Wide-field Imaging

Sept. 16th, 2011

S. Bhatnagar
Plan

- What do we mean by wide-field?
- Projection algorithms to correct for various wide-field effects
  - Relation with minor cycle algorithms
- Algorithms “unification scheme” :-)  
  - Similarity between various wide-field algorithms
- Algorithms
  - For W-term correction  
    - W-Projection, Multi facet Imaging  
  - For PB corrections  
    - A-Projection: Low and high frequency  
    - AW-Projection at low frequency bands

- Connection with Mosaicking:
  - Generalization of single pointing
What do we call Wide-field?

- Imaging that requires invoking any of the following:
  - Corrections for non co-planar baseline effects
  - Corrections for the rotational asymmetry of the PB
    - Imaging beyond 50% point, mosaicking
  - Corrections for the frequency or polarization dependent effects
    - PB, ionosphere/atmosphere

- Noise limited imaging at “low” bands (L, S and probably C Band)
  - Because of the radio brightness distribution

- Noise limited imaging of structure comparable to the PB beam-width

\[ I_{\text{Continuum}} = \int PB(\nu) \left[ I_0(\nu/\nu_o)^{\alpha(\nu)} \right] d\nu dt = \int I_0(\nu/\nu_o)^{\alpha_{pb}(\nu,t)+\alpha(\nu)} d\nu dt \]

- Mosaicking
  - By definition, imaging on scales larger than the PB beam-width
Why wide-field?

- Primarily due to improved continuum sensitivity
  
- E.g. a 1% PSF side lobe due to a source away from the center is now significantly above continuum thermal noise limit
  - This is largely independent of the total integration time

- Due to large bandwidth, EVLA is sensitive farther out in the FoV

- E.g. @L-Band, PB gain ~1 deg. away can be up to 10%
  - In the EVLA sensitivity pattern, VLA sensitivity is achieved at the location of VLA-null!
  - No null in the EVLA sensitivity pattern
Wide-field Issues

- For the same integration time, EVLA is sensitive to emission farther out

\[ \Delta S = S(R) \times PB(R) \times PSF(R) \]

- R = 1°, S(R)=1Jy, \( \Delta S = 1\, mJy - 100\, \mu Jy \)

Wide-field Issues
Effects of the W-Term

![Diagram showing effects of the W-Term with two images comparing J2000 Right Ascension and Declination. The left image shows a complex pattern, while the right image shows a simpler distribution of points.]
Non co-planar baseline: The W-term

- 2D FT approximation of the Measurement Equation breaks down
  - \( \frac{\lambda}{B_{\text{max}}} \leq \theta_f^2 \) \( \theta_f = \text{Angular distance from the phase center} \)

- We measure:
  \[ V_{12}^o = \langle E_1'(u,v,w \neq 0) E_2^*(0,0,0) \rangle \]

- We interpret it as:
  \[ V_{12} = \langle E_1(u,v,w=0) E_2^*(0,0,0) \rangle \]

- We should interpret \( E_1 \) as \( [E_1' \times \text{Fresnel Propagator}] \)
PB Effects
PB Effects: Rotation asymmetry

- Only average quantities available in the image domain
- Asymmetric PB rotation leads to time and direction dependent gains

\[ \Delta I^R = \sum_\psi \left[ PSF(\psi) - \text{avgPSF} \right] \ast \left[ (PB(\psi) - \text{avgPB}) I^0 \right] \]
PB Effects: Error Propagation

\[ \Delta I^R = \sum_\psi \delta \text{PSF}(\psi) \times [\delta \text{PB}(\psi) I^o] \]

\[ \delta \text{PB}(\psi) \quad \text{E.g. } 5 \times 10^{-3} \]

\[ \delta \text{PSF}(\psi) \quad \text{E.g. } 2 \times 10^{-2} \]
Frequency dependence of the PB

- Assume linear scaling with the frequency
Algorithms: CS Clean recap

- Compute residual using original data
  - Needs Gridding and de-Gridding during major-cycle iterations

- Most commonly used algorithm
- Every major cycle access the entire data base
  - Significant increase in I/O and computing load
- Assumes, co-planar, time- and freq-independent Measurement Equation

Cannot account for wide-field wide-band and time variability issues
Deconvolution as ChiSq Minimization

- \[ V^M = A I^M + AN \]
- \[ V_{ij} = \text{deGrid}_{ij} \text{FT}(I) \]

- Non-linear solver, to solve for the Model Image
  - Compute residuals:
    \[ V^{\text{Obs}} - AI^M \quad \text{(data domain)} \]
    \[ I^d - BI^M \quad \text{(image domain)} \]
  - Make Residual Image \( I_{\text{res}} \)
  - Find update direction: Steepest Descent Algorithm
    \[ I^c = \max \left( -2[I^{\text{Res}}] \frac{\partial \chi^2}{\partial \text{Param}} \right) \]
  - Update model:
    \[ I^M_i = T(I^M_{i-1}) \quad \text{for our discussions this is} = I^M_{i-1} + \alpha * I^c_i \]

- Since Major Cycle does model subtraction without averaging, variable terms can be included in that step
Algorithms “unification scheme”

- Incorporates direction dependent effects as part of the gridding function
  - ME: \( V_{ij} = A_{ij}I^o + N_{ij} \)
  - Construct \( D \), such that \( \frac{D_{ij}^T A_{ij}}{D_{ij}^T D_{ij}} \approx 1 \)
  - Compute residuals (major cycle): \( D_{ij} \) for forward and \( D_{ij}^T \) for reverse transform

- W- and A-Projection construct \( D \) differently
  - A-Projection has additional normalization issues:
    - Flat-noise vs. flat-sky normalization

- Mosaicking: (more in K. Golap's talk in Thursday Lecture Series)
  - The Fourier transform shift theorem
    \[
    I^{Mosaic} = \sum_k I(l_o - l_k)
    \]
    Use \( D_{ij} e^{i[(l_o - l_k) \cdot u_{ij}]} \) where \( D_{ij} \) can be \( A_{ij}, W \), or \( A_{ij} \ast W \)
    - The Fourier transform shift theorem

[https://safe.nrao.edu/wiki/pub/Software/Algorithms/WebHome/Mosaicking_aoc.pdf]
**Algorithms “unification scheme”**

- “Single polarization” case: Single element of the Mueller Matrix

- Imaging

\[ V^{Grid} = CF \ast V^{obs} \]

\[ I' = FFT \left[ V^{Grid} \right] \]

- Prediction (de-gridding):

\[ V^{Grid} = FFT^{-1} \left[ I^M' \right] \]

\[ V^M = CF^T \ast V^{Grid} \]

- CF can be A-term, W-term, AW-term, wide- or narrow-band
Projection algorithms

- Direction-dependent ("image plane") effects as convolutional terms in the visibility domain
- ME entirely in the visibility domain: $V_{ij}^O = A_{ij} I^M = M_{ij} F I^M = M_{ij} [V^M]$

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

- Diagonal: "pure" poln. products
- Off-diagonal: Include poln. leakage

$M_{pq} = J_{p,i} \ast J^*_{q,j}$

- $V_{pp}^O = M_{pp} \ast V_{pp}^M + M_{p \ p2q} \ast V_{pq}^M + M_{q \ p2q} \ast V_{qp}^M + M_{p2q \ p2q} \ast V_{qq}^M$

Generalization of the direction-independent ME
- Replace functions by complex numbers $M_{ij} = g_i g_j^*$
- Replace convolution (\(\ast\)) by complex product
W-Projection

- **W-Projection:** (CASA Imager: ftmacine=”wproject”)

\[ D = FT\left[ e^{2\pi i \sqrt{w-1}} \right] \]

- **\( D^T A = 1 \)** by construction
  - Potentially fully corrects for the effects of the W-term
  - In practice, D is computed at a finite \( w \)-resolution, with interpolation in between

- **D is non-hermitian**
  - Post deconvolution correction is not possible
  - Same as: “corrections for antenna based phase errors cannot be corrected for post-deconvolution”

[Cornwell, Golap & Bhatnagar, 2008]
W-Projection + Multi-faceting

- Multi-facet imaging (CASA Imager: facets > 1)

- Split the sky into multiple, smaller tangent-plane images

- A linear approximation of this image-plane operation is possible in the visibility plane: \( I(\mathbf{C} \mathbf{l}) \rightarrow |\text{det}(\mathbf{C})|^{-1} \mathbf{V}(\mathbf{C}^{-1} \mathbf{u}) \)
  - Advantage: leads to a single combined image in the minor cycle

- Combination of W-Projection and Multi-facet imaging possible:
  - Reduces the no. of w-planes and number of facets
A-Projection

- **A-Projection**: $D = \text{Auto-correlation of Aperture illumination function}$

- Function of time, frequency and polarization

- Since image is averaged over time and frequency, time- and frequency-dependence cannot be corrected for post-deconvolution
  - Same issue as non Hermitian nature of antenna based phase, W-term

Bhatnagar, Cornwell, Golap & Uson, 2008
A-Projection: Stokes-I Before
A-Projection: Stokes-I After
A-Projection: Stokes-V Before
A-Projection: Stokes-V After
Imaging at high frequencies

- Definition: Frequency at which the array is co-planar for the required FoV

- To the first order, aperture illumination may linearly scale with frequency (or at least with in a certain range in frequency)

- Wavelength much smaller than the physical reflecting structures
  - Geometrical ray-tracing models might be sufficient

- Can be computed once per SPW, rotated in time, and scale in frequency during imaging
  - Significantly reduces memory footprint, at the cost of computing
  - Can be computed efficiently on GPUs
Imaging at low frequencies

- Definition: Frequency at which the array is non co-planar for the required FoV

- PB variations with time
- Even D-array is non co-planar
- BW ~400 MHz

Need: Wide-band AW-Projection
**Wide-band AW-Projection**

- \( D(\nu) \neq (A*W)(\nu/\nu_o) \)

- Full-polarization case requires:

  - Can be configured for optimal usage for:
    - High frequency: A-Projection, scaling with frequency
    - Low frequency: AW-Projection
    - Heterogeneous array
Physics of “unification”

- Physics of DD terms go into the construction of D
- Multiple DD terms become “convolution of convolution functions”
  \[ W \ast \text{<convolved}> \ast A \]

- E.g. form of the phase of A-term accounts for mosaicking, pointing corrections, etc.
- Wide-band, full-pol., low-freq. Mosaic can be done naturally
  - Complexity goes in the construction of the CFs
  - Rest of the imaging / calibration framework remains oblivious