

EVLA Memo #84

Solving for the antenna based pointing errors

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

October 18, 2004

Abstract

The imaging dynamic range of an aperture synthesis telescope for mosaicing and for fields with significant flux throughout the antenna primary beams can be limited by the knowledge of the individual primary beams projected on the sky. For high dynamic range imaging of such fields, one requires an accurate measurement of the shape of the primary beams and the pointing offsets as a function of time. The effect of antenna pointing errors remain separable in the visibility domain. With at least two, well separated sources along the RA and Dec axis each to constrain the solutions, it is possible to solve for these errors in an antenna based fashion in the visibility domain.

Here we analyze the effect of antenna based pointing errors on the imaging dynamic range and fidelity and present an algorithm to solve for these errors using a model for the sky brightness distribution. For a typical L-band eVLA simulation with typical pointing errors for the VLA antennas, the RMS noise can be reduced by a factor of ~ 10 using this algorithm. The improvement in the image fidelity is even larger. The formulation given here can be further extended to include other direction dependent effects - specially for application to mosaicing observation. Extension of this work for such, more sophisticated solvers is in progress.

1 Introduction

In the absence of pointing offsets, the image made with an ideal aperture synthesis telescope with identical antennas is the true sky multiplied with the ideal primary beam pattern of the antennas. The observed visibilities

are then expressed as $E_{ij} \star V^{True}$, where $E_{ij} = E_i \star E_j^*$ and E_i is the antenna aperture illumination pattern. With no pointing offsets E_{ij} is identical for all the baselines which can be used as the gridding convolution function for computing model visibilities during the deconvolution major cycle. In the presence of antenna pointing errors, E_{ij} is potentially different for each baseline, resulting in a systematic error pattern in the residual image which limits the dynamic range of images where there is significant flux throughout the field of view.

The multiplicative antenna based complex gain, constant across the field of view, can be measured from the observation of even a single unresolved source using the SelfCal algorithm. Antenna pointing errors on the other hand result in a gain which changes across the field of view. With at least two well separated sources along the RA and the Dec axis and a good model image, it should be possible to derive the antenna based pointing offsets by minimizing the differences between the observed and the model visibilities. In the following sections we analyze the effects of antenna pointing errors and, using simulations, show that such errors can be solved for in an antenna based fashion. For a typical L-band eVLA simulation with typical pointing errors for the VLA antennas, the RMS noise can be reduced by a factor of ~ 10 using this algorithm. The improvement in the image fidelity is even larger. Section 2 contains the theory and the mathematical description of an iterative solver. Section 3 contains the details of the simulations and the results of applying the solver on the simulated data. Section 4 discusses the implications on the imaging dynamic range and fidelity and section 5 discusses the implementation details in AIPS++.

2 Formulation

The measurement equation for a single pointing observation with antenna based pointing offsets is:

$$V_{ij}^o = E_{ij} \star V_{ij}^{True} \quad (1)$$

where E_i is the complex aperture illumination pattern, $E_{ij} = E_i \star E_j^*$, V_{ij}^o and V_{ij}^{True} is the observed and the true visibility from the baseline $i-j$, and \star is the convolution operator. Let $P_i(l)$ be the primary beam pattern for antenna i given by $P_i(l) = A_i e^{-(l-l_i)^2 \sigma_i^2}$ where σ_i is the inverse of the width

and l_i is the position offset with respect to the true pointing direction. E_{ij} for $\sigma_i = \sigma_j = \sigma$ and $A_i = A_j = 1$ (identical antennas), is given by:

$$E_{ij} = \left[\frac{1}{\sigma} \sqrt{\frac{\pi}{2}} e^{-\frac{u^2 \pi^2}{2\sigma^2}} \right] \left[e^{-\frac{(l_i - l_j)^2 \sigma^2}{2}} \right] \left[e^{-\pi i u (l_i + l_j)} \right] \quad (2)$$

where u is the conjugate variable in the visibility space. The first term in Eq. 2 correspond to the convolution of the *ideal* antenna illumination patterns ($E_i^\circ \star E_j^{\circ*} = E_{ij}^\circ$). The second term is the loss in the amplitude of this ideal convolution due to the pointing offsets while the third term is the phase due to the pointing offsets. For $l_i = l_j = l_o$, the resulting power pattern in the image domain will be the same as the ideal one, but its center shifted by l_o . In the Fourier domain, this would result in only a phase ramp and no amplitude loss. On the other hand if $l_i = -l_j$, the power pattern in the image domain will be attenuated but still symmetrically centered at the nominal pointing direction. In this case there will be no phase change across the aperture but will have modified amplitude of the illumination pattern.

In a least squares minimization scheme to solve for l_i s, the step size will be computed using the derivative of the χ^2 as:

$$\nabla \chi^2 = -2\Re \left[\mathbf{V}^{\text{R}\text{T}} \nabla \mathbf{V}^{\text{R}} \right] \quad (3)$$

where \mathbf{V}^{R} is the residual visibility vector and the superscript T represents the transpose operator. Using Eq. 2, we get:

$$\nabla E_{ij} = [(l_j - l_i)\sigma^2 - \pi i u l_j] E_{ij} \quad (4)$$

E_{ij} itself, to the first order of the last two terms of Eq. 2, can be written as:

$$E_{ij} \approx E_{ij}^\circ [1 - R_{ij} - K_{ij} + R_{ij} K_{ij}] \quad (5)$$

where $R_{ij} = (l_i - l_j)^2 \sigma^2 / 2$ and $K_{ij} = \pi i (l_i + l_j) u$. If we can use a functional form for E_{ij}° , Eqs. 4 and 5 can be evaluated for every trial l_i and used as the de-gridding function to compute the two terms of Eq. 3. The last term in Eq. 5 is a third order term (the phase curvature across the aperture) and probably should be dropped (since we are working under the assumption that the pointing offsets are small). The third term is the first order error signal, which appears as a phase ramp. The second term is a second order term in pointing offsets, which appears as an amplitude modulation. This however should be retained since if we drop this term, the solver will

be insensitive to the case were the two antennas have the same pointing offsets in magnitude but opposite in direction ($l_i = -l_j$). Such a case will produce an error signal only as a change in the amplitude with respect to the ideal illumination pattern.

2.1 Point source case

For a test case of a collection of point sources of amplitude I_k located at l_k (model image $I^M(l_k) = I_k \delta(l - l_k)$), Eq. 3 gets further simplified. The corrupted visibilities are given by:

$$V_{ij}^C = E_{ij} \star V_{ij}^M = \int I^M(l) P_i(l) \cdot P_j(l) \cdot e^{2\pi i u_{ij} l} dl \quad (6)$$

and $\nabla V_{ij}^R = -2\sigma^2 [V_{ij}^{C'} - l_i V_{ij}^C]$ where $V_{ij}^{C'} = \int I^M(l) l P_i(l) P_j(l) e^{2\pi i u_{ij} l} dl$. This and Eq. 6 can be used to compute the update direction given by Eq. 3. The update equation for the estimated pointing offset at the n^{th} iteration (l_i^n) is then given by:

$$l_i^n = l_i^{n-1} + \alpha \nabla \chi^2|_{l_i^{n-1}} \quad (7)$$

where $\alpha < 1.0$ is the step size. The iterations are stopped when $|\chi_n^2 - \chi_{n-1}^2| \leq \epsilon$.

3 Simulation

The algorithm was tested using a VLA C-array multiple snapshot simulated observation at 1.4 GHz. The sky was simulated using 49 NVSS sources within the field of view with flux density in the range $\sim 2 - 200$ mJy. The primary beams of the antennas were assumed to be a gaussian of HPBW = $\lambda / (D \sqrt{\log(2)})$ (where D is the antenna diameter) and similar except for the random pointing offsets. The antenna offsets were uniformly distributed between $\pm 20''$ about the pointing center. A baseline based gaussian random RMS noise corresponding to image plane RMS noise of $\sim 1 \mu\text{Jy}$ was added to simulate the thermal noise. The dirty image without the pointing offset calibration is shown in Fig. 1 and the typical pointing offsets as a function of time are shown in Fig. 2. The convergence of the algorithm for a noise free case is shown in Fig. 3 which shows the χ^2 as a function of the iteration. The resulting residual image without applying

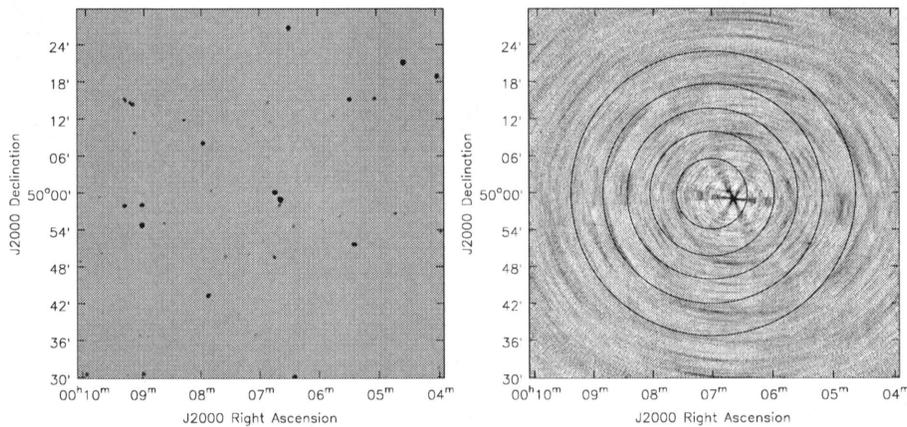


Figure 1: Figure showing the simulated dirty image (right) for a VLA C-array observation at 1.4GHz using 49 point sources extracted from the NVSS source list (left). The antenna based pointing offsets were uniformly distributed between $\pm 20''$. The overlaid contours show the ideal primary beam at 0.09, 0.18, 0.36, 0.54, and 0.72 times the peak value of 1.0. The dominant source near the center has a flux density of 195mJy.

the pointing offset calibration ($V_{ij}^R = V_{ij}^o - V_{ij}^M$) is shown in the top panel of Fig. 4. The RMS of $\sim 12\mu\text{Jy}$ in this image is dominated by the antenna based pointing errors. Eq. 7 was then used in an iterative loop to estimate the individual antenna pointing offsets with $\epsilon = 10^{-5}$. The residual pointing errors after calibration (difference between the solved and actual pointing offsets) as a function of time are also shown in Fig. 2. The residual RMS pointing offset is $\sim 1''$. The estimated pointing offsets were then used to correct the observed visibilities. The residual image using these corrected visibilities ($V_{ij}^R = V_{ij}^o - E_{ij} \star V_{ij}^M$) is shown in the lower panel of Fig. 4. The RMS noise in this image is consistent with the expected thermal noise ($\sim 1\mu\text{Jy}$). In a realistic case, where the model image may not be a collection of point sources, computation of V_{ij}^R will require the use of FFT and de-gridding. Since this can be an expensive operation, improving the rate of convergence is important. A fixed step size (α in Eq. 7) was used in this iterative scheme. Rate of convergence can probably be significantly improved by estimating the step size from the inverse Hessian.

For a typical L-band field and thermal noise of $1\mu\text{Jy}$, the RMS noise in

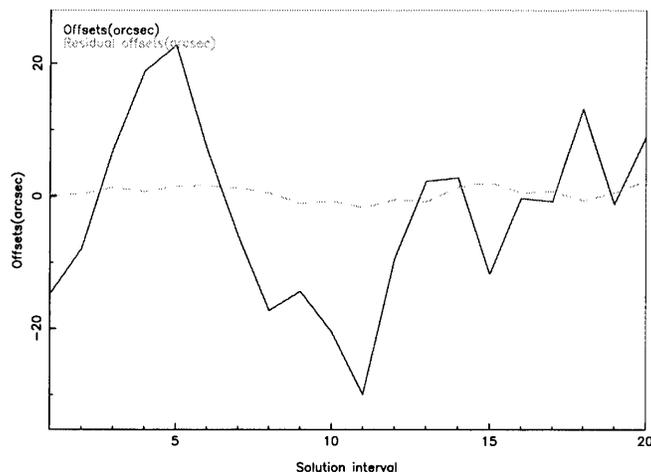


Figure 2: Figure shows the typical antenna based pointing offsets as a function of time. The continuous line (red) shows the pointing offsets used in the simulation and the dashed line (green) shows the residual pointing offsets after pointing calibration. The RMS residual pointing errors are $\sim 1''$.

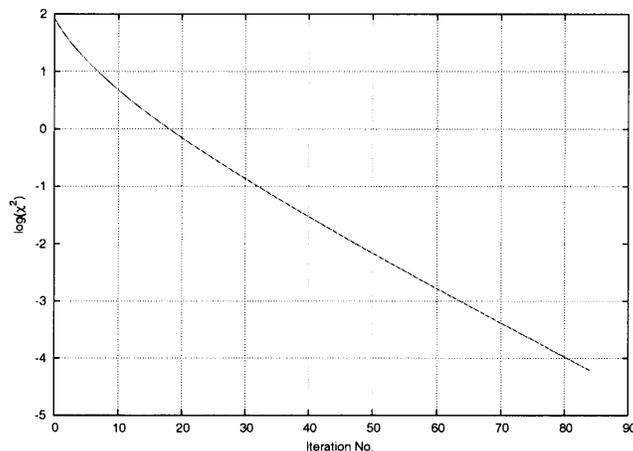


Figure 3: χ^2 as a function of the iteration number using a fixed step size for a noise-free case. For a case with noise, the χ^2 saturates at approximately the value of the square of the L^2 -norm of the noise vector.

the image due to the random antenna based pointing errors of $\pm 25''$ is $\sim 12\mu\text{Jy}$. The strongest source in this field was 195mJy and was located close to the nominal pointing center where the effects of pointing errors

is minimized. With a 0.5mJy source located at the half-power point of the primary beam, the RMS noise due to pointing errors is $\sim 50\mu\text{Jy}$. For a RMS thermal noise of $\sim 1\mu\text{Jy}$ in the image, the SNR per baseline was sufficient for estimating the antenna based pointing errors to an accuracy of $\sim 1''$.

The pointing errors for the VLA antennas is $\sim 25''$ without reference pointing. With reference pointing done every $\sim 30^m$, the residual RMS pointing error is $\sim 2''$. A more realistic pointing error model for the VLA antennas is an RMS error of $\sim 2''$ with a different mean for each antenna, uniformly distributed between the various antennas in the range of $\pm 20''$ and a coherence time of $\sim 30^m$. With such a pointing error model in the above simulations, the pointing noise was $\sim 15\mu\text{Jy}$ and the peak residual of $\sim 250\mu\text{Jy}$. The noise after pointing calibration was consistent with the thermal noise $\sim 1\mu\text{Jy}$ and the peak residual was $\sim 5\mu\text{Jy}$. The residual images before and after pointing selfcal are shown in Fig. 5. The typical antenna based pointing offsets and the residual pointing offsets after pointing calibration as a function of time are shown in Fig. 6.

4 Implications for imaging

Since the image model used in this simulation is exact, the residual visibilities are a convolution of the model visibilities and the function $\Delta E_{ij} = E_{ij}^{\circ} - E_{ij}$ where $E_{ij}^{\circ} = E_i^{\circ} \star E_j^{\circ}$ and E_i° is the ideal electric field pattern. To get an understanding of the error pattern in the residual image, consider the case where only one antenna has pointing errors. The residual visibilities on all baselines not involving this antenna will be exactly zero. This would correspond to a uv-coverage consisting of all the baselines with only this antenna and the corresponding “error PSF”. For small offsets, the error pattern will be a small residual amplitude convolved with this error PSF in the image domain. When more than one antenna has pointing errors, the modified PSF will correspond to the uv-coverage of baselines involving all these antennas. The convolving error pattern seen in the first image in Figs. 4 and 5 is this error PSF. Note that the error PSF is significantly different from the true PSF (more clearly seen in Fig. 5). Note also that the systematics in Fig. 4, where the mean of the pointing offsets of the antennas was zero but the RMS was $20''$, is less pronounced than in Fig. 5 where the antennas had a non-zero mean pointing error as a function of time with a smaller RMS. This shows that the systematics in the residual image are dominated by constant antenna pointing offsets as

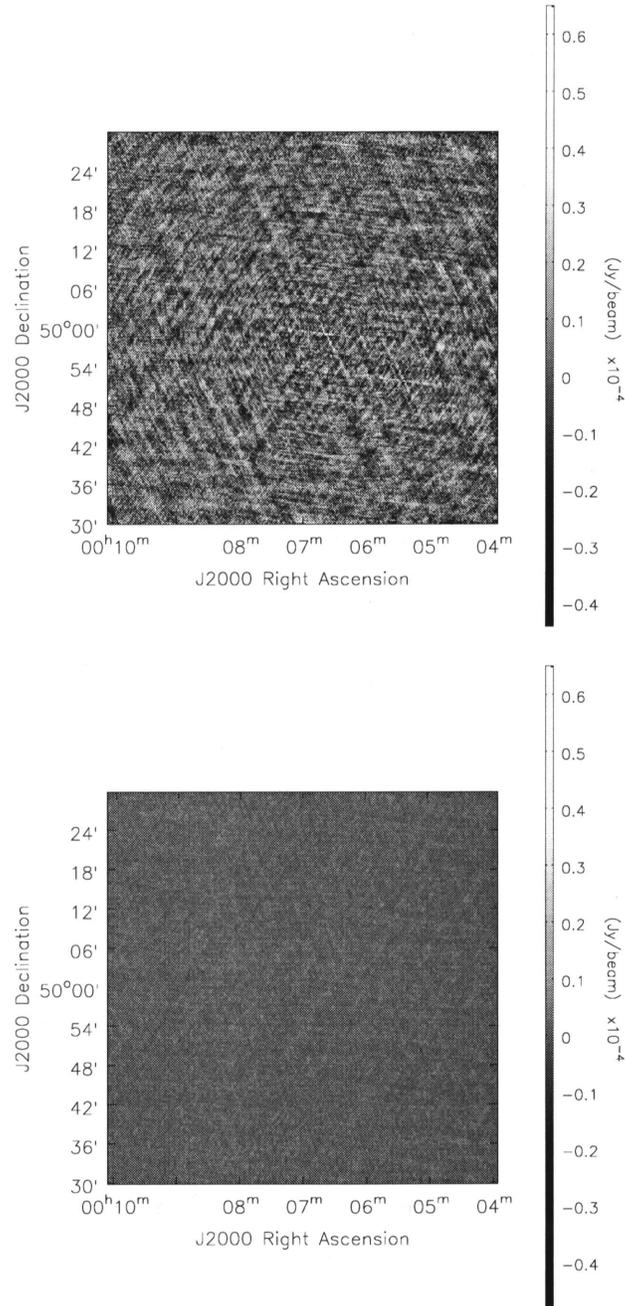


Figure 4: Figure shows the residual images before (top panel) and after applying the antenna based pointing calibration (bottom panel). The two images are shown at the same scale. The RMS noise in the top and bottom images is $\sim 12\mu\text{Jy}$ and $\sim 1\mu\text{Jy}$ respectively.

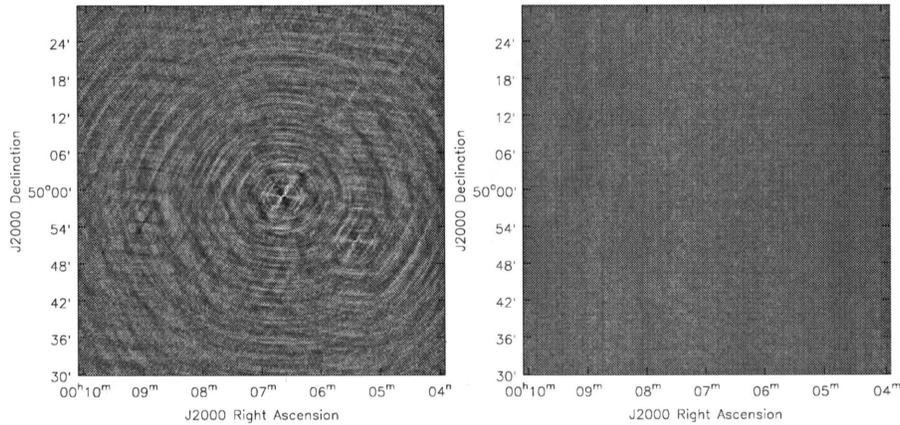


Figure 5: The image on the left is the residual image with the mean antenna pointing offsets between $\pm 25''$ with an RMS of $2''$ (see Fig. 6). The image on the right is the residual image after applying the pointing offset correction. Both images are at the same scale ($\pm 150 \mu\text{Jy}/\text{beam}$).

a function of time. The case of all antennas with the same pointing offset, will correspond to a primary beam with its peak centered away from the pointing center. This will apply a systematic amplitude gain error across the field of view. The error PSF will be same as the normal PSF, and the error pattern in the image domain will appear as the normal PSF convolved with the residual flux due to the differential gain across the field of view. For a given pointing error, the value of ΔE_{ij} will be maximum at the half power point. Its amplitude will be proportional to the square of the sum of the offsets of the two antennas (a second order effect) and the phase will be proportional to difference of the two offsets (a first order effect) (see Eq.2). Since the residual visibilities are a convolution of the true (not tapered by the primary beam) visibilities and ΔE_{ij} , pointing errors will not have a significant effect on the imaging performance if either the flux density at the half power point is small, or the pointing offsets are small compared to the primary beam. Clearly, for mosaicing observations of a source larger than the primary beam, the effect of pointing errors will be significant.

The effect of random antenna based pointing offsets is to create a complex aperture illumination pattern whose amplitude symmetrically de-

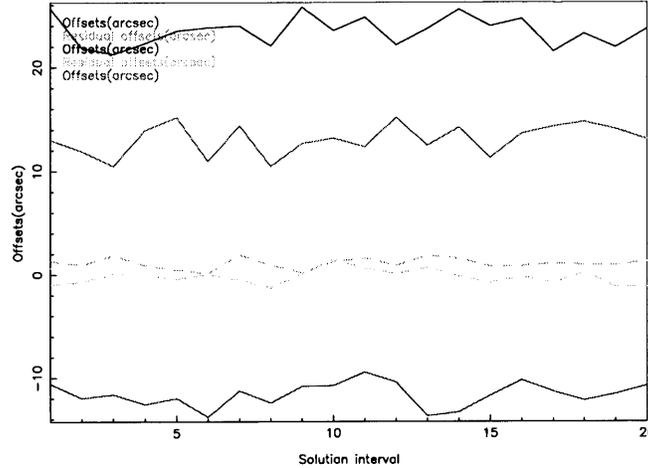


Figure 6: Figure shows the typical antenna based pointing offsets as a function of time used for the images shown in Fig. 5. The continuous lines show the pointing offsets used in the simulation and the dashed lines show the residual pointing offsets after pointing calibration. The RMS residual pointing errors are $\sim 1''$.

creases away from the center with a phase gradient across the aperture proportional to the pointing error. As a result, apart from decrease in the amplitude, the function E_{ij} has a phase ramp proportional to the difference in the pointing errors of the two antennas (Eq.2). This can masquerade as an antenna based complex gain. Roughly speaking, applying the standard amplitude gain correction to such a data will take out the average of the amplitude error as a constant term. For a trivial case of a single unresolved source in the field, the solutions for the pointing calibration and the multiplicative constant (across the aperture) complex gains are indistinguishable from each other. However since the errors due to the pointing errors are not constant across the aperture, for more than one source in the field, residual errors will be left in the data. It is easy to see that one needs at least two well separated sources along the RA and the Dec axis each to constrain the solutions for the pointing errors. To determine the extend to which the complex gain calibration will be able to remove the errors due to the antenna pointing, an amplitude and phase selfcal was done. For the simulation shown above, after correcting for complex gains the RMS noise in the image was $\sim 6\mu\text{Jy}$ (taken over the entire image). Clearly, selfcal is not sufficient to remove all the errors. Also, this residual noise will be a function of the brightness distribution and will

be significant where there is significant flux distributed all over the field of view. In general the errors due to the pointing offsets will depend on the fractional pointing offsets and the flux density at the nominal half-power point of the primary beam.

The linear phase ramp across the aperture due to the pointing errors has an intercept of zero (the phase at the peak of E_{ij} is zero). In the presence of multiplicative antenna based complex gains, this intercept will be non-zero (a phase ramp with an offset). The effect of the standard phase selfcal will be to remove this phase offset. This offset however can be easily incorporated as part of the pointing offset calibration.

The success of the pointing error calibration also depends on the accuracy of the model image. For an observation where the noise is limited by the pointing errors, typically the pointing calibration should be used in a calibration and imaging cycle. Since the pointing calibration can absorb the multiplicative complex antenna based gains, it can replace the amplitude and phase selfcal to produce an improved model image in the deconvolution phase.

5 Implementation details

For a simple case of a collection of point sources for the image model, we implemented the solver in Glish and used DFT to compute the model visibilities at each iteration. The results shown in this note were generated using only Glish script in AIPS++. The crucial and the most compute intensive steps in the algorithm are the computations of the χ^2 and its derivative. In general, this needs a modified version of the FFT based gridding/de-gridding operation where the gridding convolution function is evaluated for each baseline using the Eqs. 4 and 5. We discuss below the details of implementing such a gridded/de-gridded using existing AIPS++ classes.

The machinery for solving for antenna based quantities of the interferometric measurement equation is implemented in the class tree of the `VisJones` class. The `solve()` method of the derived class `SolvableVisJones` implements iterative solver to solve for various Jones types. For frequency independent Jones types, this method first averages the data over the available frequency range and then loops over time, using the services of the `VisEquation` to compute the χ^2 and its first derivative and the diagonal

terms of the Hessian. These are then used for the Conjugate Gradient iterations. The pseudo-code for this method is as follows:

```

SolvableVisJones::solve()
  for all chunks of the VisSet
  {
    VisEquation::initChiSquare; # Apply any existing Jones
                                # and average in frequency

    Setup the initial guess;
    ::initializeGradients();    # Resize and initialize the
                                # arrays for first and the
                                # second derivatives

    ::invalidateCache();        # Implemented in
                                # TimeVarVisJones.

    ::updateCahce();           # Use the cache for the
                                # antenna based gains
                                # (antGainCache_) to
                                # compute  $J_i * conj(J_j)$ 
                                # and store in cache
                                # for the interferometer
                                # gains (intGainCache_)

    #
    # Compute the Chisq and its derivatives. This also
    # computes the residual visibilities (using
    # VisJones::apply()).
    #
    VisEquation::gradientsChisSquared();

    for all iterations
    {
      ::updateAntGains();      # Take a step
      ::invalidateCache();
      ::updateCache();
      ::initizlizeGradients();
      VisEquation::gradientsChisSquared();
      Check for convergence;
    }
  }

```

The `VisEquation::gradientsChiSquared()` method internally uses the `SolvableVisJones::apply()` method to apply the current solutions to the model visibilities (equivalent of the left hand side of Eq. 1) referred in the code as the “corrupted visibilities” and the `SolvableVisJones::applyGradient` method to compute the first derivatives of the χ^2 (equivalent of the Eq. 3). Since the operation of applying the effects of pointing errors to the visibilities is different from that required for multiplicative errors (convolution as against multiplication), `apply()` and `applyGradient()` methods needs to be overloaded in a specialization of the `SolvableVisJones` class.

Model visibilities at the sampled (u,v) values are computed by convolving the model visibilities sampled on a regular grid by a convolution function and re-sampling the result at the observed (u,v) values. The usual convolution gridding function used is the Prolate Spheroidal function, which has a limited support and is designed to minimize the aliasing and attenuation due to its roll-off in the image domain. To incorporate the pointing offsets in the measurement equation, the de-gridding method needs to use a different gridding function for each of the baselines (namely, E_{ij}).

Gridding and de-gridding services in AIPS++ are provided by the `put()` and `get()` methods respectively of the `FTMachine` class. The `MosaicFTMachine` specialization of the `FTMachine` class implements the gridding and de-gridding algorithms for the mosaic imaging. It uses the w-projection technique (Cornwell et al. 2003) for gridding/de-gridding which is a more efficient algorithm. The `findConvFunction()` method of this class computes the gridding convolution function used by the underlying FORTRAN routine. This method needs to be overloaded in a specialization to pre-compute the functions E_{ij} for each baseline and the underlying FORTRAN routine modified to use the appropriate ones in the gridding/de-gridding operation. These services will then be used in the `SolvableVisJones::apply()` method to compute the corrupted visibilities. Note that `SolvableVisJones::` also needs to use these services, with a different convolution gridding function per baseline (Eq. 4). Pre-computation of all the convolution functions can be potentially memory demanding. Where memory usage needs to be traded for CPU time, Eq. 5 can be used to compute the approximate convolution functions. Since E_i^o is same for all the baselines, E_{ij} and its derivative need not be cached for each baseline separately.

5.1 Errors due to the approximations

Incorporation of antenna based pointing offsets in the equation implies a different gridding convolution function for every baseline (E_{ij} in Eq. 1) during de-gridding. This function needs to be computed for each minimization step making the algorithm compute intensive. A first order approximation can be used (Eqs. 4 and 5) for a faster computation of these functions on-the-fly for each baseline. For a pointing offset of 1% of the primary beam, the maximum amplitude and phase error due to the approximation is $\sim 4 \times 10^{-3}\%$ and 10^{-3} degree respectively (Fig. 7). However such an approximation will limit the level to which pointing errors can be solved. For observations which are limited by the errors due this approximation, Eq. 2 will need to be computed at a higher accuracy, possibly using look-up tables for the exponential function. E_{ij} is same for all the baselines, and can be computed once (possibly using measured primary beam shapes).

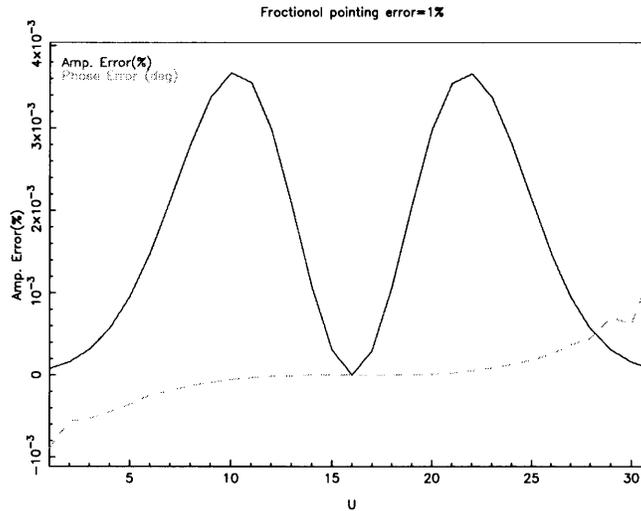


Figure 7: The percentage amplitude (the continuous curve) and phase (the dashed curve) errors across the aperture due to the approximations in Eq. 5 for a pointing error of 1% of the primary beam.

References

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