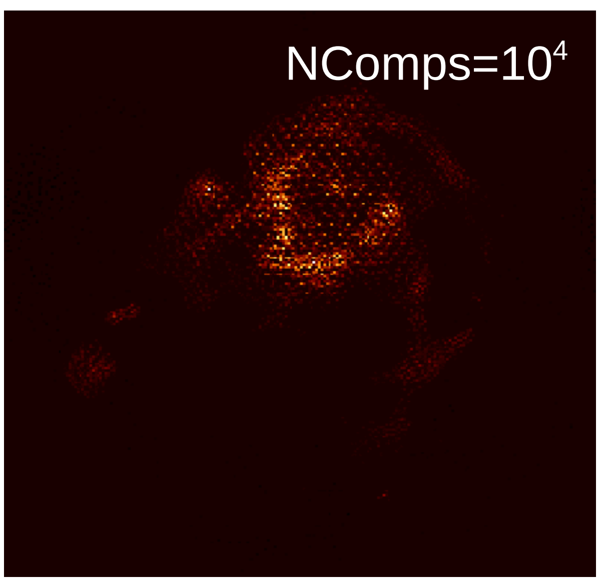


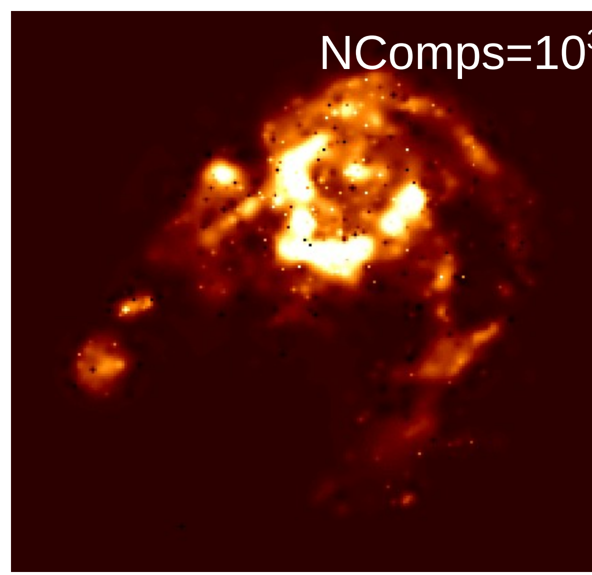


Parametrized model for sky emission

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

- The function ***I(s)*** represent sky emission
  - Information it represents is inherently in the sky domain
    - – Parametrize structure: Asp-Clean, MS-Clean
    - – Parametrize frequency dependence: MS-MFS



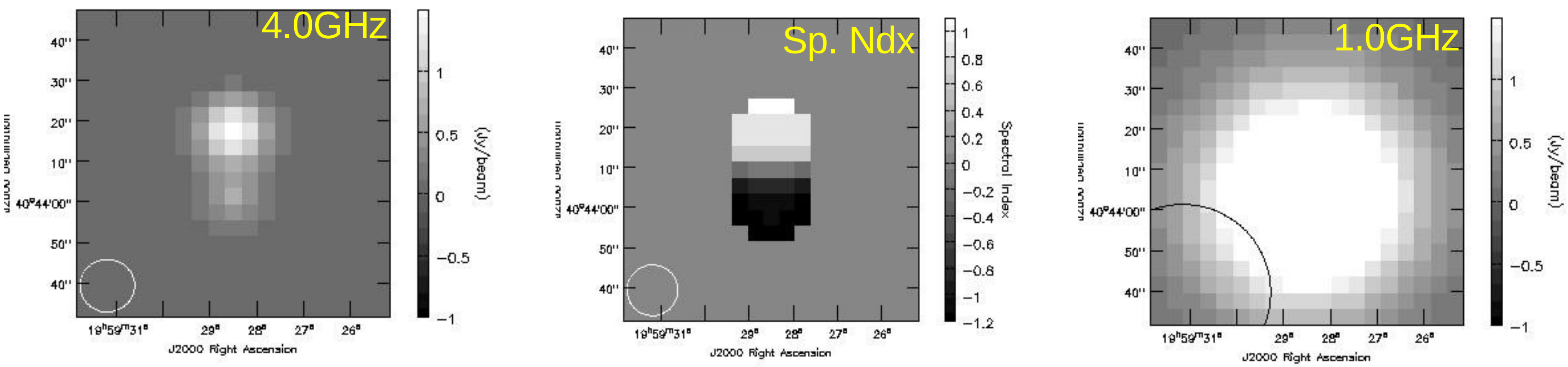


- Better parametrization in the Natural Domain

Parametrized model for sky emission

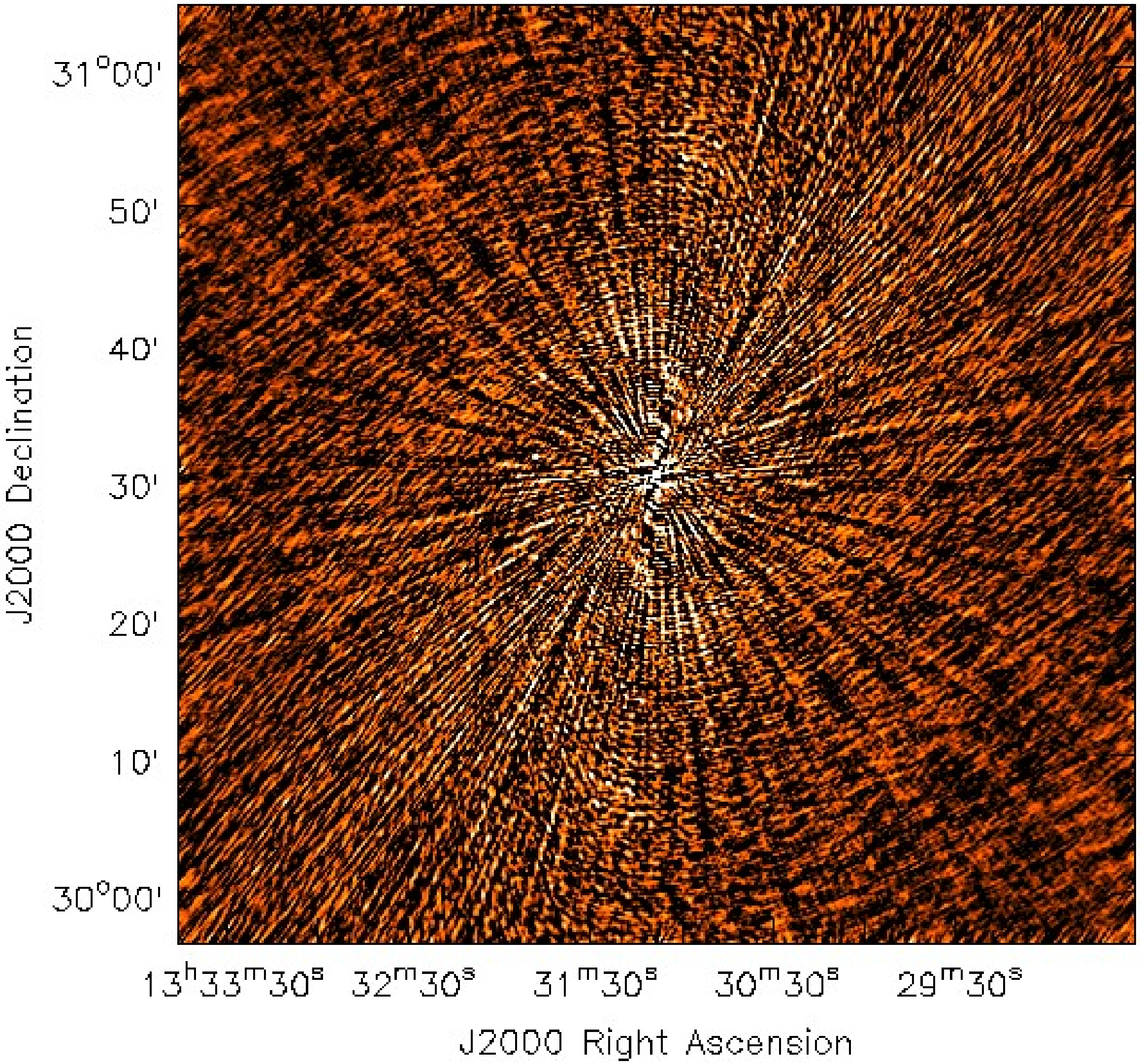
- $$V_{ij}^{Obs}(\nu) = M_{ij}(\nu, t) W_{ij} \int M_{ij}^{S}(s, \nu, t) I(s, \nu) e^{2\pi \iota (b_{ij} . s)} ds$$

- The function ***I(s)*** represent sky emission
  - Information it represents is inherently in the sky domain
    - – Parametrize structure: Asp-Clean, MS-Clean
    - – Parametrize frequency dependence: MS-MFS



Rau, PhDThesis, 2010

Wide-band imaging: Multi-Term MFS

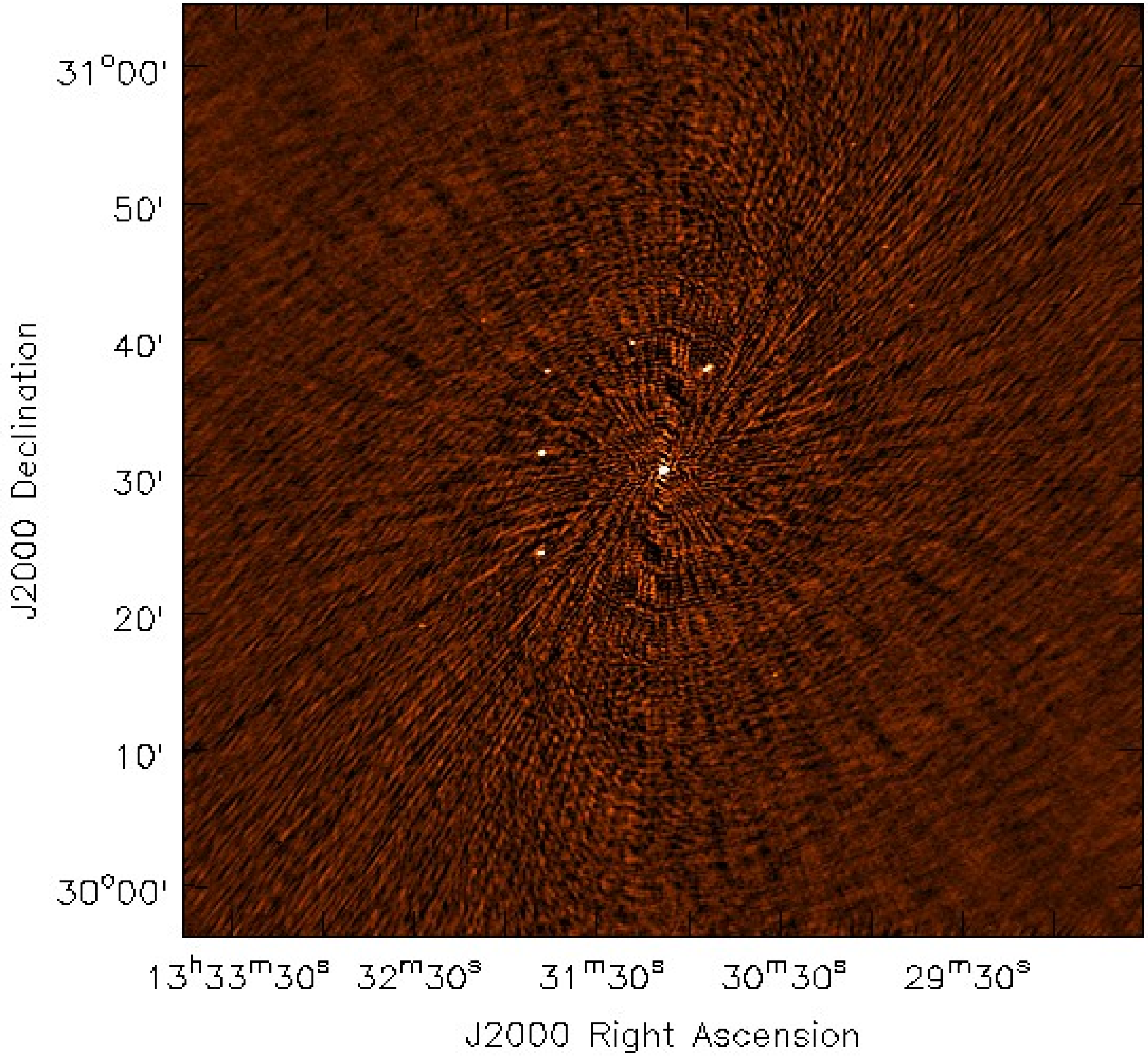


- 3C286 field
  I=14.4Jy @1.4GHz
  Sp.Ndx=-0.47
  BW = 1.1 GHz

- DR = 1600-13000

Wide-band imaging: Multi-Term MFS

MT-MFS: Collection of components whose amplitude follow a polynomial in frequency

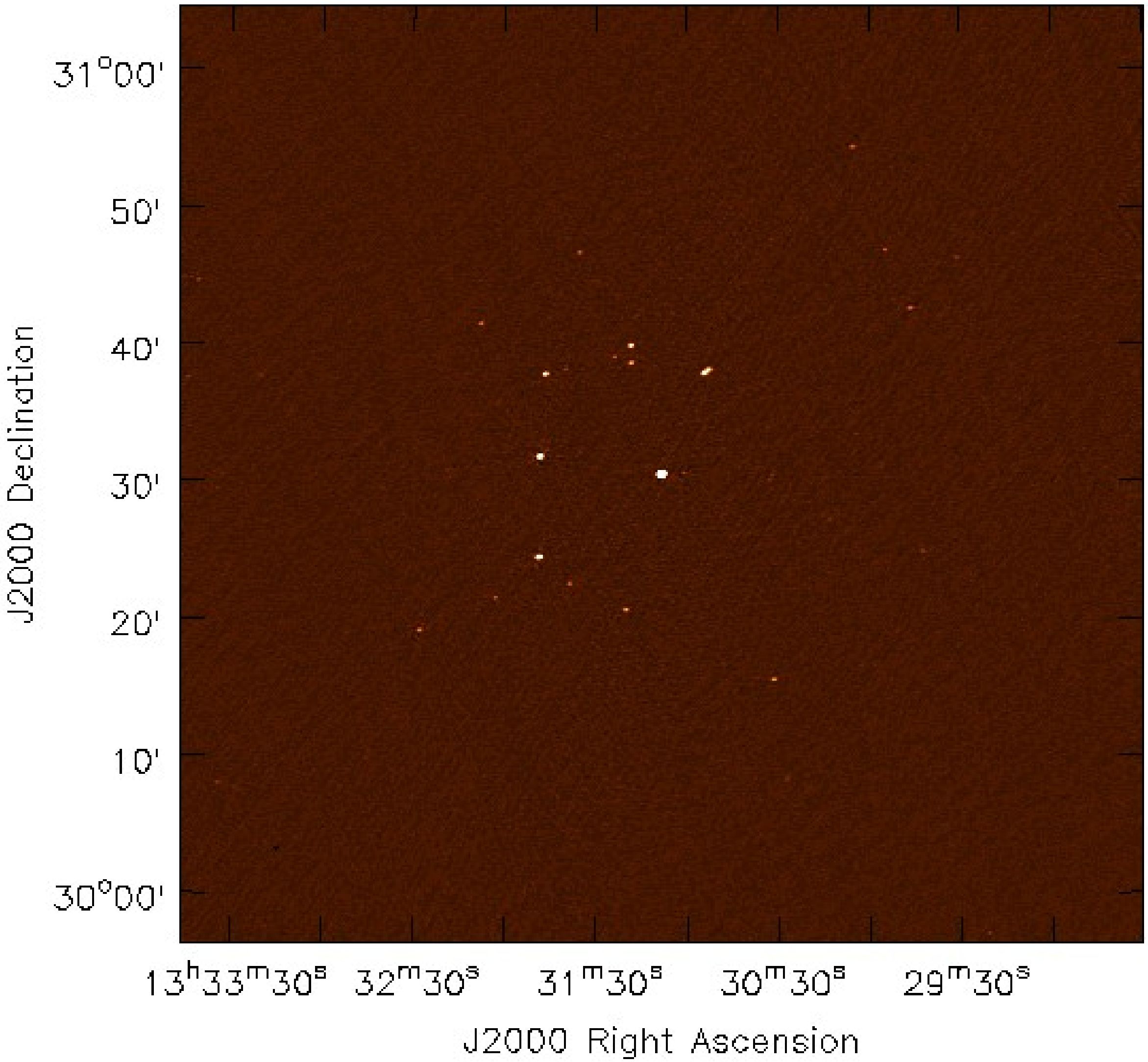


- 3C286 field
  I=14.4Jy @1.4GHz
  Sp.Ndx=-0.47
  BW = 1.1 GHz

- Multi-term MFS
- Nterm = 2

- DR = 10000-17000

Wide-band imaging: Multi-Term MFS



- 3C286 field
  I=14.4Jy @1.4GHz
  Sp.Ndx=-0.47
  BW = 1.1 GHz

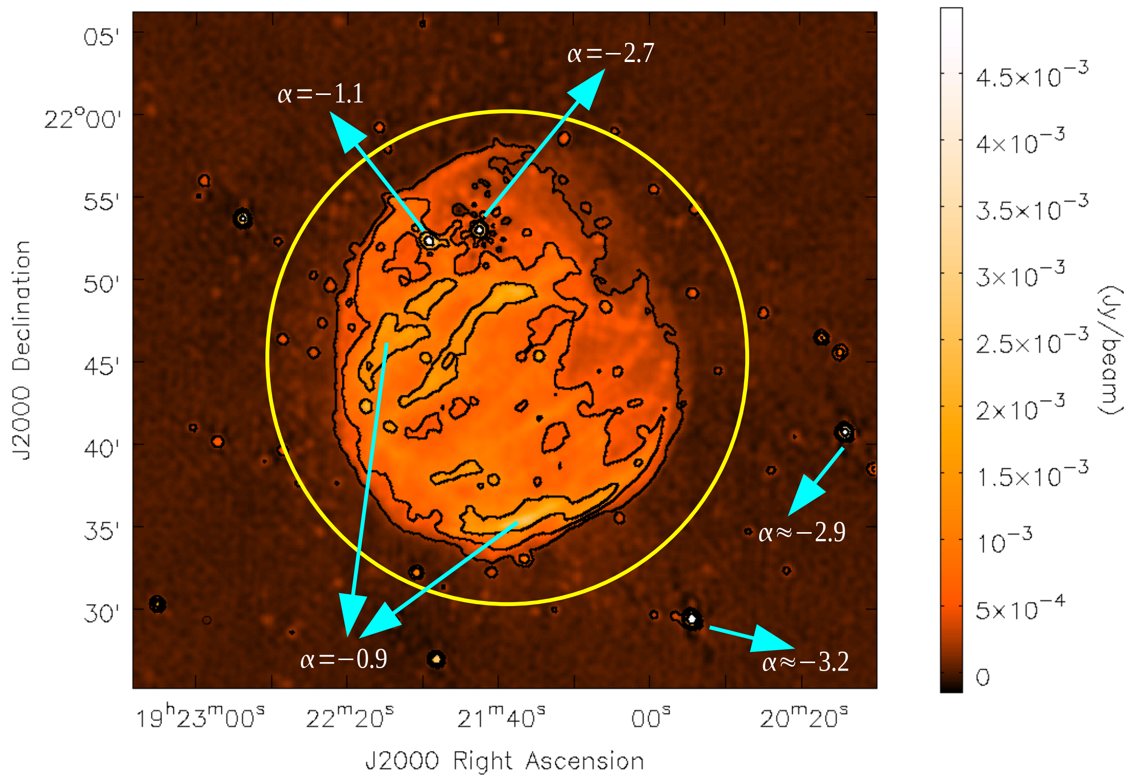
- Multi-term MFS
- Nterm = 4

•DR = 110,000-180,000

Wide-band Stokes-I imaging: MS+MT-MFS

- The sky emission varies with frequency
- Frequency dependence is also directionally dependent

$$I^D = \sum_{\nu} PSF(\nu) * \left[PB(\nu) \cdot I^{Sky}(\nu) \right]$$

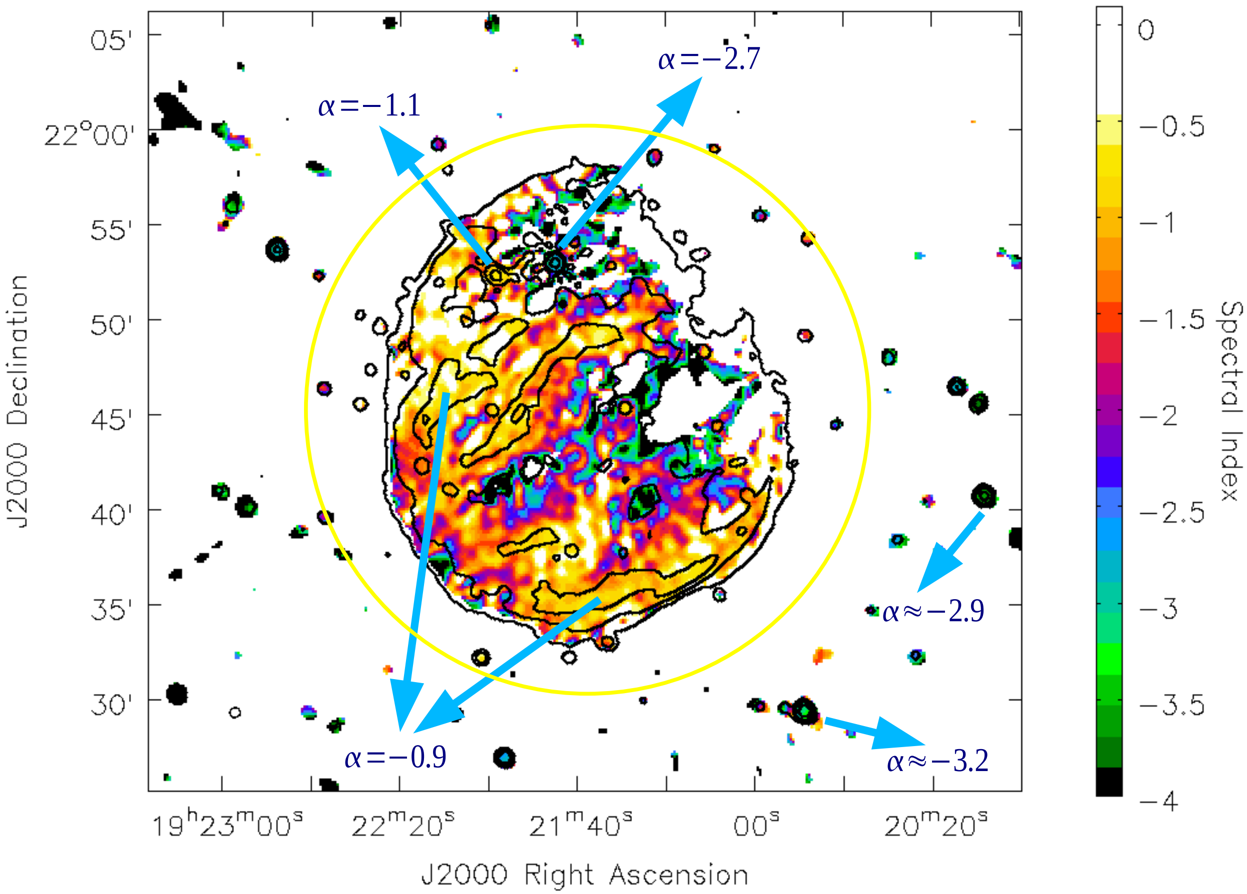


Need:
MS: For extended emission
+ MT-MFS: For DD Sp.Ndx.
+ W-Term correction
+ WB PB-correction

MS-MFS; Rau, PhDThesis, 2010

Wide-band Spectral Index Imaging: MS+MT MFS

- Spectral Index map



Need:
MS: For extended emission
+ MT-MFS: For DD Sp.Ndx.
+ W-Term correction
+ WB PB-correction

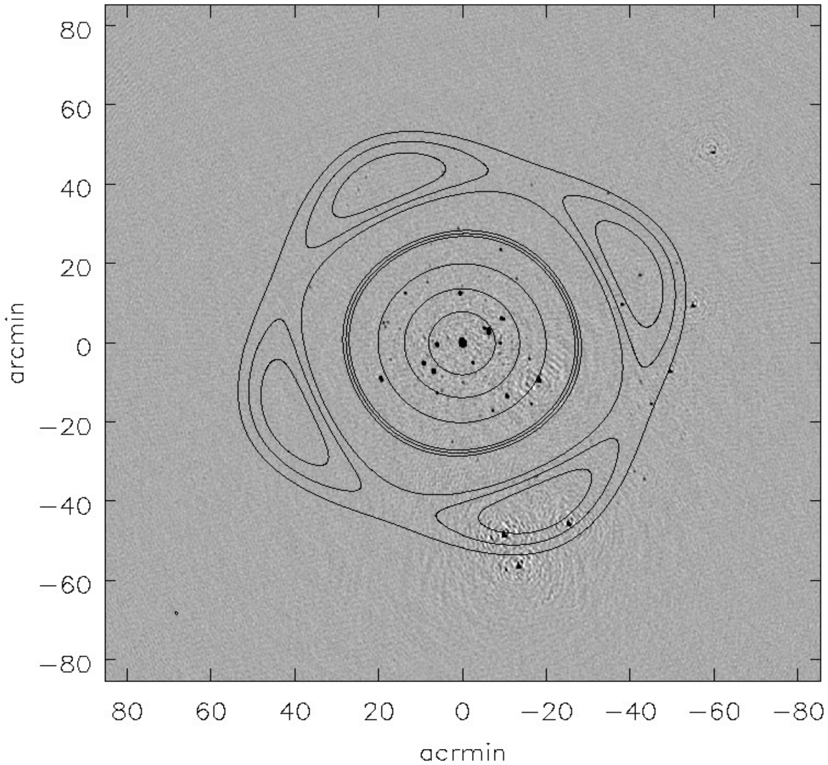
MS-MFS; Rau, PhDThesis, 2010

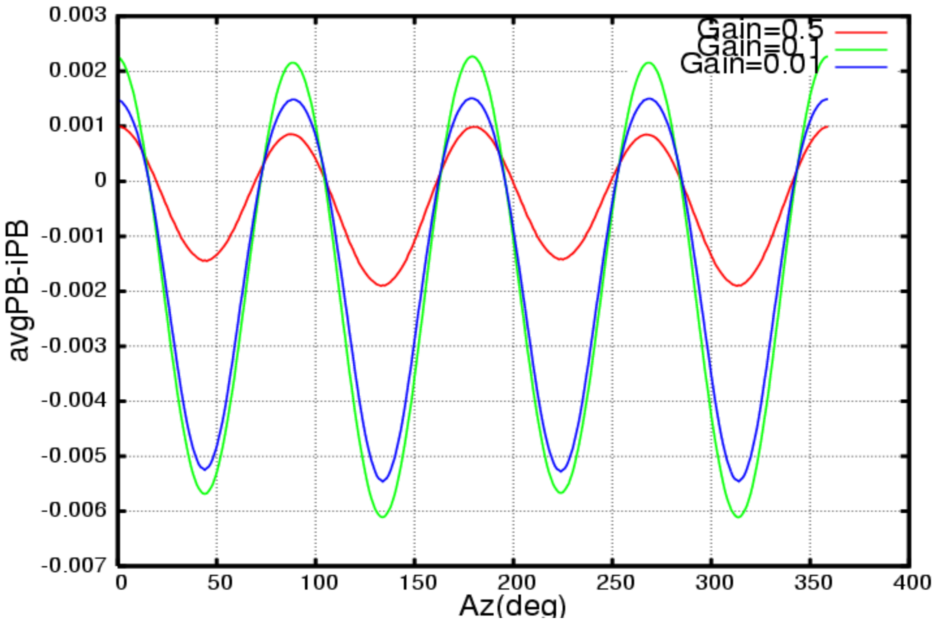
Wide-field Imaing: PB effects

- The observed data corresponds to I^{sky} multiplied by the antenna primary beam

$$I^D = \sum_t \sum_\nu PSF(\nu, t) * \left[PB(s, t) \cdot I^{Sky} \right]$$

- PB varies with time due to rotation with PA and pointing errors.
- PB gain in general is also Directionally Dependent

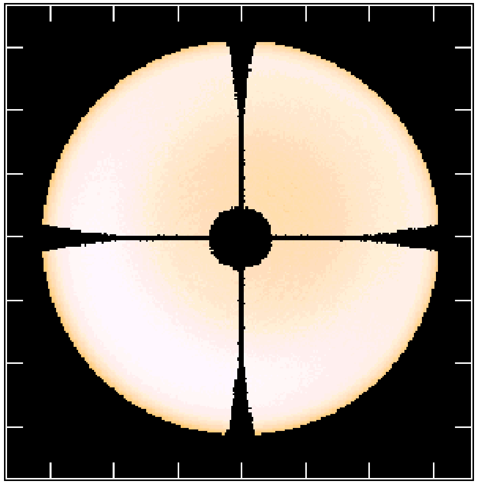




The A-Projection algorithm

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

- 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.

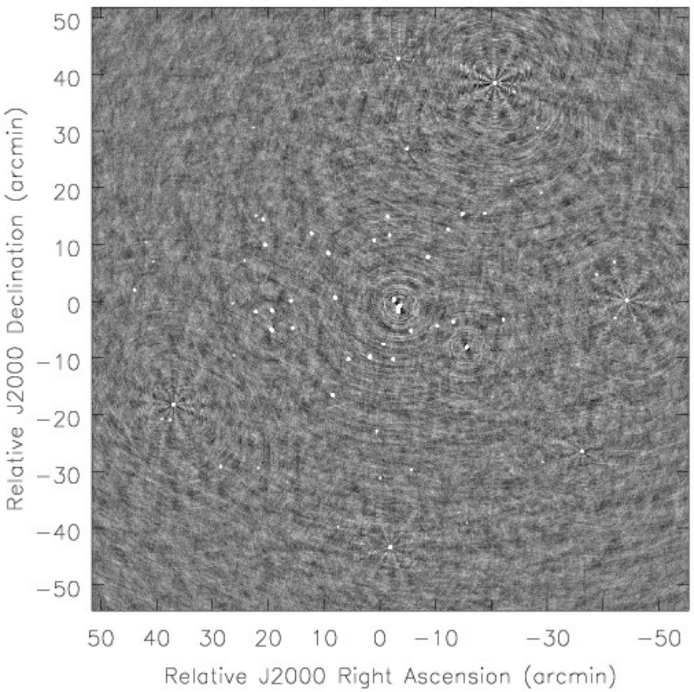
$$A(u)=A_o(u)[1+a_o Z_o(u)+...]$$

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

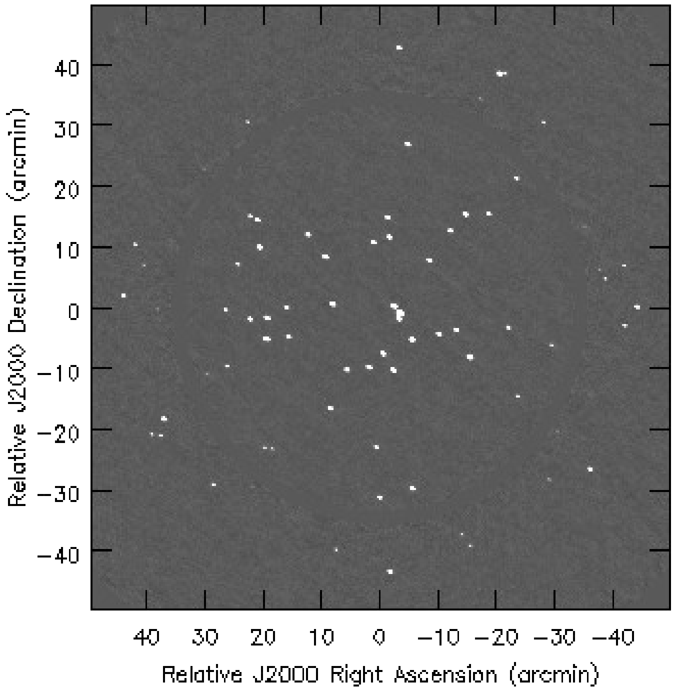
A-Projection algorithm, A&A 2008

A-Projection algorithm: Simulations

Before Correction



After Correction

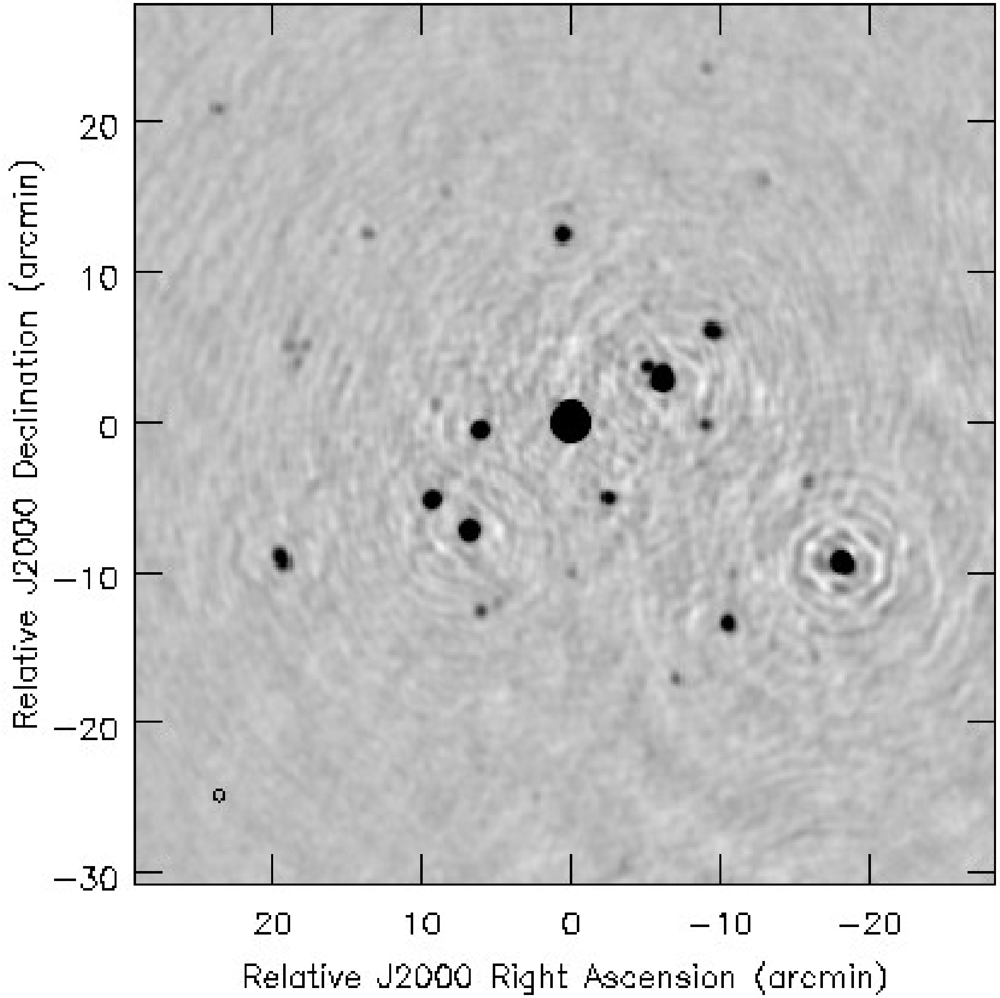


$$\textbf{Minimize}: V_{ij}^{O} - E_{ij} * [F I^{M}] \quad \textit{w.r.t.} \quad I^{M}$$

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

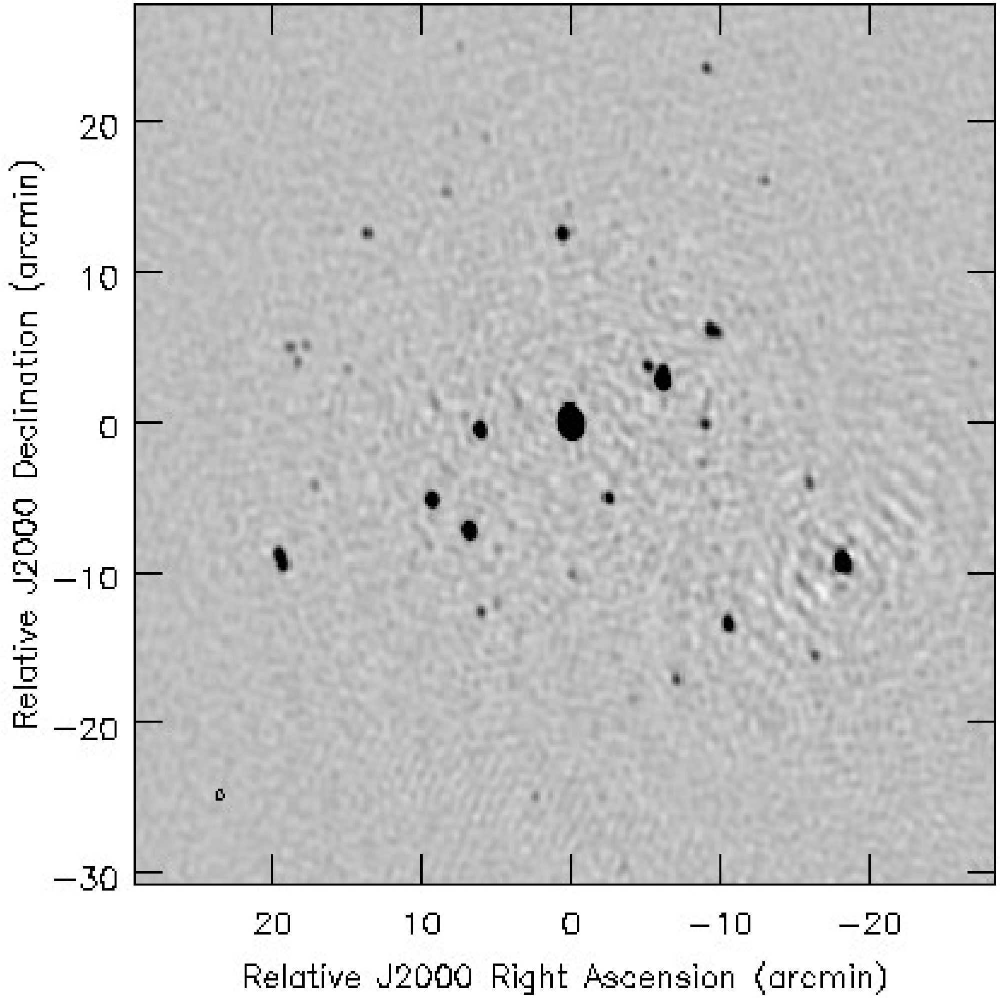
A-Projection: Bhatnagar et al., A&A,487, 2008

EVLA L-Band Stokes-I: Before correction



- 3C147 field at L-Band
- Dynamic range: ~700,000:1
- A single baseline based correction was applied

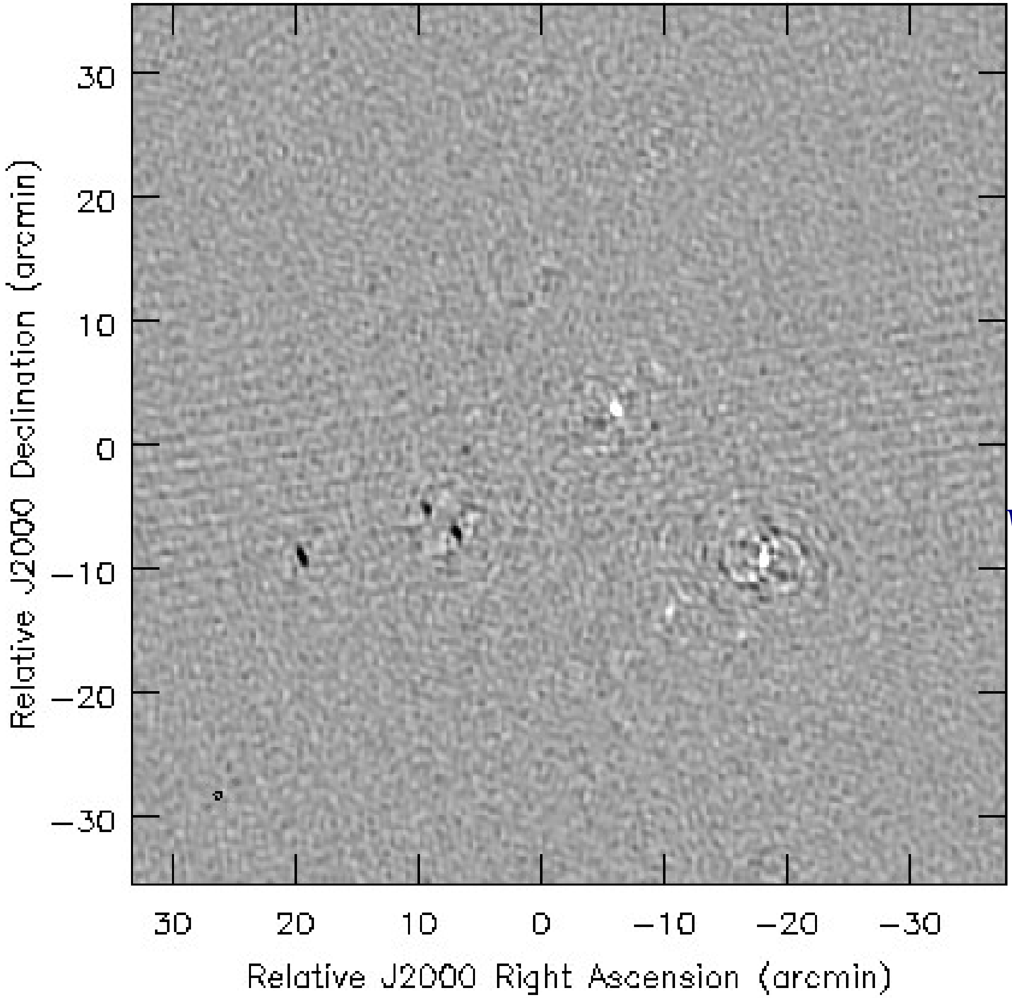
EVLA L-Band Stokes-I: After correction



- 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

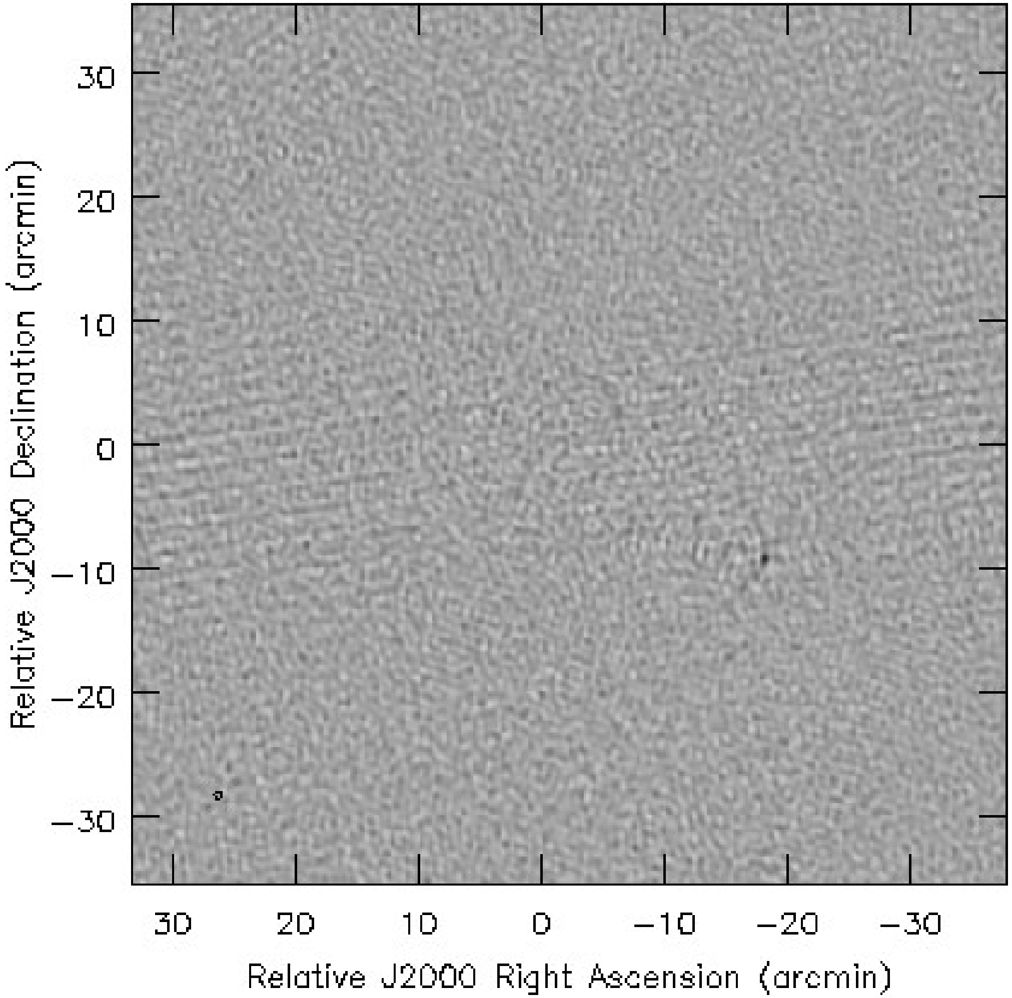


Is it $M(s, Poln)$?

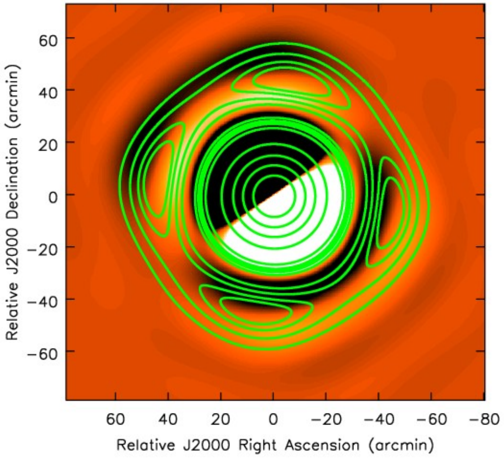
Or is it $I(s, Poln)$?

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

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

Parametrized model for aperture illumination

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

- Instrumental effects are fundamentally antenna-based
  - M^{S}_{ij} represents information inherently in the visibility domain

- **Image domain:** Only average M^{S}_{ij} is available
  - Difficult to handle the case of non-identical antennas
- **Visibility Domain:** Remains separable as antenna-based terms $$FT\left[M_{ij}^{S}\right] = FT\left[J_i\right] * FT\left[J^T\right]$$

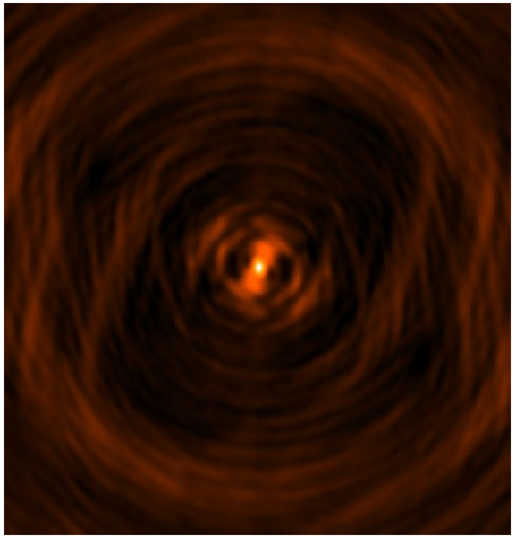
- ***Opens up algorithms for DD corrections, calibration,...***

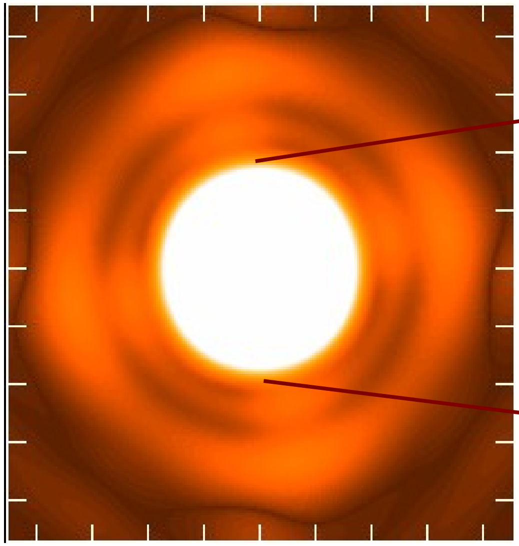
Implications for imaging: Wide-band effects

- To the first order, antenna primary beams scale with frequency
  - E.g., size of the PB changes 2x for EVLA bandwidths

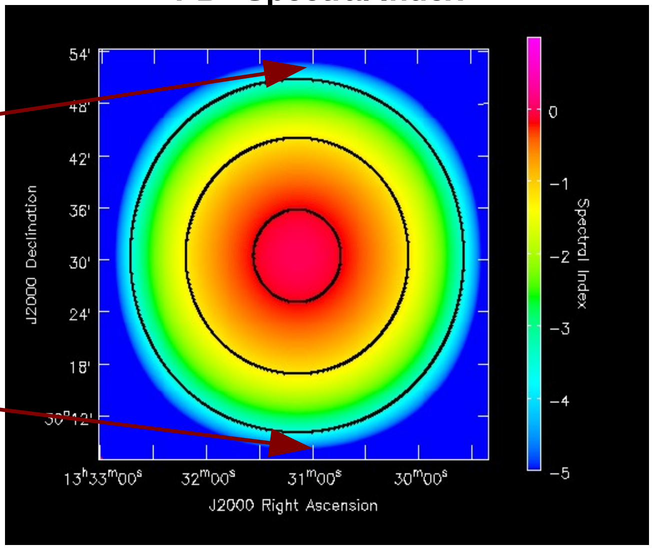
$$I^D = \sum_t \sum_\nu PSF(\nu) * \left[PB(s,t,\nu) \cdot I^{Sky} \right]$$

- PB in general is rotation asymmetric
  - Frequency dependence of the PB is also directionally dependent



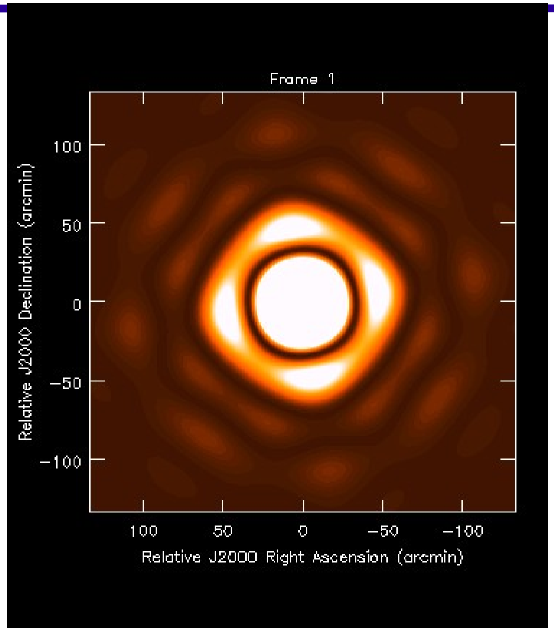


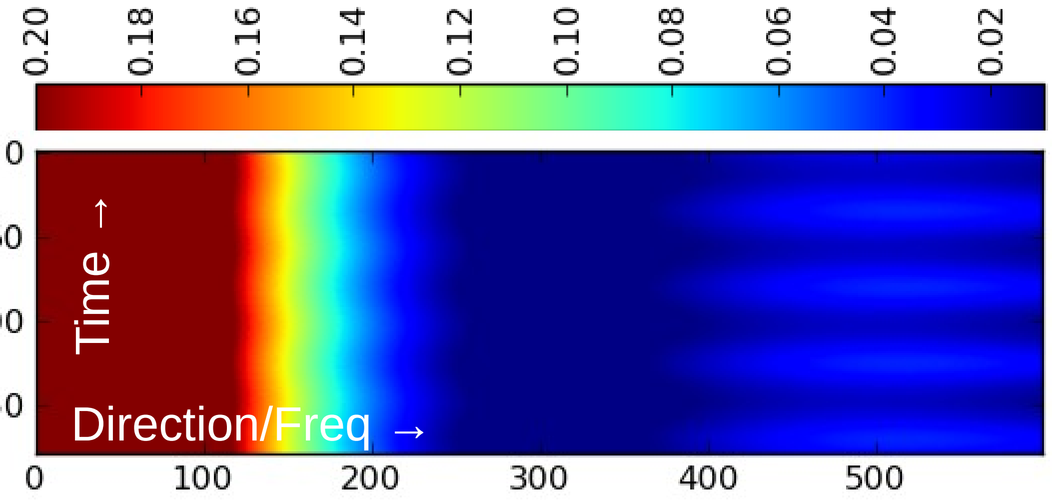
PB "Spectral Index"



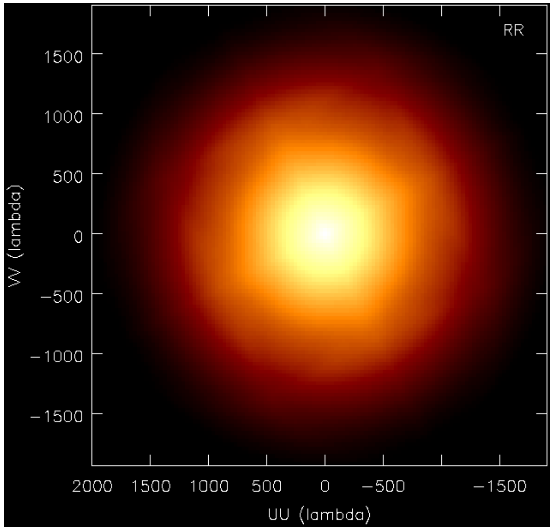
Time varying DD gains due to PB

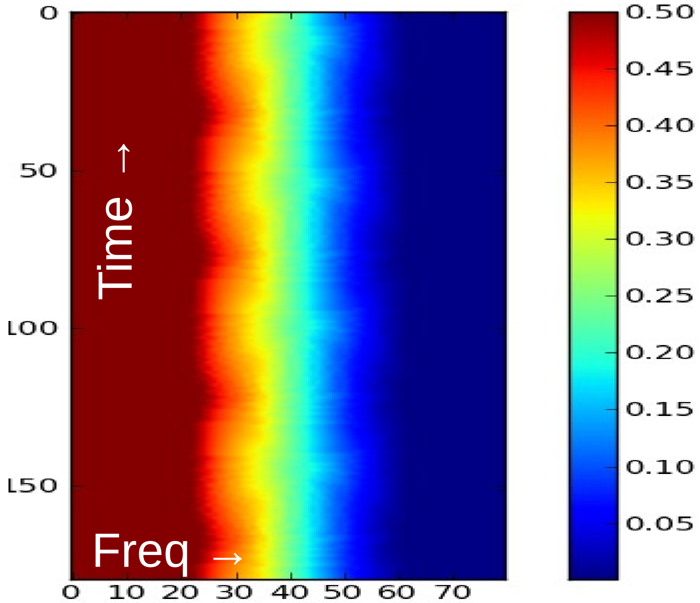
Image Plane





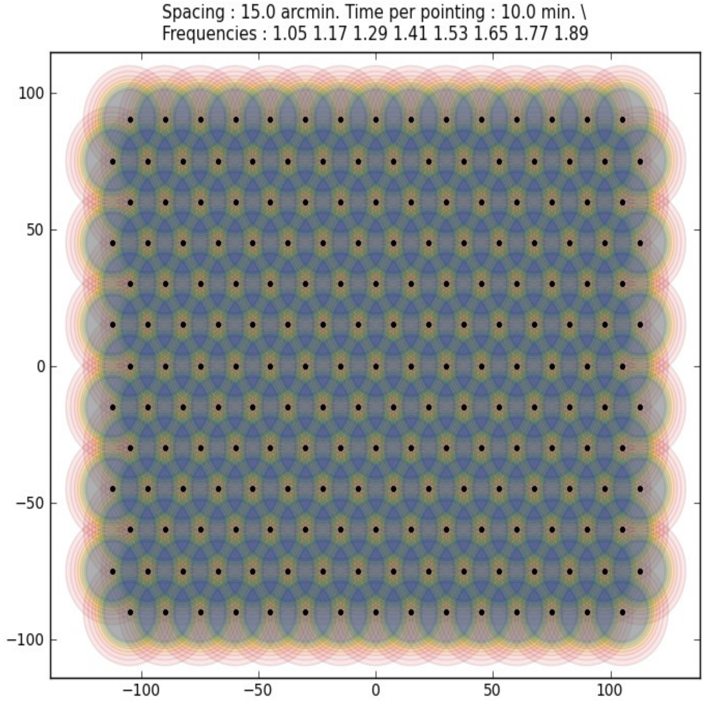
Aperture Plane





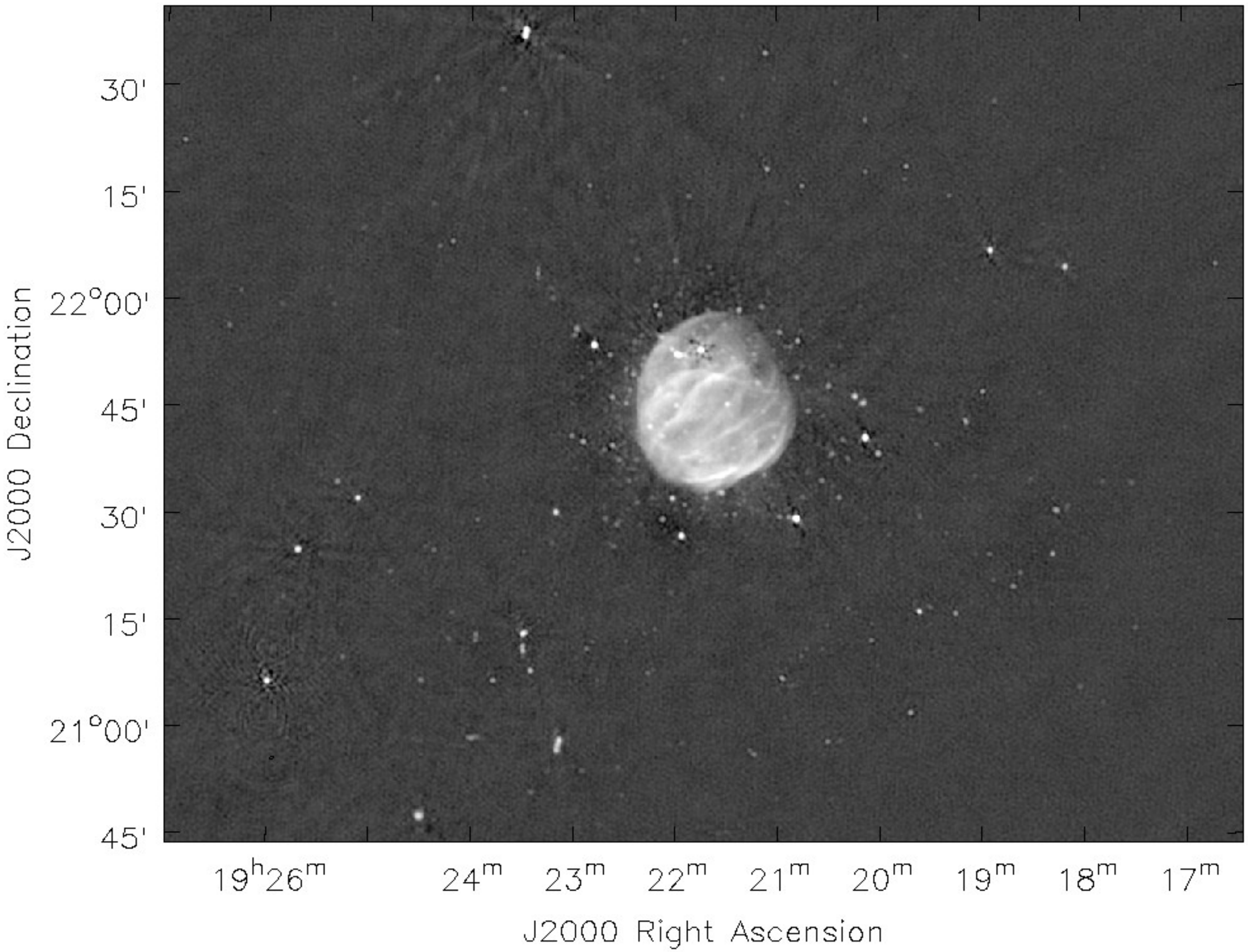
Extension to mosaicking

$$V_{ij}^{Obs}(\nu) = S_{ij}(t)\left[A_{ij}(\nu, t) * V(\nu)\right]$$



- In the data domain, PB effects correspond to convolution
  - It is included as part of the convolutional gridding operation for Projection algorithms
- Mosaicking, polarization squint, pointing errors, etc. are a matter of putting the correct phase gradient
- $A_{ij} = A_i * A_j$: The functions can be computed in a antenna dependent manner
- Naturally accounts for heterogeneous arrays (ALMA)
- DD calibration algorithms can be designed to modify A_i to fit the data (e.g. Pointing SelfCal).

Wide-field wide-band imaging with the EVLA



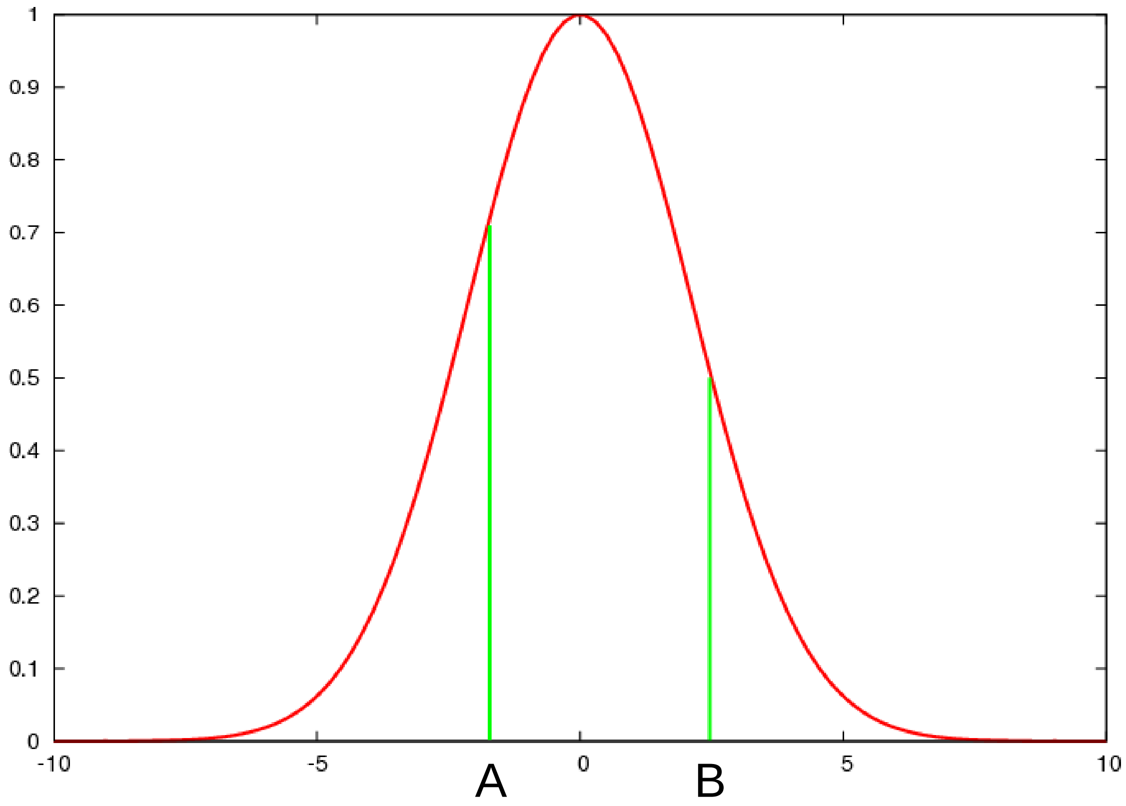
- 1.2-1.8GHz
  (4x128 MHz)
- ~25 microJy/Beam

- RSRO Projects
  (AB1345, Bhatnagar et al.)

- Scientific goals
  - Spectral Index imaging
  - RM Synthesis

  - Wide-band, wide-field imaging
  - HPC

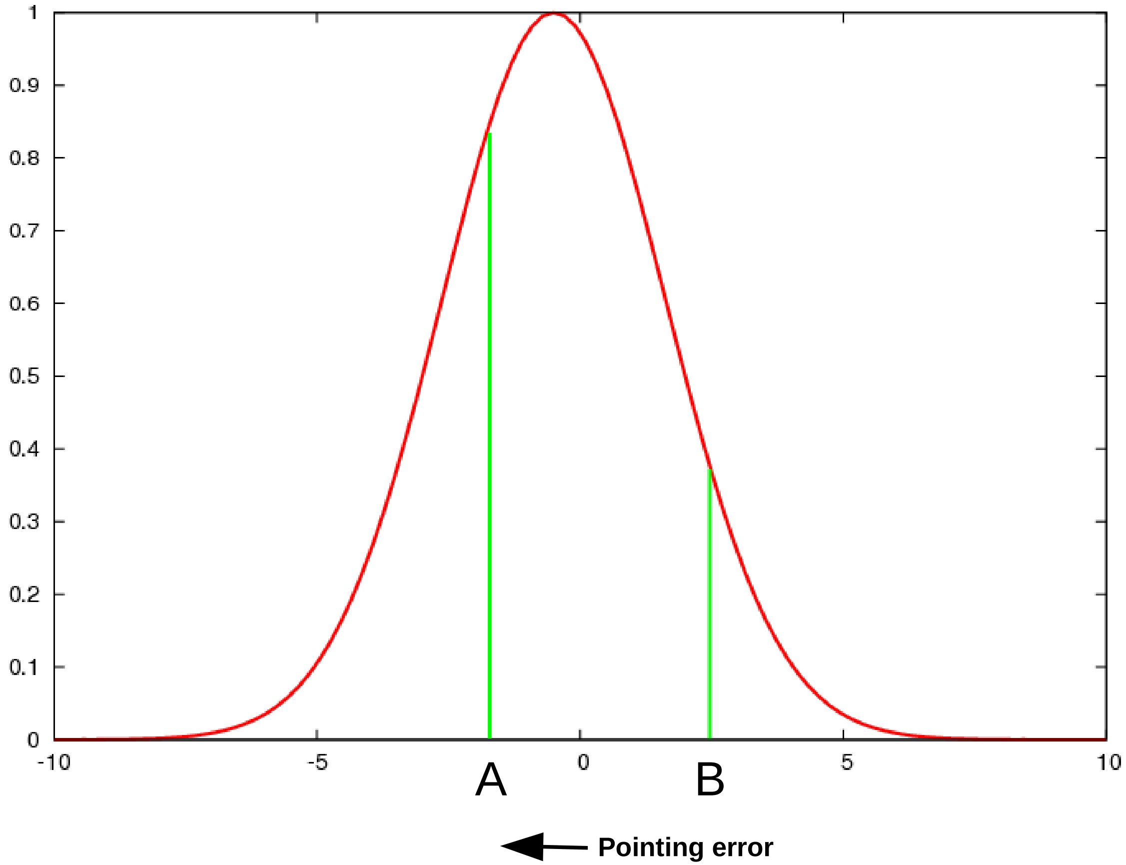
Effect of antenna pointing errors



- 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

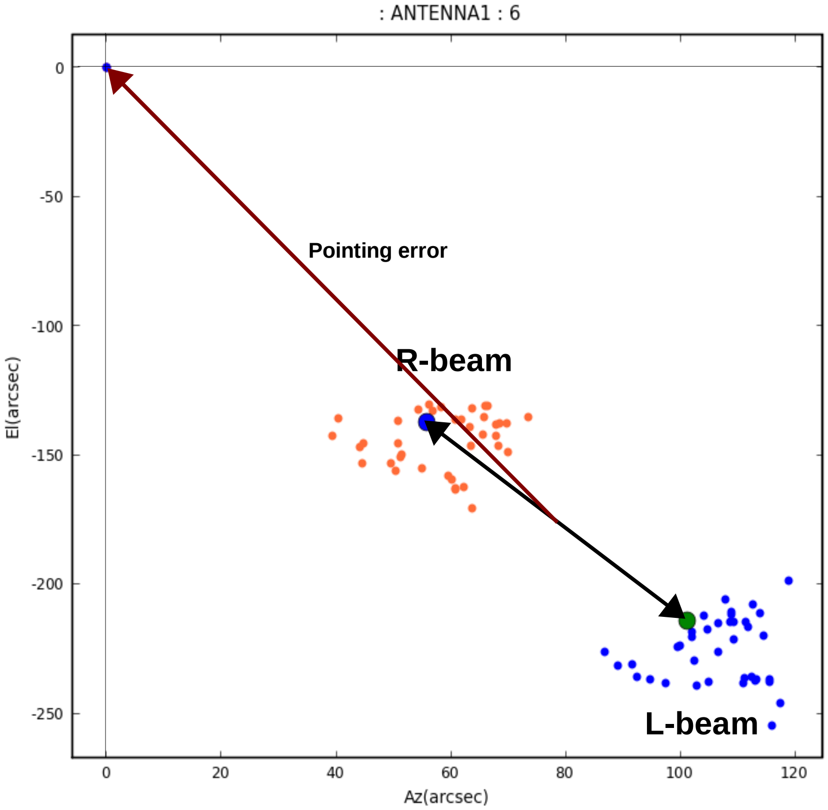
•However, there is significant in-beam phase structure –particularly for wide-field, full-Stokes imaging

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

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]

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!

I/O load

- Recent data with the EVLA: 100-500 GB

- Expect 20-50 passes through the data (flagging + calibration + imaging + human errors)
    - Effective data i/o: few TB
    - Typical disk I/O rates: 30-80 MB/s

- Exploit data parallelism
    - Distribute normal equations (SPMD paradigm)

- Deploy *computationally efficient* algorithms ('P' of SPMD) on a cluster

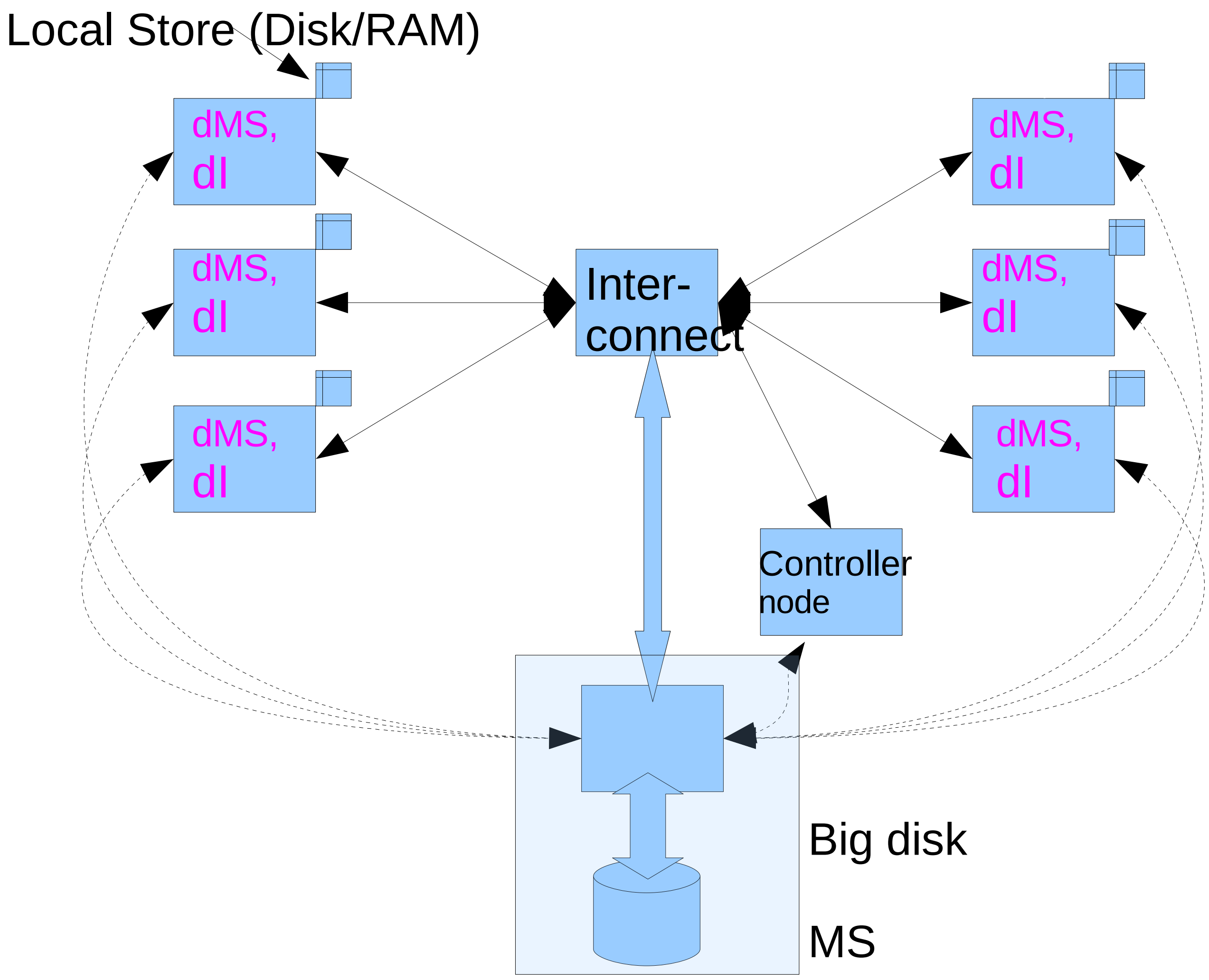
Computing load

- More data samples used for imaging
  - Few X 100-1000 frequency channels
  - 1-30 sec. Integration intervals

- More computing per gridding/de-gridding
  - Convolution support size increase for W- and A-Projection
- More images made for Multi-term MFS
  - Each term constitutes full gridding/de-gridding load

- Various optimization possible to balance between memory footprint and computing footprint
- Most operations are embarrassingly parallel

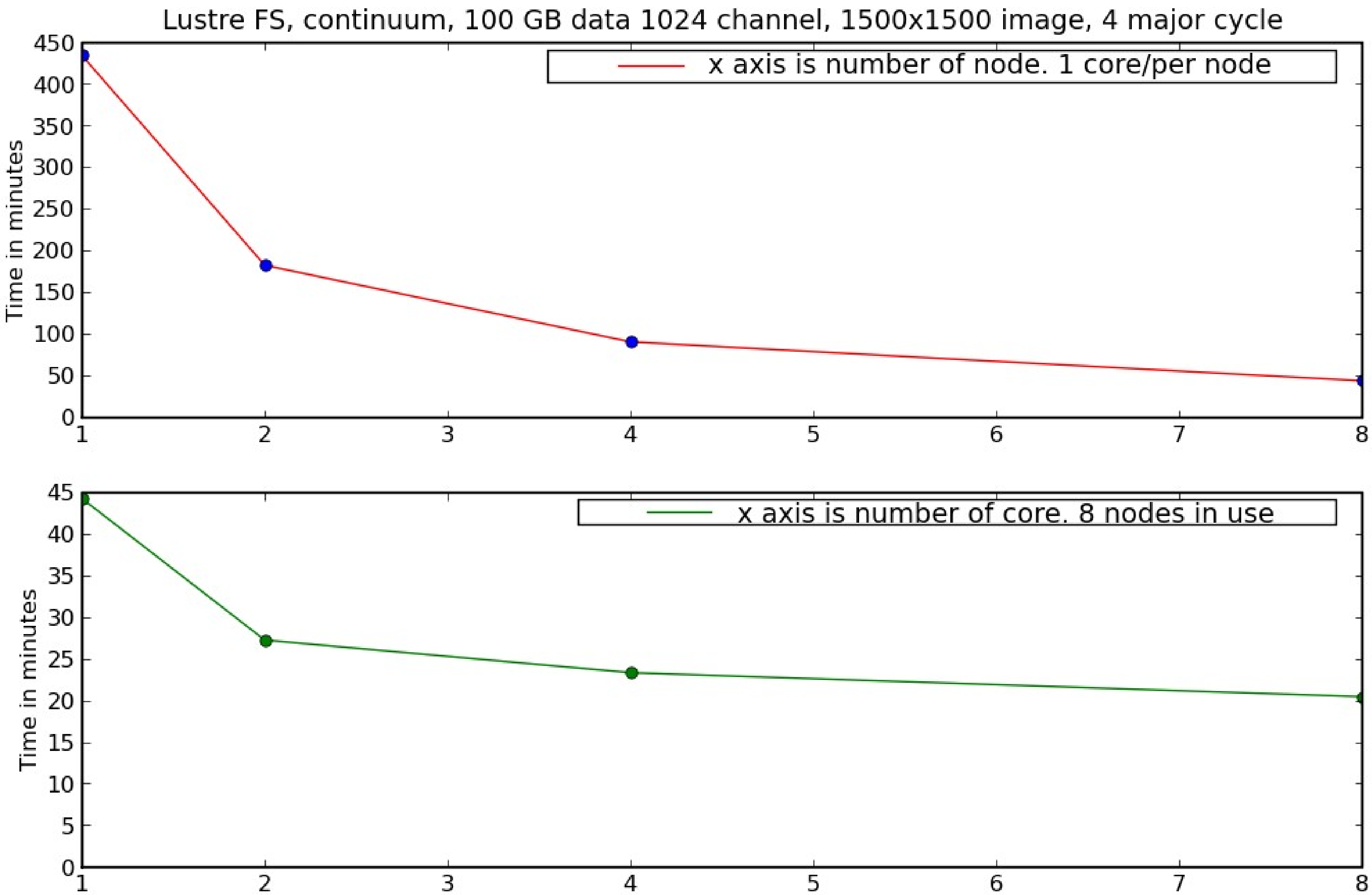
Cluster Computing



- **Goal: CPU Bound Processing**
- Gridding/de-gridding on each node using a subset of the data
- Data fed through a high speed link to fast-disk (Lustre)
  - 10 MB/s per node
- Multiple processes per node
- Asynchronous I/O
  - A single thread per node does read-ahead
  - Processing becomes CPU bound

Golap, Robnett, Jacobs, Kern

Parallelization: Initial results



Parallelization: Initial results

- Continuum imaging: (No PB-correction or MFS)
  - Requires inter-node I/O (Distribution of normal equations)
  - 4-6x speed-up using 8-cores per node
  - I/O bound without async-I/O
  - Expected close to linear speed-up with async-I/O
    - Async-I/O in the process of being deployed

- Work in progress
  - Calibration: Gain, Bandpass, polarization
  - Flagging: simple flagging + possibly auto-flagging
  - Self-Cal
  - Simple data visualization

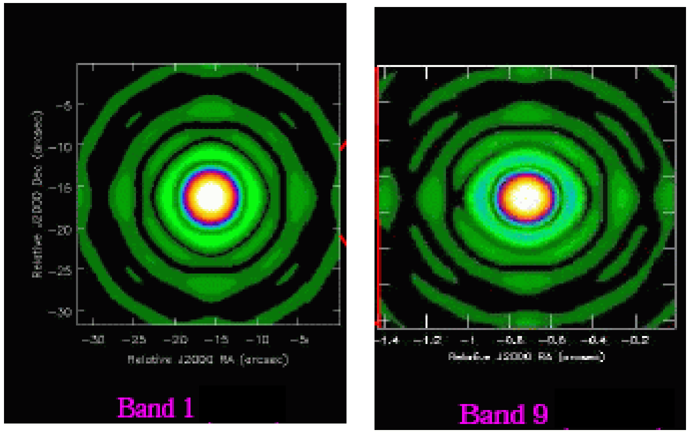
Summary

- Modeling various terms in the Natural Domain of information they represent
  - W-Term, Antenna Aperture effects in the visibility domain
    - W-Projection, WB A-Projection
    - Use in mosaicking as well
  - Extended emission, sky frequency- and polarization-dependence in the image domain
    - MS-MFS, Asp, …
    - RM-Synthesis
  - Developing DD solvers to solve for for low-order models for Aperture illumination
    - Pointing SelfCal and beyond

What keeps us busy?

- Shooting for full-sensitivity full-polarization wide-band imaging

- Bug discovery and fixes (many thanks to FO, CC, JM-J, EF, JU,...)
- Re-worked the code to enable
  - Wide-band A-Projection
  - Heterogeneous arrays (ALMA)



D. Petry, ESO

- Next steps:
  - WB A-Projection + W-Projection
  - Integrate MS-MFS and WB A-Projection
  - Extend to mosaic imaging
  - Extend to full-polarization
  - Integrate with RM-synthesis

What keeps us busy?

- Projects in various stages of R&D
  - Automatic RFI detection/removal

  - Pointing SelfCal
    - An issues for ALMA and mosaicking in general
  - Asp-Clean based MFS, RM-Synthesis(?)
    1. Memory foot-print
    2. Reduce error bars on Spectral Index images

- Integrate with parallel computing framework for deployment on HPC platforms
- Multi-threading where possible (e.g. minor cycle)
- Develop pipelines or integrate with existing pipeline processing framework

Thanks

- Various testers of the bleeding-edge code
- Various members of the EVLA commissioning team
- Computing Staff
- CASA team
- Various people whose brain I often borrow...

Era of Data Deluge

- "My" reaction to Data Deluge skeptics
  - Beginning of telephone era: People reported shock lasting days after a phone call
- Much opposition is simply romantic

- **Mathematics Tells us**
  - Information technology is not magic
  - Extracting information from data is not a sure thing
  - Specific hard work on a case-by-case basis
  - **You can** learn what must be done

"Data! Data! Data! Challenges and opportunities of the coming Data Deluge"
- David Dohono, Stanford Univ.
USNA Michelson Lecture, 2001