EVLA Memo #100
Correction of errors due to antenna power patterns during imaging

S.Bhatnagar, T.J.Cornwell* and K.Golap
NRAO, Socorro

February 23, 2006

Abstract

Errors in the observed visibilities can be classified as direction dependent and direction independent effects. Since direction dependent effects are different for each source, they must be corrected during imaging. This memo is concerned with the problem of correcting for such effects during imaging. We demonstrate that the effects of known receptor pointing errors (including polarization squint) can be corrected during imaging. An algorithm to solve for the antenna pointing errors is described in EVLA Memo 84. Errors due to such effects as well other time varying direction dependent effects are expected to limit the highest imaging dynamic range achievable for single pointing observations with the EVLA and mosaicking observations with ALMA. Error analysis and corresponding impact on computing load for wide-band and mosaic imaging is also discussed.

1 Introduction

Measurement Equation for an interferometer can be compactly written using the HBS notation (Hamaker et al. 1996) as

$$V_{ij}^{Obs} = W_{ij}\mathsf{J}_{ij}\int \mathsf{J}_{ij}^{Sky}(\boldsymbol{S})I(\boldsymbol{S})e^{\iota\boldsymbol{S}.\boldsymbol{B_{ij}}}\,d\boldsymbol{S} \tag{1}$$

where V_{ij}^{Obs} is the observed visibility Stokes vectors, J_{ij} is the Muller Matrix composed of multiplicative antenna based gains, $\mathsf{J}_{ij}^{Sky}(\boldsymbol{S})$ is the Muller Matrix of direction dependent gains as a function of direction $\boldsymbol{S}$, W_{ij} is the weight for baseline $i-j$ and I is the image. When J_{ij}^{Sky} is unity, the above equation reduces to the familiar

*Now at The Australia Telescope National Facility, Sydney

equation $V_{ij}^{Obs} = \mathsf{J}_{ij} V_{ij}^{M}$, where V_{ij}^{M} is the model visibility vector. The effects of J_{ij} can be removed prior to imaging by multiplying this equation by J_{ij}^{-1} and using the corrected visibility for making the image. The effects of J^{Sky} however cannot be similarly removed since it is an image plane effect and hence must be corrected for during imaging. Direction dependent effects which are antenna independent (and hence same for all baselines) can be removed by dividing the image by J^{Sky}. Non-squinted antenna power pattern is one such example. For El-Az mounted antennas with a squint in the power patterns of the two orthogonal polarizations, image plane polarization properties not only change across the field of view but also rotate on the sky as a function of parallactic angle (PA). As long as the squint is same across the antennas, such errors can still be corrected, to a large extent, by correcting individual snapshot dirty images for each polarization (or dirty images in some PA increment) and combining them to make the final dirty image for polarization product k which is then used as the update direction in the deconvolution iterations (Cornwell 2003) as:

$$\boldsymbol{I}_k^d = \int \mathsf{J}_k^{Sky^{-1}}(\psi) \mathcal{FT} \left[\mathsf{J_{ij}}^{-1} \boldsymbol{V}_{ij}^{Obs}(\psi) \right]_k d\psi \tag{2}$$

where ψ is the parallactic angle and the above equation is evaluated in appropriate increments of ψ. Since removing the effects of $\mathsf{J_{ij}}$ is not an issue, for the rest of this paper, $\mathsf{J_{ij}}$ will be assumed to be unity.

For significant antenna pointing errors, J^{Sky} is different for each baseline making it expensive to evaluate the inversion of Eq. 1 for each image pixel. Here we describe techniques to efficiently correct for primary beams including polarization squint and pointing errors. An algorithm to solve for the antenna pointing errors is described in EVLA Memo 84 (Bhatnagar et al. 2004).

2 Direction dependent error correction

Eq. 1, for polarization product P, can be rewritten as:

$$\boldsymbol{V}_P^{M'} = \mathsf{W E}_P \mathsf{F} \boldsymbol{I}_P^M \tag{3}$$

where W is the weights matrix, $\mathsf{E}_P = \mathsf{F J}_P^{Sky} \mathsf{F}^{\mathrm{T}}$ (the data domain representation of J^{Sky}) F is the Fourier transform operator, and $\boldsymbol{I}^M$ is the model image. The Generalized Dirty Image (Cornwell 1995) is the update direction for an iterative deconvolution, given by $\Delta \boldsymbol{I}^{\boldsymbol{D}} = -\mathsf{C} \Delta \chi^2$, where C is the covariance matrix. When J^{Sky} is same for all antennas, approximating the Hessian as a diagonal matrix, the update direction can be computed as

$$\Delta \boldsymbol{I}^{\boldsymbol{D}} = -\Re \left[\mathsf{J^{Sky}}^{\dagger} \mathsf{B} \mathsf{J^{Sky}} \right]^{-1} \Re \left[\mathsf{J}^{Sky^{\dagger}} \boldsymbol{I}^{\boldsymbol{R}} \right] \tag{4}$$

where B is the Beam Matrix and J^{Sky} is computed in appropriate increment of the Parallactic Angle (PA) (Cornwell 2003).

When J^{Sky} is antenna dependent, direction dependent effects cannot be reduced to multiplicative effects in the image domain. E_P then filters the data domain differently for each baseline. In such cases, applying the correction in the image domain becomes impractical (requires data from each baseline to be inverted for each image pixel). Such effects must then be incorporated in the evaluation of Eq. 3 for visibility prediction.

The algorithm to incorporate direction dependent effects during imaging is motivated by the w-projection algorithm (Cornwell et al. 2003). Such techniques rely on accurately computing the transform in one direction and using it in the iterative deconvolution algorithms to achieve convergence. An approximate update direction is computed and the model image is iteratively improved as $\boldsymbol{I}_i^M = \boldsymbol{I}_{i-1}^M + \alpha \max\left[\Delta \boldsymbol{I}_{\mathrm{i}}^{\mathrm{D}}\right]$ $(0 < \alpha < 1)$. Errors due to the approximation are iteratively corrected in the accurate prediction of the model visibilities to achieve convergence.

3 The forward and inverse transforms

When direction dependent errors are present, the predicted visibilities are given by Eq. 3. Typically the gridded visibilities $\mathsf{F}\boldsymbol{I}^M$ are interpolated from a regular grid to the observed (u, v, w) points and re-sampled at the measured (u, v, w) points as:

$$\boldsymbol{V}^{M'}(u, v, w) = \left(\mathsf{G}\left[\mathsf{F}\boldsymbol{I}^M\right]^g\right)(u, v, w) \tag{5}$$

where G is the interpolation operator and superscript g indicates data on a regular grid. This equation is similar to Eq. 3. In the absence of direction dependent effects, the convolution of the two antenna complex illumination patterns are identical and lossless for all baselines. The G operator is therefore real and physically correspond to "focusing" of the coherence field by the antenna surfaces. Direction dependent effects redistribute the energy in the visibility plane and the focusing operation must reflect this via G, which in general is complex valued. For visibility prediction then, Eq. 3 can be straightforwardly used by using E in place of G - provided E has appropriate aliasing properties and can be efficiently evaluated. Note that the model image in this case should be corrected for J^{Sky} for predicting model visibilities. Transform in this forward direction will be accurate and this is all that is required for the update direction computation in Pointing SelfCal algorithm to solve for E given a reasonably accurate model image (Bhatnagar et al. 2004).

The inverse transform will be $[\mathsf{FE}]^\dagger$. Using $\mathsf{E}^\dagger$ as the interpolation operator for re-sampling the data on a regular grid will correct for the effects of E in the Dirty image. An approximate update direction can be computed with correct units by dividing the image by an average J^{Sky}.

These transforms will include the effects of direction dependent terms and will be accurate in prediction of the visibilities and approximate in the other direction. Using these transforms for computing the residual visibilities and residual dirty

image, an iterative deconvolution scheme should ultimately drive the residual image to be noise-like. The errors due to the approximate inverse transform will eventually be corrected to give a deconvolved image corrected for direction dependent effects (see also the discussion in Cornwell 2003). Since the final model image is iteratively built using accurate computations only in one direction, dirty image itself will not have much physical meaning (as it never does).

4 Squint and pointing error correction

In the following section, we use superscripts p and q to represent the two orthogonal polarizations. The forward and inverse transforms used in an iterative deconvolution algorithm are described in the feed polarization basis. The minor cycle uses the Stokes images for updating the model image at each iteration. Stokes model images are converted to feed polarization basis and used in the major cycle for visibility prediction.

4.1 Forward transform

When J^{Sky} correspond to the antenna power pattern projected on the sky, E_{ij}^{pq}, the elements of $\boldsymbol{E}$ for baseline $i-j$ and polarization product pq, itself can be separated into terms that include the squint (which is assumed same for all antennas) and the antenna based pointing offsets (which are different for each antenna and a function of time) as:

$$E_{ij}^{pp} = E^{pp^\circ} f(\phi_i - \phi_j) e^{\iota(\phi_i + \phi_j)} \tag{6}$$

where ϕ_i is the pointing offset and E^{pp° is the Fourier transform of the ideal squinted voltage pattern assumed to be same for all baselines. The function f represents the de-correlation that the signal suffers at each baseline due to the antenna pointing errors. $f(0) = 1.0$ and for voltage patterns with finite support, it will be a monotonically decreasing function of its argument. The exact form of this however will depend upon the exact form of the voltage patterns. To the first order, it will be close to unity for small pointing errors (few percent of the half-power beam width).

For no pointing errors, the parallel hand predicted visibilities will correspond to a sky tapered by the corresponding power pattern (as it should be). With pointing errors, the predicted visibilities will include the effects of pointing errors.

4.2 Inverse transform

To correct for the effects of the primary beam (e.g. the pointing offsets) during imaging, $E_{ij}^{pq^\dagger}$ can be used as the interpolation operator for re-sampling V^{pq} on a regular grid:

$$\boldsymbol{V^{pq,G}} = \left(\boldsymbol{E}^{pq^\dagger} \boldsymbol{V^{pq}} \right) \tag{7}$$

The Dirty images are then computed as

$$\boldsymbol{I^d_{pq}} = \mathsf{F}^{\mathrm{T}} \boldsymbol{V^{pq,G}} \tag{8}$$

The resulting image is tapered by the product of the average power patterns for the two polarizations (E^{p^2} and E^{q^2}). Therefore to compute the dirty image consistent with the true sky (e.g., I^d_{RR} - not tapered by the square of the primary beam), forward transform should be

$$\boldsymbol{I^d_{pq}} = \overline{\mathsf{J}^{\mathsf{Sky}\dagger}_{\mathsf{pq}}}^{-1} \mathsf{F}^{\mathrm{T}} \boldsymbol{V^{pq,G}} \tag{9}$$

where $\overline{\mathsf{J}^{\mathsf{Sky}\dagger}_{\mathsf{pq}}}$ is the average over all antennas for the entire range of parallactic angle coverage.

4.3 The deconvolution algorithm

The above forward and inverse transforms are used in the major and minor cycles of the iterative deconvolution algorithms. The algorithm is as follows:

1. Set the initial model image to zero.
2. Compute the update direction using Eq.9 and the appropriate smoothing operator
3. Update the model image (minor cycle)
4. Use the forward transform to compute the residual visibilities (major cycle)
5. Goto 2 till residuals are noise-like.

The software framework for deconvolution algorithms in AIPS++ involves two basic set of C++ classes. The `SkyEquation` (and related classes) evaluate the update direction and implement the Newton-Raphson minimization. `FTMachine` line of classes implement the forward and inverse transforms, which are used in the minimization iterations. To realize the above algorithm in this framework, a new specialization of the `FTMachine` was developed (the `PBWFTMachine`) and used for forward and inverse transforms. The E_{ij} functions are evaluated using the `VLACalcIlluminationConvFunc` class and the modified `VisJones` class (the `EPJones`) to interpolate the pointing offsets in time. Rest of the machinery for the major-minor cycle iterations remains the same. At the user level, the above deconvolution algorithm can be used by setting the `ftmachine` parameter of the `imager` tool to "`pbwproject`".

5 Error analysis

Convergence of the deconvolution iterations is judged by the statistics in the residual image. The residual image is constructed as:

$$I^R = \sum_{\psi} PSF(\psi) \star [\Delta PB(\psi) I^{\circ}] \tag{10}$$

where ψ is the feed Parallactic Angle, I° is the true sky, $\Delta PB(\psi)$ is the error between the true and the assumed primary beam model at PA= ψ and $PSF(\psi)$ is the instantaneous (snap-shot) PSF. The peak residual can be estimated as a source of flux density S located at the position of the peak of the error pattern multiplied by the maximum sidelobe of the instantaneous PSF:

$$max\{I^R\} = max\{PSF_{sidelobe}(\psi_{\circ})\} max\{\Delta PB(\psi_{\circ})\} S \tag{11}$$

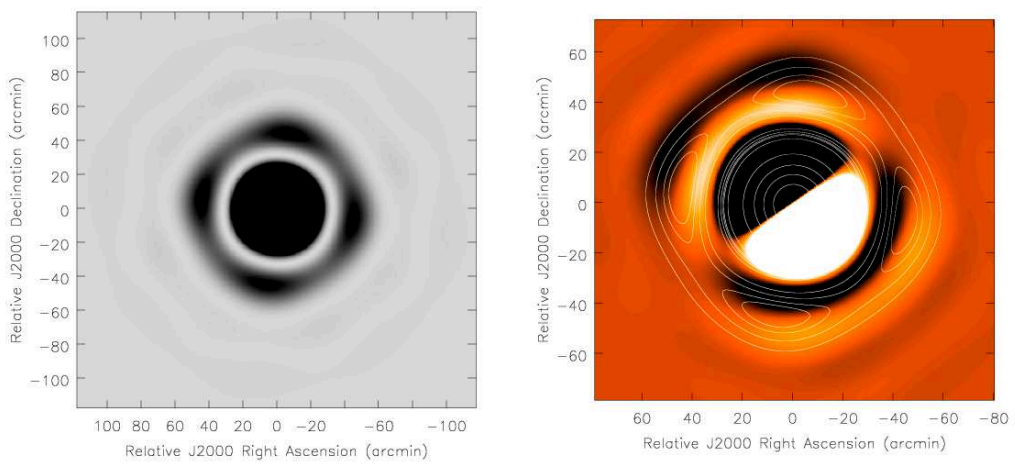

Figure 1: The model for VLA L-band antenna power pattern at Parallactic Angle $\sim 80°$. Left panel shows the Stokes-I power pattern ($[PB_{RR} + PB_{LL}]/2$). Image on the right shows the Stokes-V power pattern ($[PB_{RR} - PB_{LL}]/2$)in colour/gray scale with the contours of the Stokes-I power pattern superimposed. Dark regions in the gray scale image on the right represent negative values due to the polarization squint of VLA antenna.

An instantaneous Stokes-I and -V VLA antenna power patterns at 1.4GHz are shown in Fig. 1. This pattern rotates on the sky with the PA and results in the time varying gain across the field of view. When imaging using azimuthally symmetric PB model, the PB error pattern is given by $\Delta PB(\psi) = \overline{PB} - PB(\psi)$ where $\overline{PB}$ is the azimuthally averaged PB. Rotation of this error pattern on the sky contributes the dominant errors in the residual image (and consequently in the final deconvolved image). The Stokes-I error pattern and an azimuthal cut through this function at

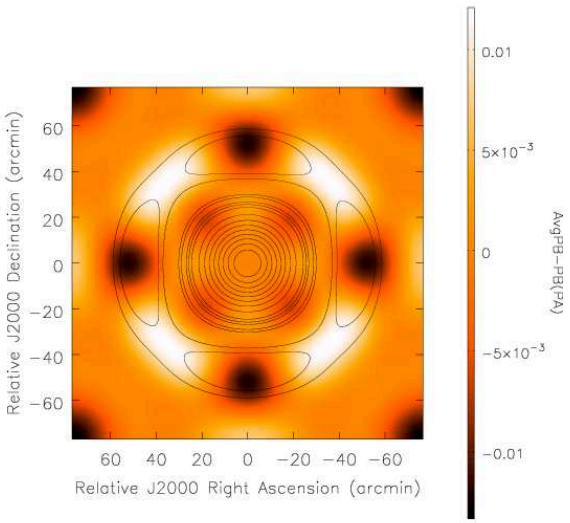

Figure 2: Difference between instantaneous Stokes-I PB and an azimuthally averaged PB. The colour image is $\Delta PB = PB(\psi_\circ) - \overline{PB}$ while the contours are for the $\overline{PB}$.

.

50%, 10% and 1% point of $\overline{PB}$ is shown in Figs. 2 and 3 respectively. The first contour in the image is at 0.95 and the sixth contour is at 0.5 of $\overline{PB}$. Within the main-lobe of the Stokes-I beam, the peak of the error pattern is where $\overline{PB} \approx 1\%$. For $S = 1$Jy, $max_{sidelobe}\{PSF(\psi_\circ)\} = 40\%$ (measured for the test L-band, C-array data) and $max\{\Delta PB(\psi_\circ)\} = 0.005$, the peak residual will be about 2mJy in Stokes-I. Peak residual in Stokes-V would at the level of about 10mJy.

The deconvolution algorithm described above is essentially approximating the function shown in Fig. 1 as a function of PA by a piece-wise constant function. The maximum error due to such an approximation can be estimated using the following equation:

$$I^R_{peak} = \Delta I^R_{peak}\,[\Delta\psi] \tag{12}$$

$$where\ \Delta I^R_{peak} = S\ [max\{PSF_{sidelobe}\}]\ \left.\frac{\partial PB}{\partial \psi}\right|_{max} \tag{13}$$

For an imaging sensitivity requirement of η, the minimum PA increment such that the imaging dynamic range is not limited by the piece-wise constant approximation would be given by

$$\Delta\psi \leq 3\eta/\Delta I^R_{peak} \tag{14}$$

For the L-band, VLA C-array test data (see below), $\left.\frac{\partial PB}{\partial \psi}\right|_{max} \approx 0.0003$. For PA increment of 10°, peak residuals in Stokes-I and -V would be about 1mJy and 5mJy

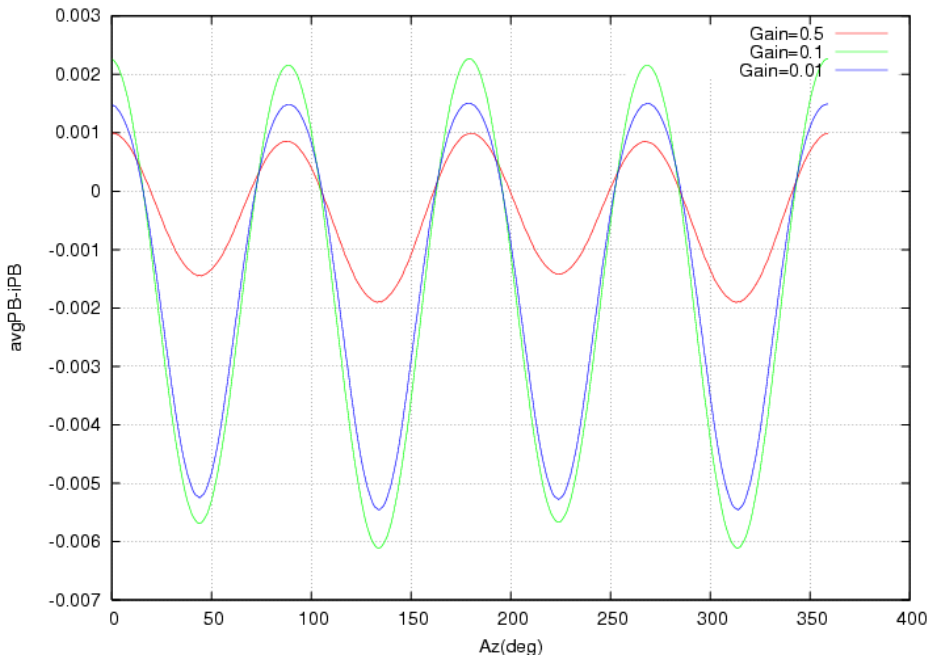

Figure 3: Azimuthal cut through the ΔPB shown in Fig. 2 at points where the gain of the average PB is at $50\%, 10\%$ and 1%.

respectively. With the expected thermal sensitivity of 0.2mJy/beam for the L-band data we used, PA increments of 1° had to be used.

The PA increments for EVLA such that imaging is not limited by the above approximation will be much higher. Use of a better interpolation of the aperture functions will significantly relax the PA-increment limit. Since image interpolation itself can be expensive, a better approach is probably to pre-compute the aperture functions at smaller PA increments in a disk-cache. The gridding cost is relatively insensitive to the number of convolution functions used. A hybrid approach of FFT based transforms plus analytical computations for the strongest sources will probably deliver optimal performance.

6 Implications for wideband and mosaic imaging

Antenna pointing errors, non-azimuthally symmetric aperture illuminations, wide bandwidths and deconvolution errors due to use of pixel basis for sky representation all leave residuals at few 10s of micro Jy level. Therefore, apart from correcting for the direction dependent effects, for the highest imaging dynamic ranges allowed by the thermal sensitivity of the EVLA, scale-sensitive decomposition of the sky will also be necessary (Bhatnagar & Cornwell 2004). While this may be more expensive

than the current deconvolution algorithms in use (Clean, MEM, etc.), it will provide the required accuracy. With higher accuracy in visibility prediction, it will relax the requirement on the sampling rate along the PA axis and may even be more efficient overall.

The algorithm described here accounts for the variations due to azimuthally asymmetric aperture illumination and consequent gain variations as a function of time. For imaging with bandwidth ratio of 2:1, the dominant error terms will be the scaling of the power pattern with frequency. Sources which will be well within the main lobe of the primary beam at the lower frequency end of the band will be outside the main lobe at the higher frequency end (and may even appear in the first sidelobe). Since the azimuthal variations in the power pattern due to feed legs/sub-reflector blockages are maximum close to the null and in the first sidelobe, frequency scaling of the aperture illumination will contribute the first order error.

Scaling with frequency in observations with wide bandwidths can be incorporated by computing the aperture illumination functions at appropriate increments in frequency. Since the computing cost scales weakly with the number of convolution functions used, the extra computing load will not be too high. The cost of computing the aperture functions itself will be significant, but it is a one-time cost.

The sidelobe gain variation as a function PA is $\sim 50\%$. For mosaic observations, this will contribute significant time varying flux in most individual pointings. Assuming the peak PSF sidelobe of 10%, sidelobe gain variations will contribute an error of few percent of the flux density at the location of the PB sidelobes. This clearly will significantly limit the mosaicking dynamic range and will be the first order effect that will need to be corrected for. The second order error due to antenna pointing errors will also need to be corrected for. Neither of these dominant effects are included in existing simulations to estimate the imaging performance of ALMA.

The algorithm described above includes the pointing error term. The problem of mosaicking is a straight forward generalization of this work where the vector for the center of each pointing with respect to the mosaic field center can be used as the antenna pointing offset. The resulting phase gradient in the aperture plane and proper normalization in the image domain is internally computed already in the `PBWFTMachine` and can be straight forwardly used for forward and inverse mosaic transforms. Rest of the machinery for iterative deconvolution remains unchanged. Since this FTMachine nominally corrects for the time varying sidelobe gains and antenna pointing errors, one can expect these dominant errors to be significantly reduced. With some more work (related to internal book-keeping of convolution functions), inhomogeneous arrays like CARMA/ALMA (not all antennas in the array are identical even to the first order) can be easily handled as well. (Ultimately, at the very high dynamic range imaging with telescopes like the SKA, nominally identical antenna elements may have variations at levels higher than the thermal noise sensitivity limit. In that sense, such telescopes will also need to be treated as inhomogeneous arrays).

7 Results

7.1 Simulations

The algorithm was tested for squint correction by imaging visibilities simulated using the standard AIPS++ simulator which can simulate VLA polarization squint for *azimuthally symmetric primary beam patterns*. A model for typical sky emission at 1420MHz was generated using the NVSS source list. The PA increment of 10° was used in order to rotate the squinted R- and L-beams on the sky. The visibilities were simulated for VLA B-array and an RMS noise of $\sim 1\mu$Jy/beam was added.

Fig. 4 shows the Stokes-V deconvolved images with and without squint correction applied during deconvolution. Squint correction results in noise figure improvement by a factor of ~ 10 and more than 20 times improvement in the image fidelity.

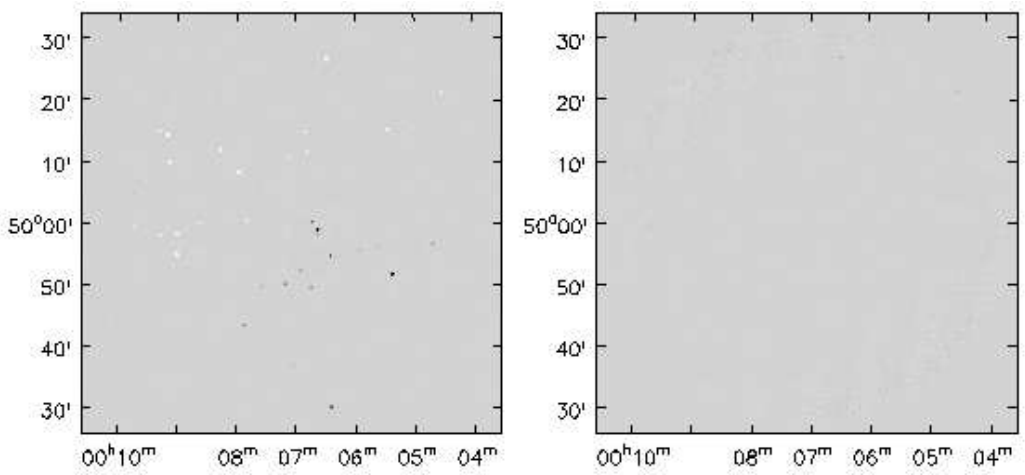

Figure 4: The deconvolved Stokes-V images: The image on the left had no squint correction applied while squint correction was applied to the image on the right. The peak and RMS noise in the images are 2mJy and $\approx 10\mu$Jy/beam and 70μJy and $\approx 1\mu$Jy/beam respectively.

AIPS++ simulator (or any other simulator that we know of) cannot simulate visibilities with time varying antenna pointing errors (for the same reasons that makes it impractical to apply known pointing errors during imaging). The visibilities with pointing errors where therefore simulated by predicting the model visibilities using the `PBFTMachine` in increments of 10° in PA. A model for the VLA aperture illumination pattern (Brisken 2003) was used to generate a non-azimuthally symmetric power pattern. The mean of the pointing errors were randomly distributed between $\pm 25''$ with an RMS of $5''$. The images where then deconvolved using the above algorithm and the results are shown in Figs. 5 and 6. Fig. 5 shows the Stokes-I images. As expected, the deconvolution errors were maximum for the sources around the half-power point of the primary beam. These errors were significantly reduced

(close to the noise level) when the pointing and squint corrections were applied during deconvolution. Fig. 6 shows the Stokes-V images (without and with pointing and squint corrections) which separates the effect of only pointing offsets.

7.2 VLA L-band data

The algorithm was also tested for Stokes-I and -V imaging using a VLA L-band, 2×7hr observation (without solving for pointing offsets). The field contains two strong ($\sim 1Jy$) sources on either side of the pointing center, with one of them located close to the half-power point of the primary beam. The aperture illumination pattern for each antenna was assumed to be the same and was computed using the model for VLA antennas (Brisken 2003). The model includes geometry of the sub-reflector and the feed position as well as the aperture blockage due to the feed legs and the sub-reflector. The aperture illuminations were computed as a function for PA increments of 1°. The expected thermal noise for this data is ~ 0.2mJy. The results of the imaging run with and without the correction for time varying primary beam gains and polarization squint are shown in Figs.7 and 8. The peak residual in the Stokes-V images without and with primary beam corrections are 7mJy and 0.3mJy respectively.

Acknowledgments

We thank W. Brisken for providing the code for VLA antenna aperture illumination modeling and various related impromptu discussions and J. Uson for providing the VLA L-band test data.

References

Bhatnagar, S. & Cornwell, T. J. 2004, A&A, 426, 747

Bhatnagar, S., Cornwell, T. J., & Golap, K. 2004, Solving for the antenna based pointing errors, Tech. rep., EVLA Memo 84

Brisken, W. 2003, Using Grasp8 To Study The VLA Beam, Tech. rep., EVLA Memo 58

Cornwell, T. J. 1995, The Generic Interferometer: II Image Solvers, Tech. rep., AIPS++ Note 184

Cornwell, T. J. 2003, Full primary beam Stokes I,Q,U,V imaging, Tech. rep., EVLA Memo 62

Cornwell, T. J., Golap, K., & Bhatnagar, S. 2003, W-Projection: A new algorithm for non-coplanar baselines, Tech. rep., EVLA Memo 67

Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137

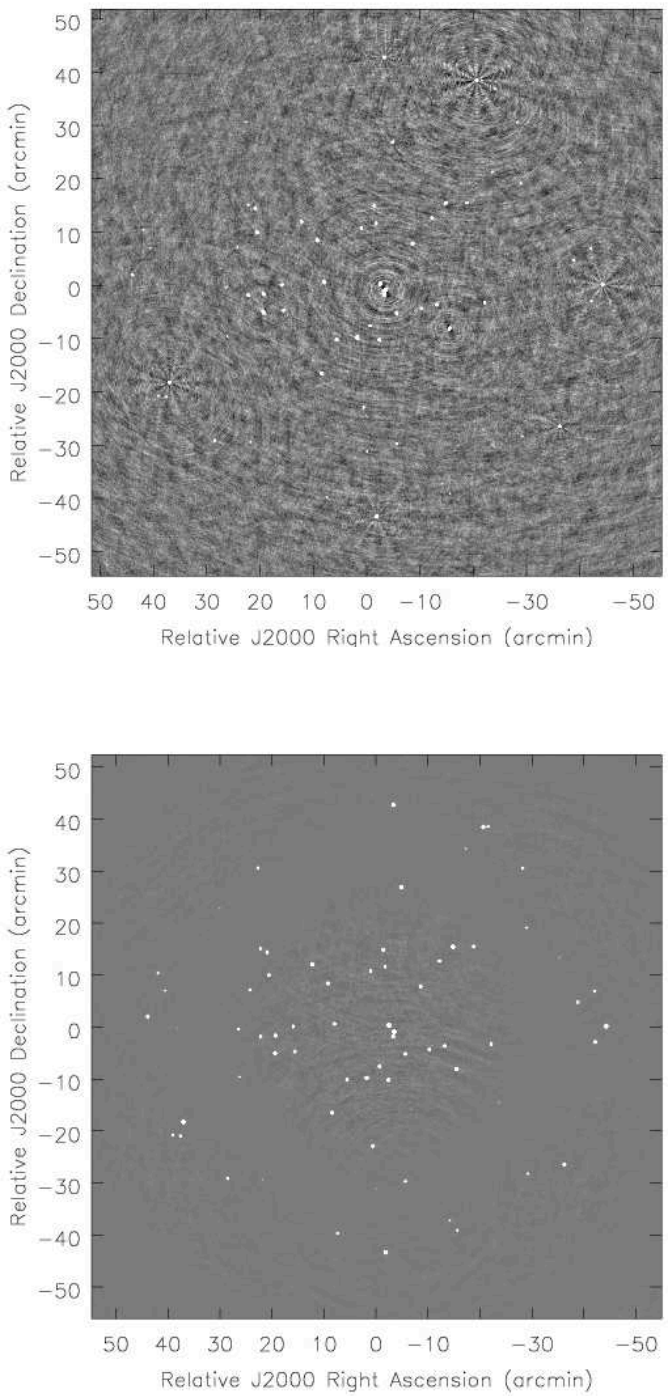

Figure 5: Stokes-I imaging: Image in the top panel was made without squint and pointing correction while bottom image had both corrections applied. The systematic deconvolution errors seen at the location of the strongest sources around the half-power point are due to the antenna pointing errors.

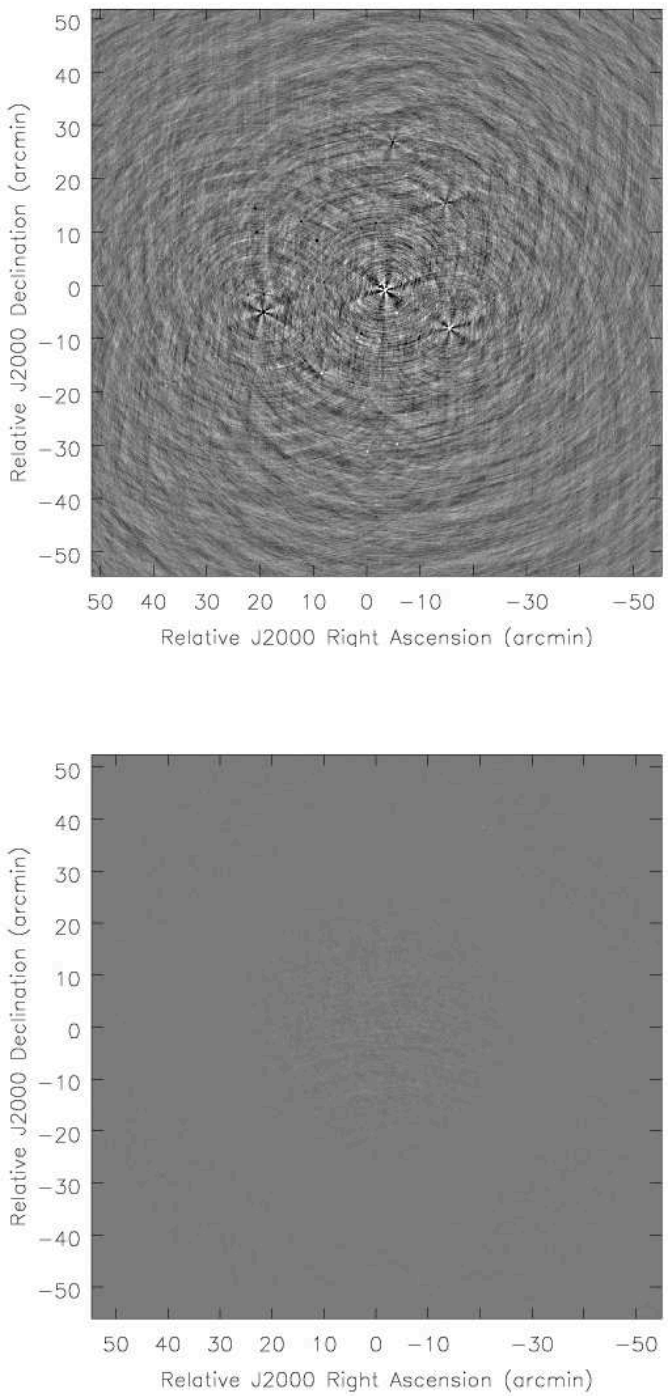

Figure 6: The residual Stokes-V images: Only VLA polarization squint corrections were applied to the image on the top. Both, antenna pointing offset and squint corrections were applied to the image in the bottom panel.

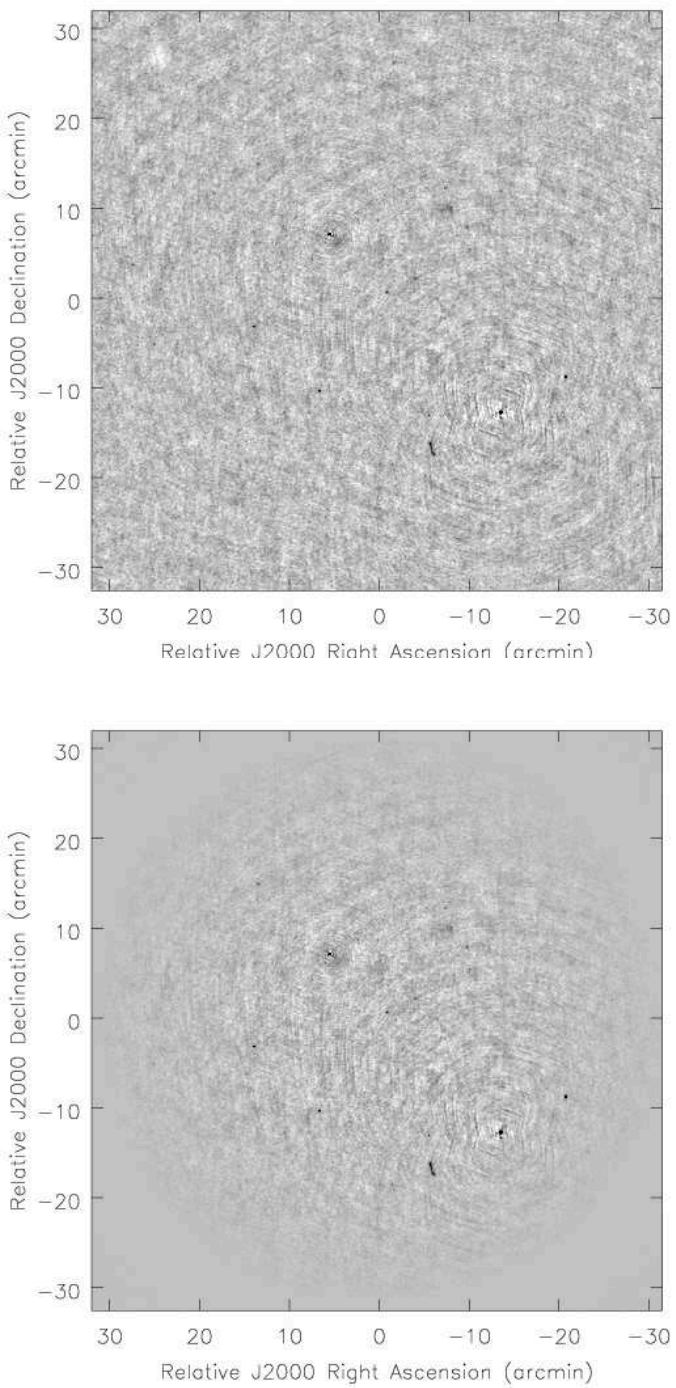

Figure 7: The Stokes-I images for a L-band VLA C-array data. Top panel shows the deconvolved image without correcting for the antenna power pattern variations as a function of the Parallactic Angle. Bottom panel shows the result of the deconvolution using the algorithm described above. The two dominant sources, on either side of the pointing center have a flux density of $\sim$ 1Jy each.

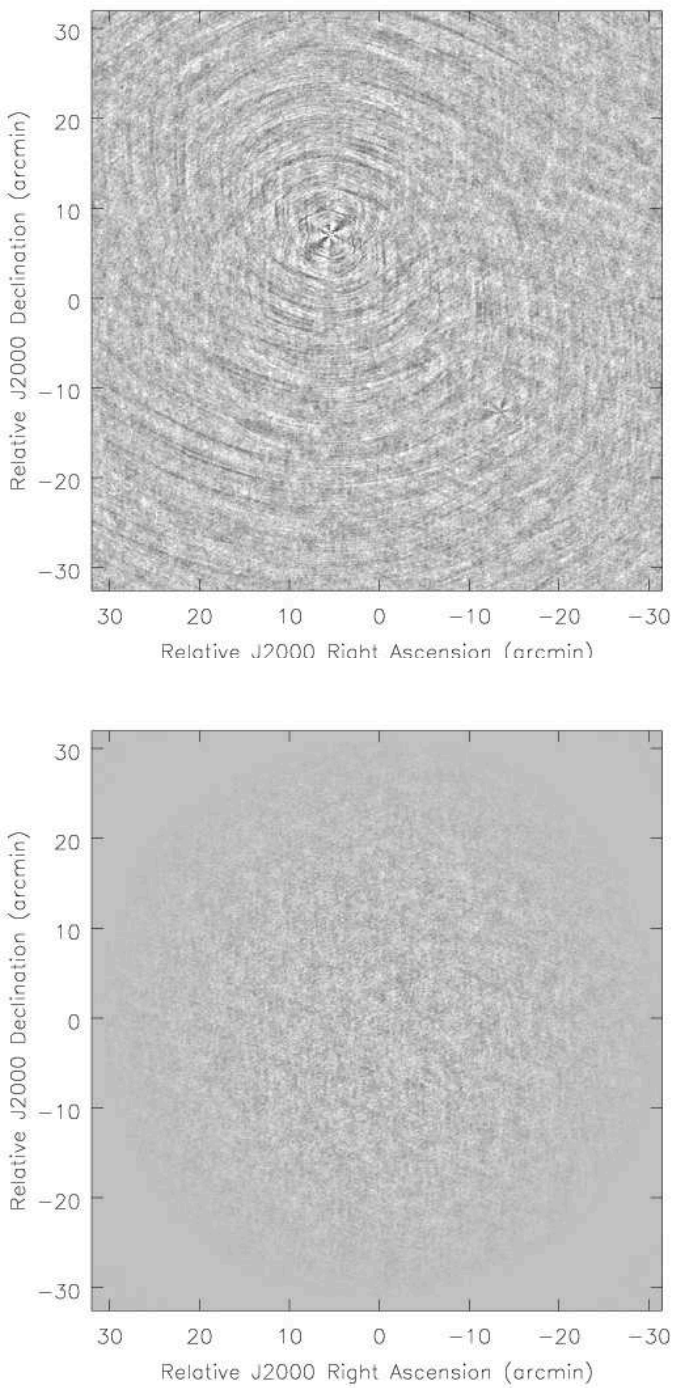

Figure 8: The Stokes-V images for a L-band VLA C-array data. The top and bottom panels show the images without and with PB-corrections. Errors due to the polarization squint are clearly above thermal noise in the top image while the bottom image is consistent with the thermal noise (0.3mJy/beam).