On–axis Instrumental Polarization Calibration for Circular Feeds

DRAFT W. D. Cotton, September 15, 2012

***Abstract*—Various instrumental and atmospheric effects corrupt the response of an interferometer to polarized signals. In the case of high dynamic range imaging, uncorrected, these effects can also degrade the total intensity image. These effects must be estimated and removed in order to produce images of the polarized emission, or high dynamic range total intensity images. This memo describes the reimplementation in Obit of the AIPS feed ellipticity–orientation modeling for arrays with circular feeds. Examples are given using wideband VLA data. While a substantial improvement in the quality of the polarized images is obtained, there appears to be a limit to the accuracy of the corrections to polarized images, perhaps as a result of time or geometry dependent instrumental polarization. Tests using only calibrators in the same part of the sky support this possibility.**

***Index Terms*—interferometry, polarization, calibration**

I. Introduction

RADIO interferometric imaging of polarized celestial emission provides a powerful probe of various physical processes as well as the propagation through intervening media. Instrumental and atmospheric effects corrupt the response of an interferometer to polarized signals. In particular, the array detectors do not respond to precisely the intended polarization state which leads to a spurious polarized response.

Detectors sensitive to the electric field of the incoming wave, as used in radio heterodyne systems, respond to a single polarization state; in order to fully measure the polarization of the wave, detectors measuring orthogonal polarization states are needed. In the correlation process, all four products of the two states at each antenna are produced. The most commonly used systems are right- and left-hand circular polarization and orthogonal linear polarizations. In practice, the detected polarizations are never precisely the desired ones. The exact polarization states detected must be determined in order to transform the measured visibilities into the corrected form.

This memo describes an implementation in the Obit package [1] [1] of the nonlinear ellipticity–orientation model used in the AIPS task PCAL. This discussion is only relevant for arrays using alt-az mounted antennas and using detectors (AKA "feeds") sensitive to orthogonal circular polarizations. Such arrays are the VLA and VLBA.

II. Interferometric Polarimetry

An arbitrary electromagnetic wave can be described as elliptically polarized, circular and linear polarization are extreme cases. One way of modeling the polarization state to which a detector responds is the ellipticity and orientation of the ellipse to which the detector responds.

An alternate, and popular approach is to model the response as the desired state plus a fraction of the orthogonal state. The fraction of the orthogonal state is referred to as the "leakage term". This model has the advantage that it can be linearized allowing for faster fitting. In the past, computation speed was a serious consideration and most popular polarization implementations used a linearized "d term" model. However, modern wideband systems, especially using circular feeds have relatively poor polarization purity while the higher sensitivity calls for higher accuracy. Wideband data with spectral resolution also exposes the variation of instrumental polarization with frequency; the modeling of the feed response also needs to be done on a frequency basis. See [2][3][4] for more detailed descriptions of the response to polarized radiation.

A. Effects of prior calibration

Discussions of instrumental polarization such as those given in [2] generally do not include the effects of prior calibration on the data. In particular, the practice of applying corrections for the parallactic angle prior to any other calibration is highly recommended yet has a profound effect on the model of the feed response. In the following, it is assumed that this correction has been made. For arrays with circularly polarized feeds, the common practice is to calibrate the two parallel circular systems of visibilities independently.

Since interferometers only measure differences of the phase of a wavefront between pairs of antennas, the absolute phase is undetermined and phases are referred to that at a reference antenna whose phase has been arbitrarily set to zero. This allows for an arbitrary phase and delay offset between the independent circular systems. See [5] for a discussion of correcting these offsets. After the independent phase calibration of the parallel hand systems the effects of the orientation of the reference feed is removed from the parallel hand systems but leaves the difference of the orientations of the two reference feeds in the cross hand products.

B. Response by Circular Feeds

For an interferometer with perfectly circular feeds observing an unresolved, partially polarized source and forming all four correlation products the expected visibility is given by

$$\mathbf{S} = [ipol + vpol,\ qpol + j\ upol,\ qpol - j\ upol,\ ipol - vpol] \quad (1)$$

where $ipol$, $qpol$, $upol$ and $vpol$ are the Stokes' I, Q, U and V of the source and j is $\sqrt{-1}$. In this case, the cross–polarized

National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA, 22903 USA email: bcotton@nrao.edu

[1]http://www.cv.nrao.edu/∼bcotton/Obit.html

components (RL, LR) give the linearly polarized response to the source. In practice, the feeds are never precisely circular and some fraction of the Stokes I "leaks" into the cross–polarized visibilities giving rise to a spurious "instrumental polarization".

The following adopts the "ORI-" model used in the AIPS task PCAL. Since this model has not been previously documented it is described in some detail here. This model describes each feed in terms of its ellipticity, θ and the the orientation of this ellipse, ϕ. The response of a given interferometer as a function of the parallactic angle, χ, and taking into account the previously applied calibration is given by a Muller matrix (4 $\times$ 4 complex matrix). This matrix multiplied by the true source polarization vector gives the model value of the visibilities. The parallactic angle given by:

$$\chi \;=\; \tan^{-1}\Big(\frac{\cos\,\lambda\,\sin\,h}{\sin\,\lambda\,\cos\,\delta\;-\;\cos\,\lambda\,\sin\,\delta\,\cos\,h}\Big) \quad (2)$$

where δ is the source declination, λ is the latitude of the antenna and h is the hour angle of the source.

A Jones matrix which includes the effects of calibration can be constructed for each feed, the elements are:

$$\begin{aligned}
J_{00} =& \frac{1}{\sqrt{2}}\;(cos(\theta_{R\,i}) + sin(\theta_{R\,i}))e^{j(\phi_{R\,ref})}\\
J_{01} =& \frac{1}{\sqrt{2}}\;(cos(\theta_{R\,i}) - sin(\theta_{R\,i}))\;e^{2j\;\phi_{R\,i}}\times\\
& e^{-2j\;\chi_i}\;e^{j\phi_{R\,ref}}\\
J_{10} =& \frac{1}{\sqrt{2}}\;(cos(\theta_{L\,i}) + sin(\theta_{L\,i}))\;e^{-2j\;\phi_{L\,i}}\times\\
& e^{+2j\chi_i}e^{-j(\phi_{L\,ref}+PD)}\\
J_{11} =& \frac{1}{\sqrt{2}}\;(cos(\theta_{L\,i}) - sin(\theta_{L\,i}))\;e^{-j(\phi_{L\,ref}+PD)}
\end{aligned}$$

and the Jones matrix is:

$$\mathbf{J}_i = \begin{vmatrix} J_{00} & J_{01} \\ J_{10} & J_{11} \end{vmatrix}$$

The Muller matrix, M_{ik}, for baseline i-k is then the outer product of $\mathbf{J}_i$ and $\mathbf{J}_k^*$.

$$\mathbf{M}_{ik} \;=\; \mathbf{J}_i \;\otimes \mathbf{J}_k^*$$

Computation of the Muller matrix is described in more detail in the Appendix. Note: the formulation given above for the Jones matrix assumes that the parallactic angle corrections have been previously applied.

The predicted correlation vector is then

$$\mathbf{V}_{model\,ik} \;=\; \mathbf{M}_{ik}\;\mathbf{S} \quad (3)$$

where

$$\mathbf{V}^T_{model\,ik} \;=\; [RR_{ik},\;RL_{ik},\;LR_{ik},\;LL_{ik}],$$

are the four correlations of the R and L detectors.

C. Source and instrumental polarization

In the general case, the polarization of the calibrator is unknown and must be solved for jointly with the instrumental polarization. For arrays with alt-az antenna mounts, the variation in the parallactic angle at which a calibrator is observed introduces a different effect on the source and instrumental polarization. Since the instrumental polarization is introduced in the frame of the antenna, it is constant with parallactic angle. For antennas using circularly polarized feeds, the source polarization will rotate with parallactic angle. After the data are corrected for the effects of parallactic angle, the reverse is true. To separate the two contributions to the polarized response requires that at least one calibrator be observed over a sufficient range of parallactic angle to separate the two effects. How much is enough depends on the SNR but generally a radian or more is desirable. For a calibrator of known polarization, including none, any distribution of parallactic angle is usable.

III. Solutions for Model Parameters

The parameters of the model described in Section II-B can be fitted to a data set using any of a number of techniques. In the following two such techniques are explored.

A. Relaxation fitting

The first technique considered is a relaxation technique in which corrections are cyclically made to each model parameter being fitted by determining the effect of that parameter on the χ^2. The χ^2 of the fit is defined as

$$\chi^2 = \sum_{i=0}^{n}(model_i - obs_i)^2/\sigma_i^2$$

where $model_i$ is the model value for observation i (real or imaginary part of visibility), obs_i is the observation for i and σ_i^2 is the variance of the amplitude of observation i. The model values are components of the vector V_{model} given by Eq. 3. For each parameter, P, in each iteration, n, the revised value is given by:

$$P_{n+1} = P_n \;+\; tan^{-1}\Big(\frac{\frac{\partial\chi^2}{\partial P}}{\frac{\partial^2\chi^2}{\partial P^2}}\Big)$$

subject to the constrain that the corrected value, P_{n+1}, leads to a decrease in χ^2. Should this not be the case, the correction is reduced in magnitude until the χ^2 decreases. Note: the evaluation of χ^2 only need consider data which depend on the given parameter. The various partial derivatives are given in the appendix. In the following, this is referred to as the "fast" technique.

B. General nonlinear least squares

Most scientific software packages provide a generalized nonlinear least squares facility such as the Levenberg-Marquardt least squares technique. These can provide robust solutions as well as an error analysis but can be expensive to compute. An initial estimate of the parameters using the

relaxation technique can reduce the total computing cost. These generalised nonlinear least squares packages generally need the derivatives of either the χ^2 or the model wrt the various parameters; these derivatives are given in the appendix.

IV. Correcting Observed Visibilities

Once the model parameters describing the instrumental polarization are known, they can be used to remove the instrumental effects from the data. The corrected visibility vector V_{corr} can be obtained from the observed visibility vector V_{obs} by

$$\mathbf{V}_{cor} \;=\; \mathbf{M}^{-1}\,\mathbf{V}_{obs} \tag{4}$$

where M^{-1} is the inverse of the Muller matrix. A useful property of the outer product is that the the outer product of the inverses of two Jones matrices is the inverse of the outer product on the Jones matrices themselves. The inverse of a 2×2 matrix given by:

$$\mathbf{J} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

is

$$\mathbf{J}^{-1} = \frac{1}{ad-bc} \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}$$

V. Obit implementation: PCal

Implemention of this calibration technique in Obit is in the task PCal which uses the software class ObitPolnCalFit. Calibration is per channel or sliding window of channels. Fitted feed parameters are stored in an AIPS PD table with the "Real" part being the ellipticity and the "Imaginary" being orientation. The table is labeled as type "ORI-ELP".

Fitting always uses the "fast" Relaxation method optionally followed by a Levenberg-Marquardt least squares using the GSL package nonlinear fitter. The initial value of the parameters are for perfect feeds for the first channel fitted and the results of the previous channel in subsequent fittings.

The parameters to PCal allow specifying the source polarizations for some subset of the calibrators in terms of the fractional polarization, the EVPA (as the R-L phase difference) and the rotation measure. In this case, the right–left phase offset in the data should also be solved for and the results stored in the AIPS PD table as well as an AIPS BP table.

Application of instrumental polarization corrections to the data computes the inverse Muller matrix from the outer product of the inverse of the Jones matrices. The antenna Jones matrices as a function of parallactic angle are derived as described in Section II-B.

VI. Examples with VLA data

A. Single Unpolarized Calibrator

A test of the technique was made using a VLA wideband dataset consisting of ≈10 min on the strong, weakly polarized calibrator 3C84. The data were in C band (5 GHz) and consisted of 2 IFs of 64 2 MHz channels. Data were calibrated using Obit VLA calibration scripts [6] assuming 1 Jy for the flux density of the source. End channels in each IF were flagged. Right–left delay calibration using Obit task RLDly [5] could not provide an adequate solution as the cross–polarizations are strongly dominated by instrumental polarization which is not coherent amoung all antennas. Spectra for "typical" and "poor" baselines are shown in Figures 1 and 2.

Task PCal was used both with only the "fast" solutions as well as including the Levenberg-Marquardt least squares fitting to compare the results. The cpu time on a 2.2 GHz processor was 178 sec for the "fast" calibration and 541 sec for the Levenberg-Marquardt fitting. The fitting did not solve for the source polarization which was assumed to be zero. The calibration averaged data to 5 minutes and solved for each channel independently. The data were then corrected and averaged to 2 minutes. Plots of the two baselines using the "fast" calibration are given in Figures 3 and 4. Data calibrated using the Levenberg-Marquardt fits are shown in Figures 5 and 6.

B. Multiple Polarized Calibrators

In general it is necessary to determine the instrumental polarization from observations of one or more partly polarized sources, the polarization state of which may be unknown. The technique is to use the different effect of parallactic angle on source and instrumental polarizations to separate the components. A second set of tests were performed on a wideband set of VLA observations covering from 6 to 8 GHz using three calibrators, J1331+3030 (=3C286), J1504+1029 and J1651+0129 with approximate flux densities of ∼5, 1. and 0.5 Jy respectively. These were observed in a number of ∼15 sec scans over approximately 7.5 hours; J1331+3030 5 times, J1504+1029 7 times and J1651+0129 32 times. The parallactic angle range covered was approximately + to - 1 radian. The VLA was in the "A" configuration so the image resolution is 0.3". Total intensity calibration used the standard Obit EVLA calibration scripts including right–left delay calibration. A final pass at calibration used self calibrated models of each of the calibrators and data were edited for outlyers from the self cal model visibilities.

Several calibrations were tested, the first fitting for the source polarizations as well as the instrumental terms in blocks of 5 channels (10 MHz) using the "fast" method. After the application of this calibration, the right-left phase was adjusted (Obit/RLPass) using the J1331+3030 data for which the value is known to be 66 degrees over at least the frequency range of 1 to 100 GHz. This is called calibration 1 in the following. This calibration took 1.6 hours with some speedup from multi–threading.

The first technique has the disadvantage that the source polarizations are allowed to vary arbitrarily with frequency Therefore, the second calibration was to fix the source polarizations to a given fraction of the total intensity and a constant right-left phase per source. J1331+3030 is strongly polarized and is known to have essentially zero rotation measure. The other calibrators did not seem to vary over the frequency range in the initial calibration. In this calibration, the right-left phase of the data was also determined in the fit. This is called calibration 2 in the following. The polarization models are shown in Table I. This used the "fast" method and took 2.3 hours with multi threading.

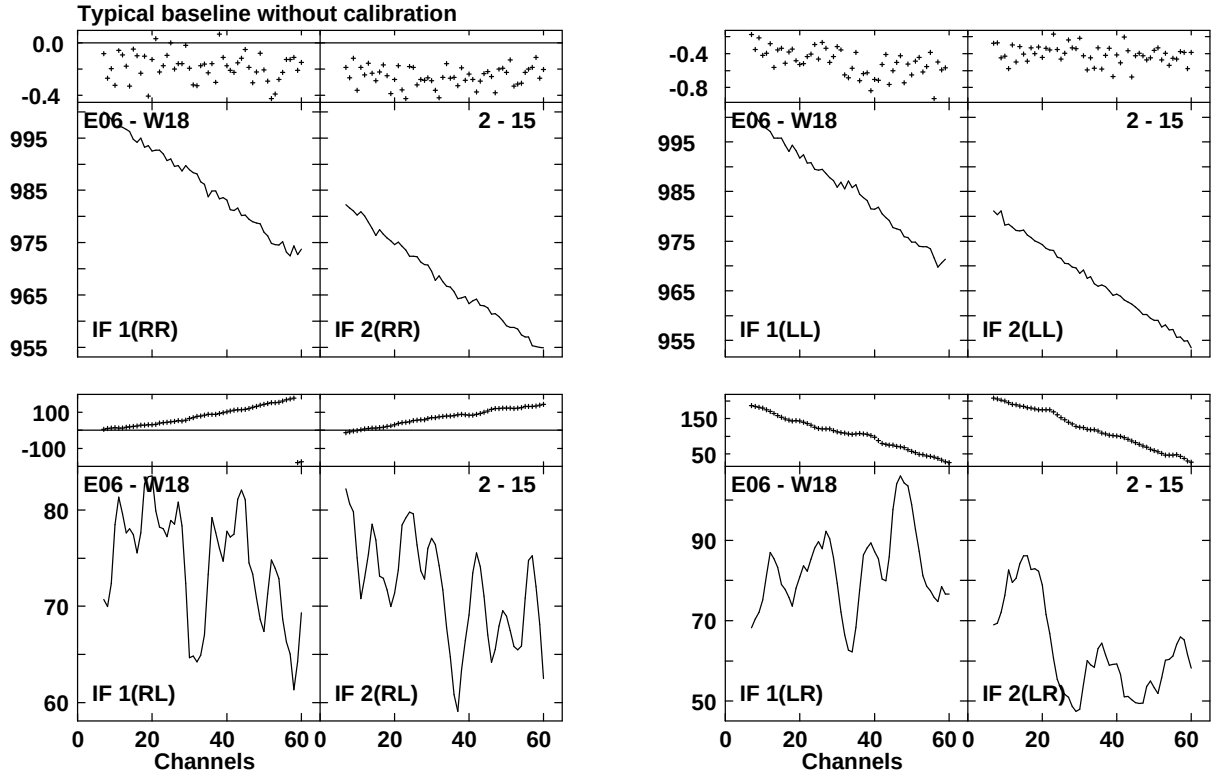

Fig. 1. Visibility spectra for all correlation products on a typical baseline of data on 3C84 without polarization calibration. Parallel products above, cross below; in each panel phases are on the top in degrees and below are amplitudes in units of 1/1000 of the total flux density. Cross hand spectra show residual cross hand group delay errors.

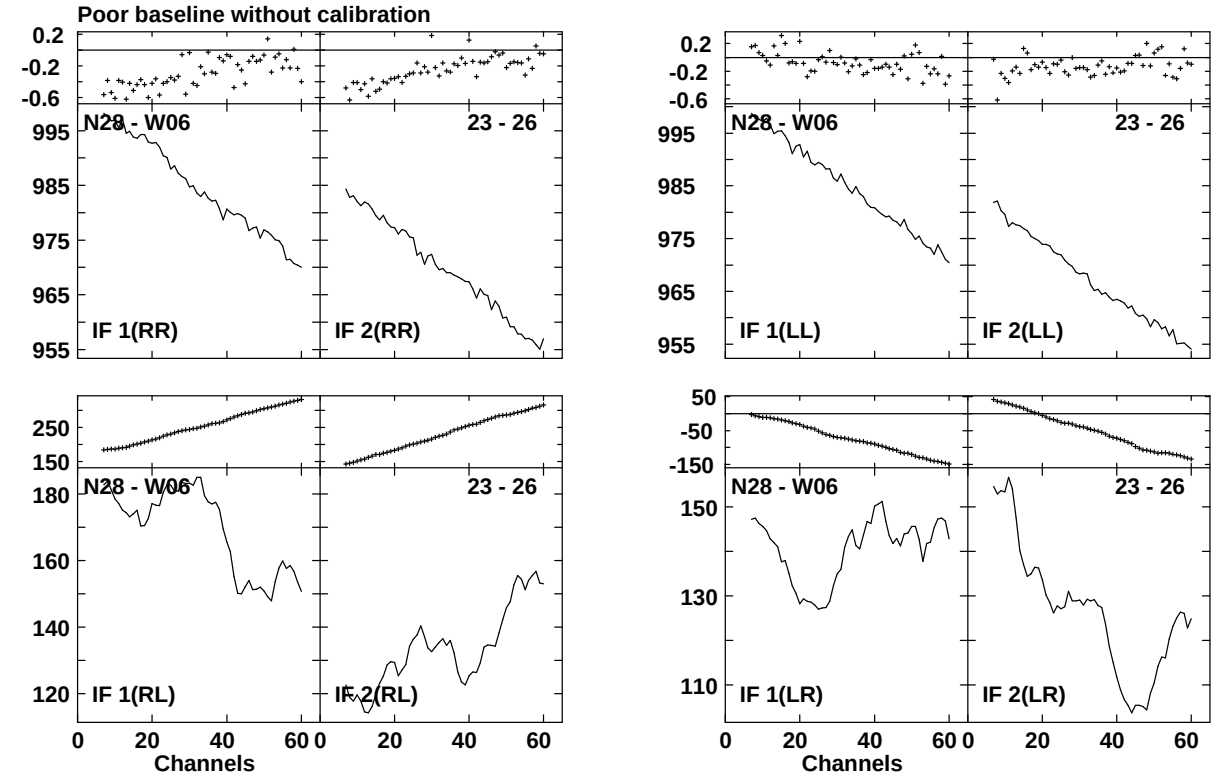

Fig. 2. Like Figure 1 but for a poorly behaved baseline.

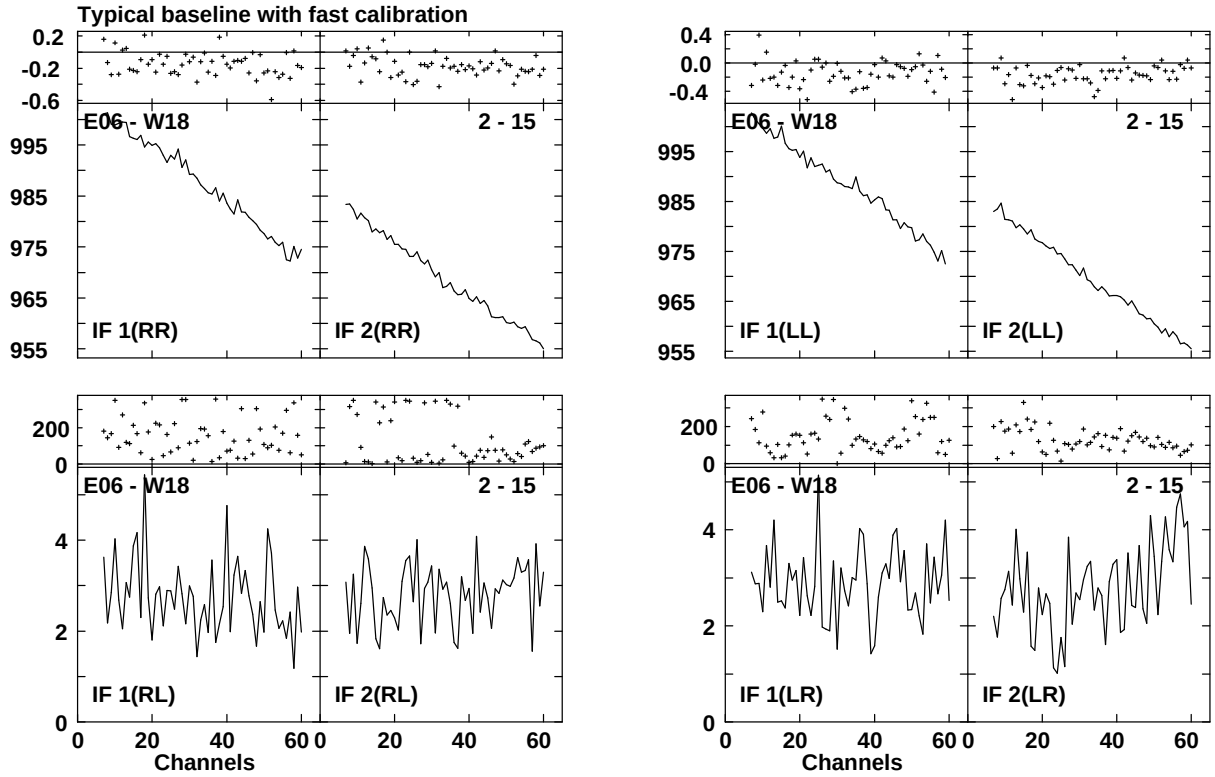

Fig. 3. Like Figure 1 but for a poorly behaved baseline but after applying the calibration from the "fast" method and averaging to 2 minutes.

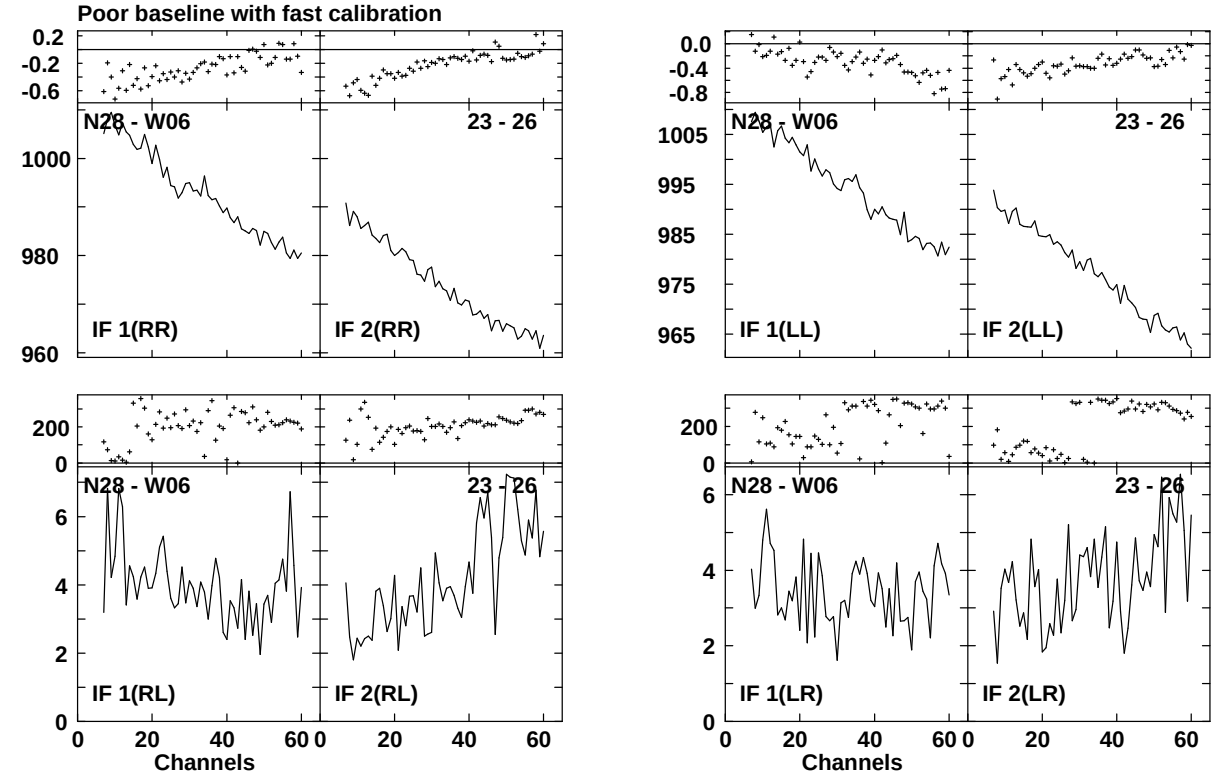

Fig. 4. Like Figure 1 but for a poorly behaved baseline after applying the calibration from the"fast" method and averaging to 2 minutes.

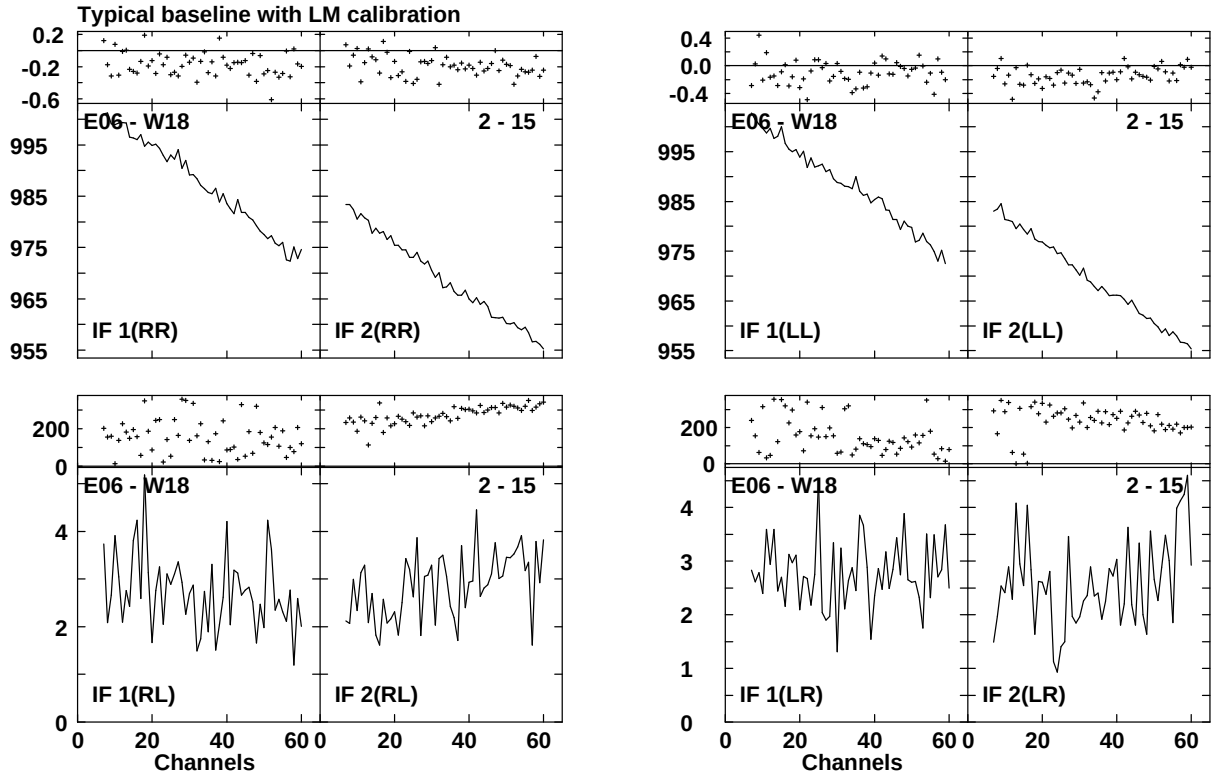

Fig. 5. Like Figure 1 after applying the calibration from the Levenberg-Marquardt least squares method and averaging to 2 minutes.

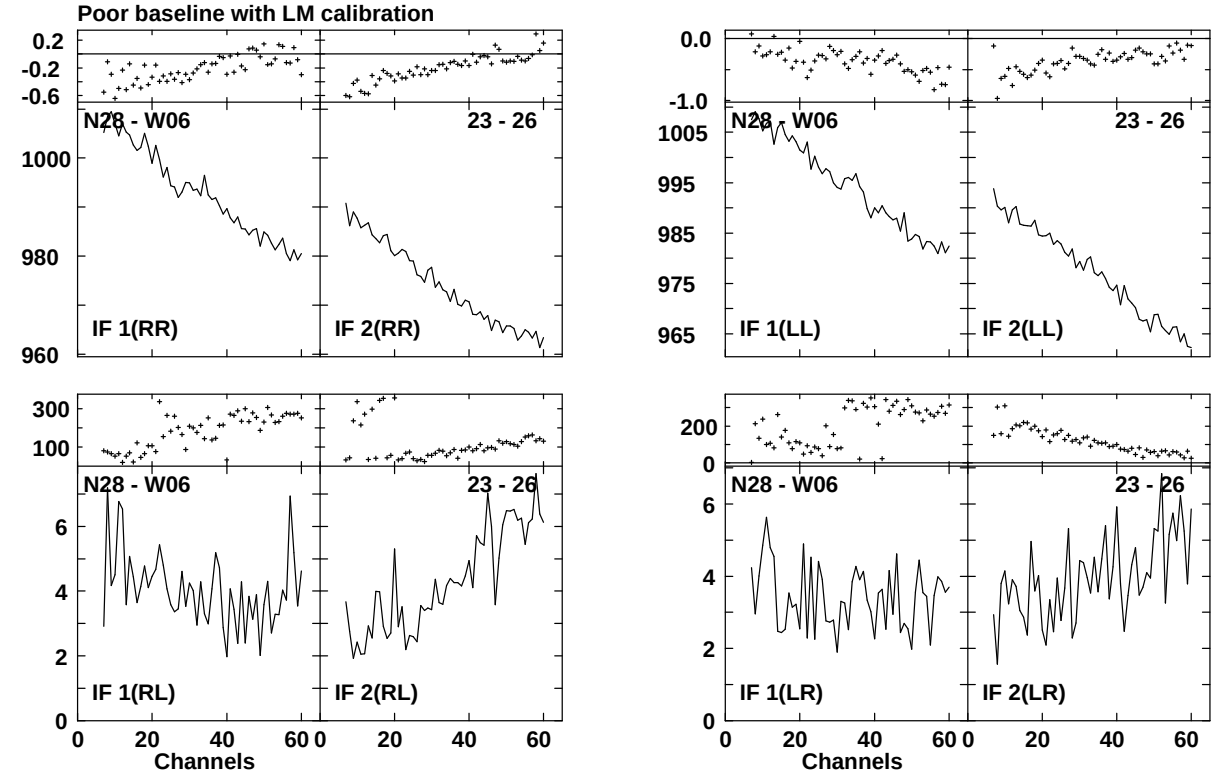

Fig. 6. Like Figure 1 but for a poorly behaved baseline after applying the calibration from the Levenberg-Marquardt least squares method and averaging to 2 minutes.

As a third test, the previous calibration was rerun using the expensive Levenberg-Marquardt ("LM") method; this was calibration 3. This took 15.2 hours and while it was multithreaded, the GSL least squares routines were run single threaded.

J1331+3030 is in a different part of the sky from the two calibrators used to evaluate the accuracy of the calibration. Any geometry dependent effects will cause the use of this calibrator might skew the instrumental solutions. As a final test (calibration 4) the previous calibration was rerun without J1331+3030 (16.8 hr run time). This results in a significant reduction in the level of polarization artifacts.

J1331+3030 has strong enough polarized emission that single baseline spectra are sufficient to show the efficacy of the calibration. The spectra from a poorly behaved baseline without and with two of the calibrations are shown in Figures 7 – 9.

The data for J1504+1029 and J1651+0129 were imaged without and with each of the polarization calibrations. Note: the calibration procedure corrects the data for the effects of parallactic angle which allows for imaging polarized data without the correction for instrumental polarization. Imaging used the Obit wideband imager MFImage using the 16 IFs as the coarse frequency bins and with phase and delay self calibration followed by amplitude and phase self calibration. The selected images are shown in Figures 10 and 11. The RMS in a box near the source was measured in each image and the results given in Table II.

VII. Discussion

The technique works well in the first test case reducing the 5–20% instrumental terms to under ≈0.5%, possibly limited by the actual source polarization. In this test, results for the "fast" solution method are comparable to the Levenberg-Marquardt fitting and ran in about 1/3 of the cpu time (176 v. 541 sec.).

The second test (Table II) showed that application of polarization calibration improved the near source Stokes I residuals by about 10% for the two calibrators for source dynamic ranges of 40,000-50,000:1. The calibration (1) fitting the source polarizations in each of the solution blocks reduced the near source polarized RMS by roughly a factor of 5 and fixing the source polarizations using the "LM" fitting reduced the RMS by approximately a factor of 8. The "LM" solution (calibration 3) took 7 times longer (15.2 v. 2.3 hours) but reduced the polarized artifacts in this test by about 15% over the relaxation only fitting.

In all cases, the polarized images had visible calibration artifacts and the near source Q and U RMSes were approximately 50% higher than the corresponding Stokes I. It is unclear what is responsible for the residual instrumental polarization but variability with time and/or observing geometry would produce such an effect. The result of the fourth test supports the possibility that the instrumental polarization is direction independent; when onlyt the two calibrators used to evaluate the calibration were used in the calibrtation, the results were significantly better (24%) than when J1331+3030 was included in the calibration. The two calibrators involved are fairly close on the sky.

TABLE I
Calibrator polarization models

Source	frac. pol	R-L phase °
J1331+3030	0.122	66.0
J1504+1029	0.020	12.0
J1651+0129	0.036	-135.0

TABLE II
Calibrator polarization RMS

Source	Cal	I RMS μJy/bm	Q RMS μ/bm	U RMS μ/bm
J1504+1029	no Cal	26.9	288.	358.
J1504+1029	Cal. 1	23.7	57.2	58.8
J1504+1029	Cal. 2	23.7	42.8	46.7
J1504+1029	Cal. 3	23.7	36.1	41.2
J1504+1029	Cal. 4	23.6	28.9	31.8
J1651+0129	no Cal	11.3	106.	118.
J1651+0129	Cal. 1	10.2	22.9	19.9
J1651+0129	Cal. 2	10.2	16.5	17.3
J1651+0129	Cal. 3	10.2	14.1	15.6
J1651+0129	Cal. 4	10.2	10.7	11.0

Acknowledgment

I would like to thank Fred Schwab for help with the mathematics.

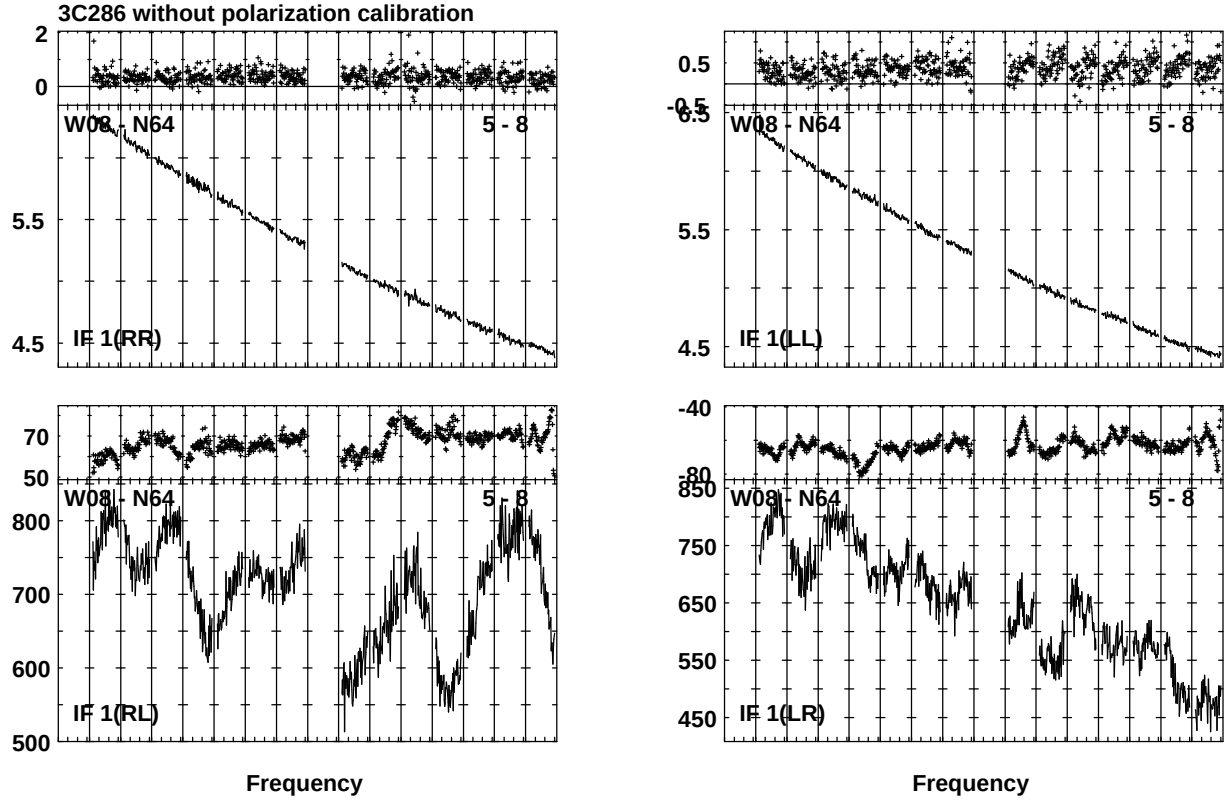

Fig. 7. Averaged spectrum of 3C286 on a single troublesome baseline without polarization calibration. Parallel products above, cross below; in each panel phases are on the top in degrees and below are amplitudes in Jy for RR and LL and mJy for LR and RL.

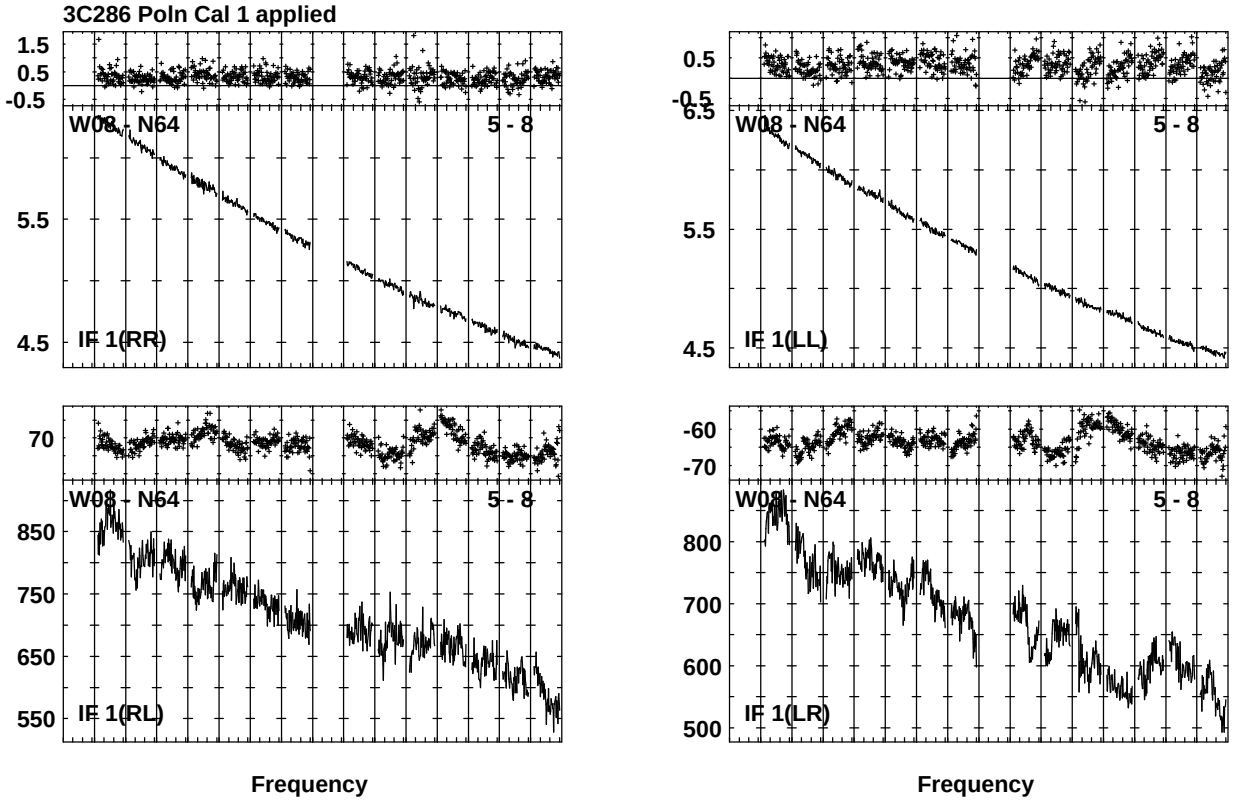

Fig. 8. Like Figure 7 but after the application of polarization calibration 1.

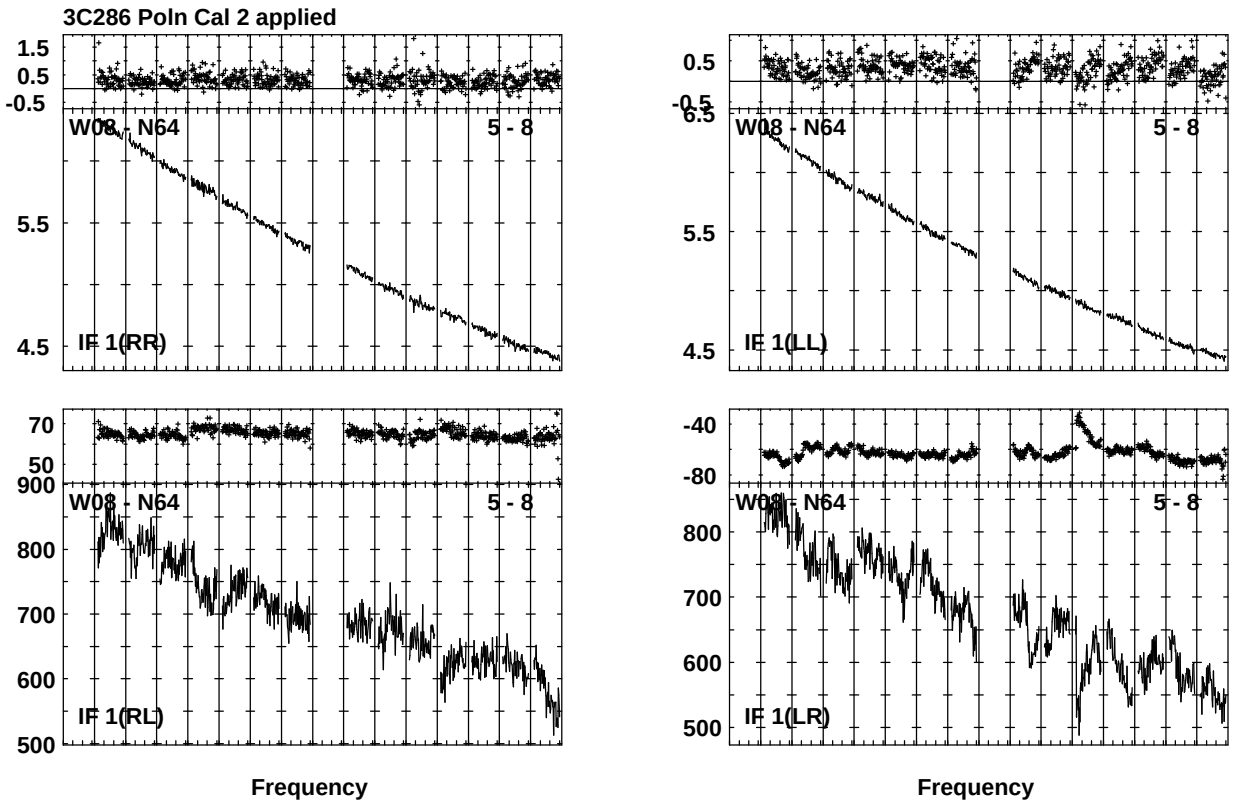

Fig. 9. Like Figure 7 but after the application of polarization calibration 2.

APPENDIX

Muller matrix

In order to compute the Muller matrix for a given baseline, for each antenna *i* and the reference antenna *ref* define

$$
\begin{aligned}
S_{Ri} &= cos(\theta_{Ri}) + sin(\theta_{Ri}) \\
D_{Ri} &= cos(\theta_{Ri}) - sin(\theta_{Ri}) \\
S_{Li} &= cos(\theta_{Li}) + sin(\theta_{Li}) \\
D_{Li} &= cos(\theta_{Li}) - sin(\theta_{Li}) \\
P_{Ri} &= e^{j\ 2\ \phi_{Ri}} \\
P_{Li} &= e^{-j\ 2\ \phi_{Li}} \\
\sigma_{Ri} &= \frac{1}{\sqrt{2}}\ S_{Ri} \\
\delta_{Ri} &= \frac{1}{\sqrt{2}}\ D_{Ri}\ P_{Ri} \\
\sigma_{Li} &= \frac{1}{\sqrt{2}}\ S_{Li}\ P_{Li} \\
\delta_{Li} &= \frac{1}{\sqrt{2}}\ D_{Li} \\
PA_i &= e^{-i\ 2\ \chi_i} \\
R_{Rref} &= e^{j\ \phi_{Rref}} \\
R_{Lref} &= e^{-j\ (\phi_{Lref}+PD)} \\
\psi_{RL} &= R_{Rref}\ R^*_{Lref} \\
\psi_{LR} &= R_{Lref}\ R^*_{Rref}
\end{aligned}
$$

where θ is the ellipticity, ϕ the orientation, χ the parallactic angle, *PD* is the right-left phase difference of the reference antenna and * denotes the complex conjugate.

The elements of the Muller matrix are then given by:

$$
\begin{aligned}
M_{00} &= \sigma_{R1}\ \sigma^*_{R2} \\
M_{01} &= \sigma_{R1}\ \delta^*_{R2}\ PA^*_2 \\
M_{02} &= \delta_{R1}\ \sigma^*_{R2}\ PA_1 \\
M_{03} &= \delta_{R1}\ \delta^*_{R2}\ PA_1\ PA^*_2 \\
M_{10} &= \psi_{RL}\ \sigma_{R1}\ \sigma^*_{L2}\ PA_2 \\
M_{11} &= \psi_{RL}\ \sigma_{R1}\ \delta^*_{L2} \\
M_{12} &= \psi_{RL}\ \delta_{R1}\ \sigma^*_{L2}\ PA_1\ PA_2 \\
M_{13} &= \psi_{RL}\ \delta_{R1}\ \delta^*_{L2}\ PA_1 \\
M_{20} &= \psi_{LR}\ \sigma_{L1}\ \sigma^*_{R2}\ PA^*_1 \\
M_{21} &= \psi_{LR}\ \sigma_{L1}\ \delta^*_{R2}\ PA^*_1\ PA^*_2 \\
M_{22} &= \psi_{LR}\ \delta_{L1}\ \sigma^*_{R2} \\
M_{23} &= \psi_{LR}\ \delta_{L1}\ \delta^*_{R2}\ PA^*_2 \\
M_{30} &= \sigma_{L1}\ \sigma^*_{L2}\ PA^*_1\ PA_2 \\
M_{31} &= \sigma_{L1}\ \delta^*_{L2}\ PA^*_1 \\
M_{32} &= \delta_{L1}\ \sigma^*_{L2}\ PA_2 \\
M_{33} &= \delta_{L1}\ \delta^*_{L2}
\end{aligned}
\tag{5}
$$

and the Muller matrix is

$$
\begin{matrix}
M_{00}\ M_{01}\ M_{02}\ M_{02} \\
M_{10}\ M_{11}\ M_{12}\ M_{12} \\
M_{20}\ M_{21}\ M_{22}\ M_{22} \\
M_{30}\ M_{31}\ M_{32}\ M_{32}
\end{matrix}
$$

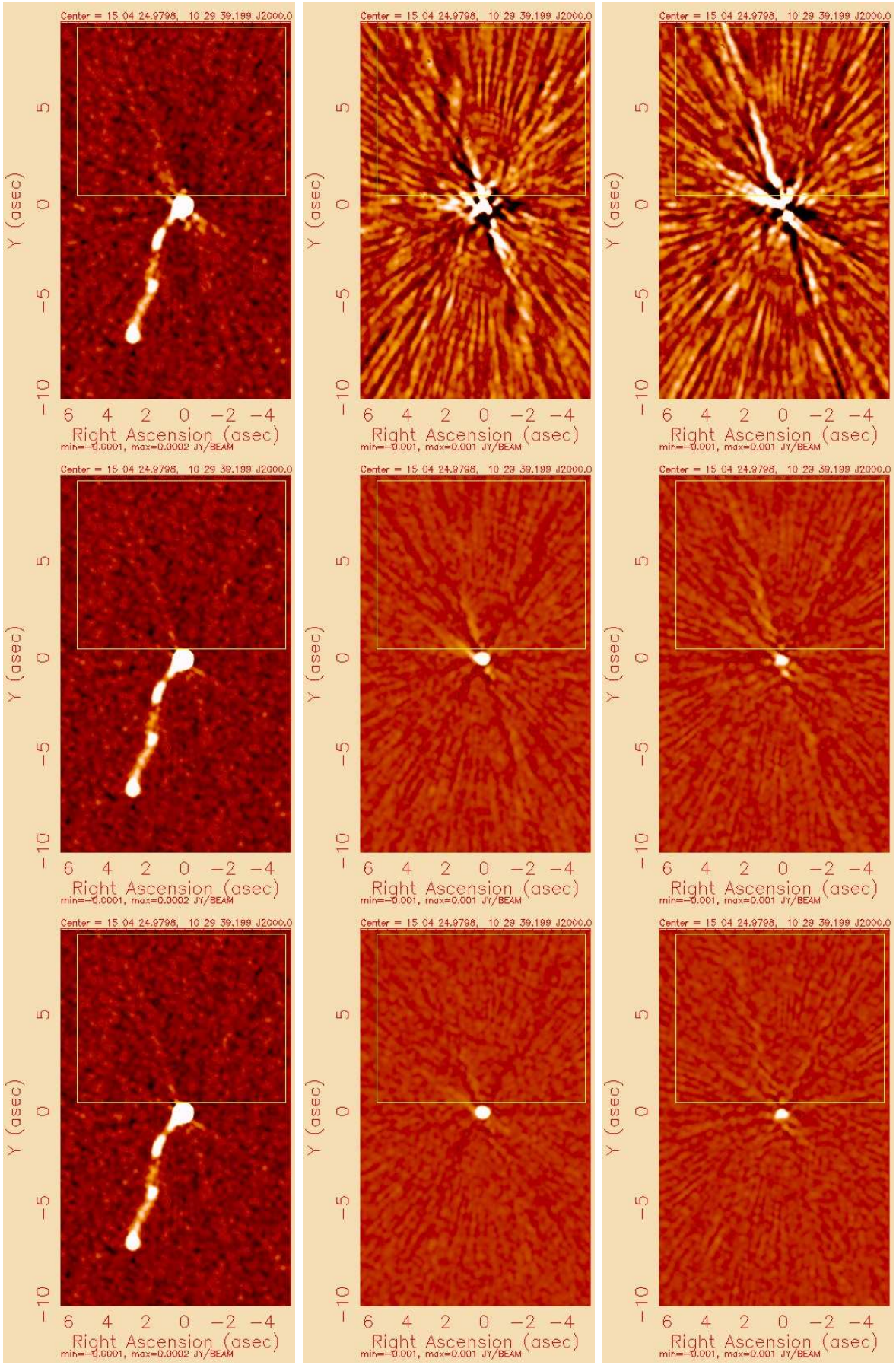

Fig. 10. Images of the calibrator J1504+1029. On the left is Stokes I clipped at -100 and +200 μJy/bm, in the center is Stokes Q and right Stokes U both clipped to between -1 and +1 mJy/bm. Top row is without polarization calibration, the middle row is with polarization calibration 1 and the bottom row polarization calibration 3. The yellow box indicates the region used to measure the off-source RMSes given in Table II.

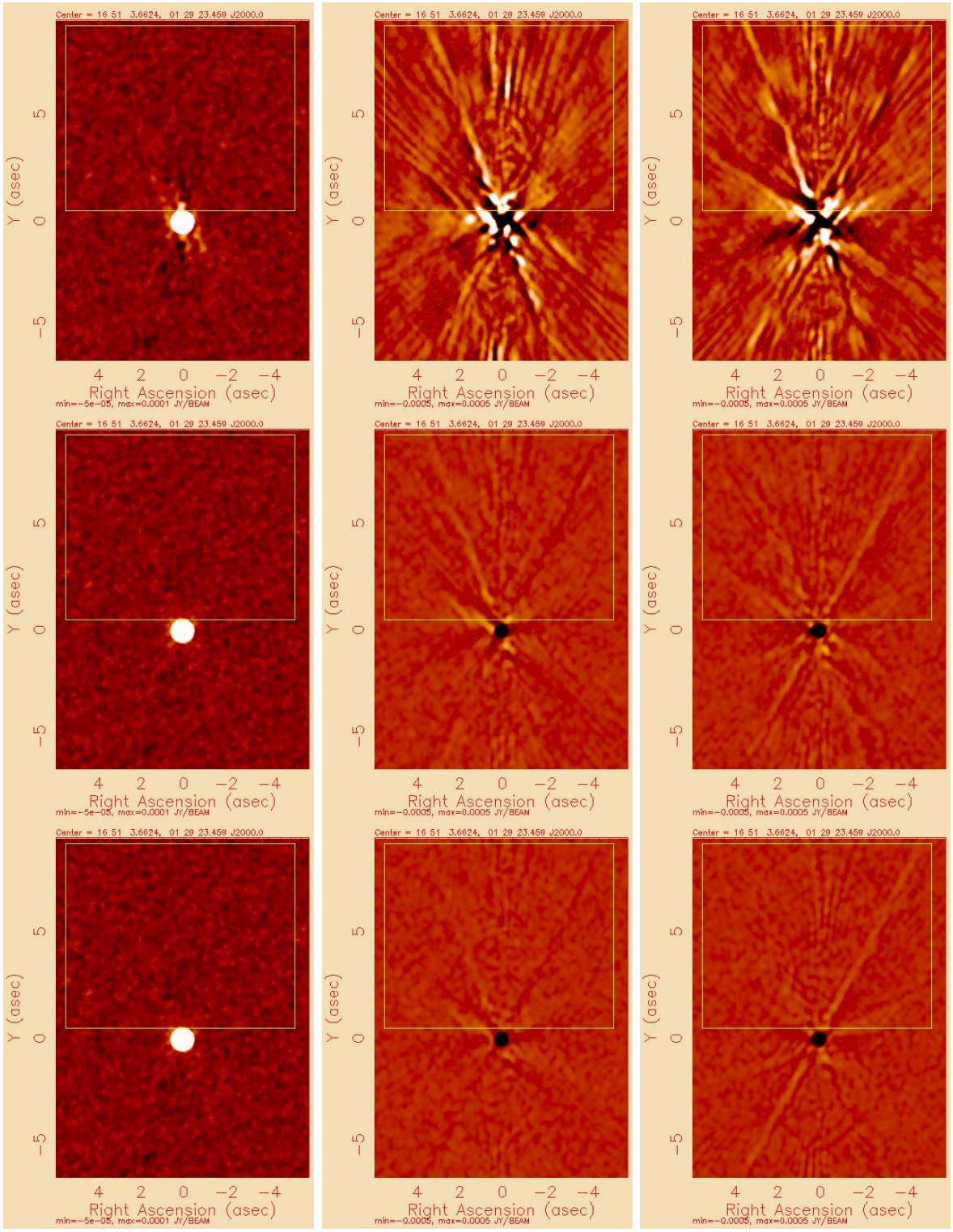

Fig. 11. Images of the calibrator J1651+0129. On the left is Stokes I clipped at -50 and +100 μJy/bm, in the center is Stokes Q and right Stokes U both clipped to between -0.5 and +0.5 mJy/bm. Top row is without polarization calibration, the middle row is with polarization calibration 1 and the bottom row polarization calibration 3. The yellow box indicates the region used to measure the off-source RMSes given in Table II.

Derivatives of model wrt parameters

Source terms:

$$\begin{aligned}
\frac{\partial RR}{\partial ipol} &= M_{00} + M_{03}, & \frac{\partial^2 RR}{\partial ipol^2} &= 0\\
\frac{\partial RR}{\partial qpol} &= M_{01} + M_{02}, & \frac{\partial^2 RR}{\partial qpol^2} &= 0\\
\frac{\partial RR}{\partial upol} &= j\ (M_{01} + M_{02}), & \frac{\partial^2 RR}{\partial upol^2} &= 0\\
\frac{\partial RR}{\partial vpol} &= M_{00} - M_{03}, & \frac{\partial^2 RR}{\partial vpol^2} &= 0\\
\frac{\partial RL}{\partial ipol} &= M_{10} + M_{13}, & \frac{\partial^2 RL}{\partial ipol^2} &= 0\\
\frac{\partial RL}{\partial qpol} &= M_{11} + M_{12}, & \frac{\partial^2 RL}{\partial qpol^2} &= 0\\
\frac{\partial RL}{\partial upol} &= j\ (M_{11} + M_{12}), & \frac{\partial^2 RL}{\partial upol^2} &= 0\\
\frac{\partial RL}{\partial vpol} &= M_{10} - M_{13}, & \frac{\partial^2 RL}{\partial vpol^2} &= 0\\
\frac{\partial LR}{\partial ipol} &= M_{20} + M_{23}, & \frac{\partial^2 LR}{\partial ipol^2} &= 0\\
\frac{\partial LR}{\partial qpol} &= M_{21} + M_{22}, & \frac{\partial^2 LR}{\partial qpol^2} &= 0\\
\frac{\partial LR}{\partial upol} &= j\ (M_{21} + M_{22}), & \frac{\partial^2 LR}{\partial upol^2} &= 0\\
\frac{\partial LR}{\partial vpol} &= M_{20} - M_{23}, & \frac{\partial^2 LR}{\partial vpol^2} &= 0\\
\frac{\partial LL}{\partial ipol} &= M_{30} + M_{33}, & \frac{\partial^2 LL}{\partial ipol^2} &= 0\\
\frac{\partial LL}{\partial qpol} &= M_{31} + M_{32}, & \frac{\partial^2 LL}{\partial qpol^2} &= 0\\
\frac{\partial LL}{\partial upol} &= j\ (M_{31} + M_{32}), & \frac{\partial^2 LL}{\partial upol^2} &= 0\\
\frac{\partial LL}{\partial vpol} &= M_{30} - M_{33}, & \frac{\partial^2 LL}{\partial vpol^2} &= 0
\end{aligned}$$

Non zero terms in ϕ_R:

$$\begin{aligned}
\frac{\partial RR}{\partial \phi_{R\,1}} &= -2\ j\ (S_2\ M_{02} + S_3 M_{03}), & \frac{\partial^2 RR}{\partial \phi_{R\,1}^2} &= -2\ j\ \frac{\partial RR}{\partial \phi_{R\,1}}\\
\frac{\partial RR}{\partial \phi_{R\,2}} &= +2\ j\ (S_1\ M_{01} + S_3 M_{03}), & \frac{\partial^2 RR}{\partial \phi_{R\,2}^2} &= +2\ j\ \frac{\partial RR}{\partial \phi_{R\,2}}\\
\frac{\partial RL}{\partial \phi_{R\,1}} &= -2\ j\ (S_2\ M_{12} + S_3 M_{13}), & \frac{\partial^2 RL}{\partial \phi_{R\,1}^2} &= -2\ j\ \frac{\partial RL}{\partial \phi_{R\,1}}\\
\frac{\partial LR}{\partial \phi_{R\,2}} &= +2\ j\ (S_1\ M_{21} + S_3 M_{23}), & \frac{\partial^2 LR}{\partial \phi_{R\,2}^2} &= +2\ j\ \frac{\partial LR}{\partial \phi_{R\,2}}
\end{aligned}$$

If one of the antennas is the reference antenna the following terms need be added:

$$\begin{aligned}
\frac{\partial RL}{\partial \phi_{R\,1}} &+= +\ j\ V_{RL}, & \frac{\partial^2 RL}{\partial \phi_{R\,1}^2} &+= +\ j\ \frac{\partial RL}{\partial \phi_{R\,1}}\\
\frac{\partial LR}{\partial \phi_{R\,2}} &+= -\ j\ V_{LR}, & \frac{\partial^2 LR}{\partial \phi_{R\,2}^2} &+= -\ j\ \frac{\partial LR}{\partial \phi_{R\,2}}
\end{aligned}$$

where V_{RL} and V_{LR} are the model values of the RL and LR correlations.

Non zero terms in ϕ_L:

$$\begin{aligned}
\frac{\partial LL}{\partial \phi_{L\,1}} &= +2\ j\ (S_0\ M_{30} + S_1 M_{31}), & \frac{\partial^2 LL}{\partial \phi_{L\,1}^2} &= +2\ j\ \frac{\partial LL}{\partial \phi_{L\,1}}\\
\frac{\partial LL}{\partial \phi_{L\,2}} &= -2\ j\ (S_0\ M_{30} + S_2 M_{32}), & \frac{\partial^2 LL}{\partial \phi_{L\,2}^2} &= -2\ j\ \frac{\partial LL}{\partial \phi_{L\,2}}\\
\frac{\partial RL}{\partial \phi_{L\,2}} &= +2\ j\ (S_0\ M_{10} + S_2 M_{12}), & \frac{\partial^2 RL}{\partial \phi_{L\,2}^2} &= +2\ j\ \frac{\partial RL}{\partial \phi_{L\,2}}\\
\frac{\partial LR}{\partial \phi_{L\,1}} &= -2\ j\ (S_0\ M_{20} + S_1 M_{21}), & \frac{\partial^2 LR}{\partial \phi_{L\,1}^2} &= -2\ j\ \frac{\partial LR}{\partial \phi_{L\,1}}
\end{aligned}$$

If one of the antennas is the reference antenna the following terms need be added:

$$\begin{aligned}
\frac{\partial LR}{\partial \phi_{L\,1}} &+= -\ j\ V_{LR}, & \frac{\partial^2 LR}{\partial \phi_{L\,1}^2} &+= -\ j\ \frac{\partial LR}{\partial \phi_{L\,1}}\\
\frac{\partial LR}{\partial \phi_{L\,2}} &+= +\ j\ V_{RL}, & \frac{\partial^2 LR}{\partial \phi_{L\,2}^2} &+= +\ j\ \frac{\partial LR}{\partial \phi_{L\,2}}
\end{aligned}$$

Non zero terms in θ_R:

$$\frac{\partial RR}{\partial \theta_{R\,1}} = \frac{1}{\sqrt{2}} D_{R\,1}(S_0 \sigma^*_{R\,2} + S_1 \delta^*_{R\,2} PA^*_2) - \frac{1}{\sqrt{2}} S_{R\,1} P_{R\,1}(S_2 \sigma^*_{R\,2} PA_1 + S_3 \delta^*_{R\,2} PA_1 PA^*_2)$$

$$\frac{\partial^2 RR}{\partial \theta^2_{R\,1}} = -\frac{1}{\sqrt{2}} S_{R\,1}(S_0 \sigma^*_{R\,2} + S_1 \delta^*_{R\,2} PA^*_2) - \frac{1}{\sqrt{2}} SDR\,1 P_{R\,1}(S_2 \sigma^*_{R\,2} PA_1 + S_3 \delta^*_{R\,2} PA_1 PA^*_2)$$

$$\frac{\partial RR}{\partial \theta_{R\,2}} = \frac{1}{\sqrt{2}} D_{R\,2}(S_0 \sigma_{R\,1} + S_2 \delta_{R\,1} PA_1) - \frac{1}{\sqrt{2}} S_{R\,2} P^*_{R\,2}(S_1 \sigma_{R\,1} PA^*_2 + S_3 \delta_{R\,1} PA_1 PA^*_2)$$

$$\frac{\partial^2 RR}{\partial \theta^2_{R\,2}} = -\frac{1}{\sqrt{2}} D_{R\,2}(S_0 \sigma_{R\,1} + S_2 \delta_{R\,1} PA_1) - \frac{1}{\sqrt{2}} D_{R\,2} P^*_{R\,2}(S_1 \sigma_{R\,1} PA^*_2 + S_3 \delta_{R\,1} PA_1 PA^*_2)$$

$$\frac{\partial RL}{\partial \theta_{R\,1}} = \frac{1}{\sqrt{2}} D_{R\,1}(\psi_{RL} S_0 \sigma^*_{L\,2} PA_2 + \psi_{RL} S_1 \delta^*_{L\,2}) - \frac{1}{\sqrt{2}} S_{R\,1} P_{R\,1}(\psi_{RL} S_2 \sigma^*_{L\,2} PA_1 PA_2 + \psi_{RL} S_3 \delta^*_{L\,2} PA_1)$$

$$\frac{\partial^2 RL}{\partial \theta^2_{R\,1}} = -\frac{1}{\sqrt{2}} S_{R\,1}(\psi_{RL} S_0 \sigma^*_{L\,2} PA_2 + \psi_{RL} S_1 \delta^*_{L\,2}) - \frac{1}{\sqrt{2}} D_{R\,1} P_{R\,1}(\psi_{RL} S_2 \sigma^*_{L\,2} PA_1 PA_2 + \psi_{RL} S_3 \delta^*_{L\,2} PA_1)$$

$$\frac{\partial LR}{\partial \theta_{R\,2}} = \frac{1}{\sqrt{2}} D_{R\,2}(\psi_{LR} S_0 \sigma_{L\,1} PA^*_1 + \psi_{LR} S_2 \delta_{L\,1}) - \frac{1}{\sqrt{2}} S_{R\,2} P^*_{R\,2}(\psi_{LR} S_1 \sigma_{L\,1} PA^*_1 PA^*_2 + \psi_{LR} S_3 \delta_{L\,1} PA^*_2)$$

$$\frac{\partial^2 LR}{\partial \theta^2_{R\,2}} = -\frac{1}{\sqrt{2}} S_{R\,2}(\psi_{LR} S_0 \sigma_{L\,1} PA^*_1 + \psi_{LR} S_2 \delta_{L\,1}) - \frac{1}{\sqrt{2}} D_{R\,2} P^*_{R\,2}(\psi_{LR} S_1 \sigma_{L\,1} PA^*_1 PA^*_2 + \psi_{LR} S_3 \delta_{L\,1} PA^*_2)$$

Non zero terms in θ_L:

$$\frac{\partial LL}{\partial \theta_{L\,1}} = \frac{1}{\sqrt{2}} D_{L\,1} P_{L\,1}(S_0 \sigma^*_{L\,2} PA^*_1 PA_2 + S_1 \delta^*_{L\,2} PA^*_1) - \frac{1}{\sqrt{2}} S_{L\,1}(S_2 \sigma^*_{L\,2} PA_2 + S_3 \delta^*_{L\,2})$$

$$\frac{\partial^2 LL}{\partial \theta^2_{L\,1}} = -\frac{1}{\sqrt{2}} S_{L\,1} P_{L\,1}(S_0 \sigma^*_{L\,2} PA^*_1 PA_2 + S_1 \delta^*_{L\,2} PA^*_1) - \frac{1}{\sqrt{2}} D_{L\,1}(S_2 \sigma^*_{L\,2} PA_2 + S_3 \delta^*_{L\,2})$$

$$\frac{\partial LL}{\partial \theta_{L\,2}} = \frac{1}{\sqrt{2}} D_{L\,2} P^*_{L\,2}(S_0 \sigma_{L\,1} PA^*_1 PA_2 + S_2 \delta_{L\,1} PA_2) - \frac{1}{\sqrt{2}} S_{L\,2}(S_1 \sigma_{L\,1} PA^*_1 + S_3 \delta_{L\,1})$$

$$\frac{\partial^2 LL}{\partial \theta^2_{L\,2}} = -\frac{1}{\sqrt{2}} S_{L\,2} P^*_{L\,2}(S_0 \sigma_{L\,1} PA^*_1 PA_2 + S_2 \delta_{L\,1} PA_2) - \frac{1}{\sqrt{2}} D_{L\,2}(S_1 \sigma_{L\,1} PA^*_1 + S_3 \delta_{L\,1})$$

$$\frac{\partial RL}{\partial \theta_{L\,2}} = \frac{1}{\sqrt{2}} D_{L\,2} P^*_{L\,2}(\psi_{RL} S_0 \sigma_{R\,1} PA_2 + \psi_{RL} S_2 \delta_{R\,1} PA_1 PA_2) - \frac{1}{\sqrt{2}} S_{L\,2}(\psi_{RL} S_1 \sigma_{R\,1} + \psi_{RL} S_3 \delta_{R\,1} PA_1)$$

$$\frac{\partial^2 RL}{\partial \theta^2_{L\,2}} = -\frac{1}{\sqrt{2}} S_{L\,2} P^*_{L\,2}(\psi_{RL} S_0 \sigma_{R\,1} PA_2 + \psi_{RL} S_2 \delta_{R\,1} PA_1 PA_2) - \frac{1}{\sqrt{2}} D_{L\,2}(\psi_{RL} S_1 \sigma_{R\,1} + \psi_{RL} S_3 \delta_{R\,1} PA_1)$$

$$\frac{\partial LR}{\partial \theta_{L\,1}} = \frac{1}{\sqrt{2}} D_{L\,1} P_{L\,1}(\psi_{LR} S_0 \sigma^*_{R\,2} PA^*_1 + \psi_{LR} S_1 \delta^*_{R\,2} PA^*_1 PA^*_2) - \frac{1}{\sqrt{2}} S_{L\,1}(\psi_{LR} S_2 \sigma^*_{R\,2} + \psi_{LR} S_3 \delta^*_{R\,2} PA^*_2)$$

$$\frac{\partial^2 LR}{\partial \theta^2_{L\,1}} = -\frac{1}{\sqrt{2}} S_{L\,1} P_{L\,1}(\psi_{LR} S_0 \sigma^*_{R\,2} PA^*_1 + \psi_{LR} S_1 \delta^*_{R\,2} PA^*_1 PA^*_2) - \frac{1}{\sqrt{2}} D_{L\,1}(\psi_{LR} S_2 \sigma^*_{R\,2} + \psi_{LR} S_3 \delta^*_{R\,2} PA^*_2)$$

Non zero terms in PD:

$$\frac{\partial RL}{\partial PD} = -\,j\; V_{LR}, \qquad \frac{\partial^2 RL}{\partial PD^2} = -\,j\; \frac{\partial RL}{\partial PD}$$

$$\frac{\partial LR}{\partial PD} = +\,j\; V_{RL}, \qquad \frac{\partial^2 LR}{\partial PD^2} = +\,j\; \frac{\partial LR}{\partial PD}$$

Derivatives of χ^2 wrt parameters

The χ^2 of the fit is defined as

$$\chi^2 = \sum_{i=0}^{n} (model_i - obs_i)^2 / \sigma_i^2$$

where $model_i$ is the model value for observation i (real or imaginary part of visibility), obs_i is the observation for i and

σ_i^2 is the variance of the amplitude of observation i. Denote $model_i - obs_i$ as $resid_i$.

The derivative of χ^2 wrt each parameter p is:

$$\frac{\partial \chi^2}{\partial p} = \sum_{i=0}^{n} 2 \left(resid_i \frac{\partial model_i}{\partial p} \right) / \sigma_i^2$$

This expression is complex, to get the real value, instead use

$$\frac{\partial \chi^2}{\partial p} = \sum_{i=0}^{n} 2 \left(\mathrm{Real}(resid_i)\mathrm{Real}(\frac{\partial model_i}{\partial p}) + \mathrm{Imag}(resid_i)\mathrm{Imag}(\frac{\partial model_i}{\partial p}) \right) / \sigma_i^2$$

where Real and Imag denote the real and imaginary parts of the argument.

The second derivative of χ^2 wrt each parameter p is:

$$\frac{\partial^2 \chi^2}{\partial p^2} = \sum_{i=0}^{n} 2 \left(\frac{\partial model_i}{\partial p} \frac{\partial model_i}{\partial p} + resid_i \frac{\partial^2 model_i}{\partial p^2} \right) / \sigma_i^2$$

This expression is also complex, to get the real value, instead use

$$\frac{\partial^2 \chi^2}{\partial p^2} = \sum_{i=0}^{n} 2 \left(\mathrm{Real}(\frac{\partial model_i}{\partial p})\mathrm{Real}(\frac{\partial model_i}{\partial p}) + \mathrm{Imag}(\frac{\partial model_i}{\partial p}\mathrm{Imag}(\frac{\partial model_i}{\partial p}) + \mathrm{Real}(resid_i)\mathrm{Real}(\frac{\partial^2 model_i}{\partial p^2}) + \mathrm{Imag}(resid_i)\mathrm{Imag}(\frac{\partial^2 model_i}{\partial p^2}) \right) / \sigma_i^2$$

REFERENCES

[1] W. D. Cotton, "Obit: A Development Environment for Astronomical Algorithms," *PASP*, vol. 120, pp. 439–448, 2008.

[2] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy, 2nd Edition*, Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr., Ed. Wiley-Interscience, 2001.

[3] W. D. Cotton, "Polarization in Interferometry," in *Synthesis Imaging in Radio Astronomy II*, ser. Astronomical Society of the Pacific Conference Series, G. B. Taylor, C. L. Carilli, and R. A. Perley, Eds., vol. 180, 1999, pp. 111–+.

[4] R. J. Sault, J. P. Hamaker, and J. D. Bregman, "Understanding radio polarimetry. II. Instrumental calibration of an interferometer array." *A&A Suppl.*, vol. 117, pp. 149–+, 1996.

[5] W. D. Cotton, "A New Method for Cross Polarized Delay Calibration of Radio Interferometers," *JAI*, vol. 1 in press, 2012.

[6] W. D. Cotton, "EVLA Continuum Scripts: Outline of Data Reduction and Heuristics," *Obit Development Memo Series*, vol. 29, pp. 1–25, 2012.