Computation Of Antenna Dependent Complex Gains

Sanjay Bhatnagar
N.C.R.A., Pune

January, 1999

1 Introduction

This document describes a method for computing the antenna dependent amplitude and phase from the visibility data for a point source. The method can be extended to include data from extended sources for which a good source model is known.

Interpretation of the antenna dependent quantities in terms of the antenna and system temperatures is also explained. Using the antenna dependent quantities from a database with good signal-to-noise ratio, the effective system temperature valid for interferometry data can be monitored.

The equations given here are exact and are implemented in the programs `antsol`[1] and `rantsol`[2], which operates on the GMRT visibility database written in the native LTA format.

2 Problem definition

The normalized cross-correlation function (the correlator output), measured by an interferometer using two antennas, antenna i and antenna j, in the limit $I \ll T_{sys_i}/\eta_i$, can be written as:

$$\rho_{ij}^{Obs} = \rho^{Obs}(u_{ij}, v_{ij}, w_{ij}) = \iint_{-\infty}^{+\infty} I(l,m)\sqrt{\frac{\eta_i \eta_j}{T_{sys_i} T_{sys_j}}}\; e^{2\pi\iota(u_{ij}l + v_{ij}m + w_{ij}\sqrt{1-l^2-m^2} + \phi_i - \phi_j)} \frac{dl\ dm}{\sqrt{(1-l^2-m^2)}} + \epsilon_{ij} \tag{1}$$

[1]http://langur.ncra.tifr.res.in/~sanjay/Offline/offline
[2]http://langur.ncra.tifr.res.in/~sanjay/Offline/offline

where $I(l,m)$ is the sky surface brightness, η_i is the sensitivity and T_{sys_i} the system temperature of the antenna i in units of Kelvin/Jy and Kelvin respectively, ϵ_{ij} is the additive noise on the baseline i-j, and ϕ_i is the antenna based phase of the signal. The rest of the symbols have the usual meaning.

In practice however, the antenna based amplitude ($\sqrt{\eta_i/T_{sys_i}}$) and phase (ϕ_i) are potentially time varying quantities. This could be due to changes in the ionosphere, temperature variations, ground pick up, antenna blockage, noise pick up by various electronic components, background temperature, etc. Treating the quantities under the square root in the above equation as the antenna dependent amplitude gain, these antenna dependent quantities can be written as complex gains $g_i = a_i e^{-\iota\phi_i}$ where $a_i = \sqrt{\eta_i/T_{sys_i}}$. For an unresolved source at the phase tracking center, variations in this amplitude will be indistinguishable from a variations in the ratio of η and T_{sys}.

In terms of g_is, we can write Eq. 1 as

$$\rho_{ij}^{Obs} = g_i g_j^{\star} \rho_{ij}^{\circ} + \epsilon_{ij} \tag{2}$$

where

$$\rho_{ij}^{\circ} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} I(l,m)\ e^{2\pi\iota(u_{ij}l+v_{ij}m+w_{ij}\sqrt{1-l^2-m^2})} \frac{dl\ dm}{\sqrt{(1-l^2-m^2)}} \tag{3}$$

The use of the word "antenna based gains" for g_is result into confusion for many and needs some clarifications. g_is are called complex "gains" since they multiply with the complex quantity ρ_{ij}. For an unresolved source, $|g_i|$ represents the fraction of correlated signal and $arg(g_i)$ represents the phase of the correlated part of the signal from the antenna with respect to the phase reference (usually the reference antenna). It is in this sense that it is referred to as "antenna based" gains. g_is are antenna based but a function of direction in the sky since, as defined here, they include T_{sys} which in turn includes the sky background temperature. However, here we assume that the angular scale over which g_is vary is larger than the antenna primary beam (isoplanatic case).

For an unresolved source at the phase tracking center, all terms in the exponent of ρ_{ij}° are exactly zero. ρ_{ij}° in this case would be proportional to the flux density of the source.

Given ρ_{ij}^{Obs} and knowing ρ_{ij}° the goal is to determine the antenna dependent complex gains g_is.

3 Solution for the complex gains

Assuming that the antenna dependent complex gains are independent, with a gaussian probability density function (this implies that the real and imaginary parts are independently gaussian random processes), one can estimate g_is by minimizing, with respect to g_is, the function S given by

$$S = \sum_{\substack{i,j \\ i \neq j}} \left| \rho_{ij}^{Obs} - g_i \; g_j^\star \; \rho_{ij}^\circ \right|^2 \; w_{ij} \tag{4}$$

where $w_{ij} = 1/\sigma_{ij}^2$, σ_{ij} being the variance on the measurement of ρ_{ij}^{Obs}

Dividing the above equation by ρ_{ij}° (the source model, which is presumed to be known - it is trivially known for an unresolved source), and writing $\rho_{ij}^{Obs}/\rho_{ij}^\circ = X_{ij}$, we get

$$S = \sum_{\substack{i,j \\ i \neq j}} \left| X_{ij} - g_i \; g_j^\star \right|^2 \; w_{ij} \tag{5}$$

If ρ_{ij}° represents the structure of the source accurately, X_{ij} will have no source dependent terms and is purely a product of the two antenna dependent complex gains.

Expanding Eq. 5, we get

$$S = \sum_{\substack{i,j \\ i \neq j}} \left[\; |X_{ij}|^2 - g_i^\star g_j X_{ij} - g_i g_j^\star X_{ij}^\star + g_i g_i^\star g_j g_j^\star \right] \; w_{ij} \tag{6}$$

Evaluation $\frac{\partial S}{\partial g_i^\star}$ and equating it to zero [3], we get

$$\frac{\partial S}{\partial g_i^\star} \;=\; \sum_{\substack{j \\ j \neq i}} \left[-g_j X_{ij} w_{ij} + g_i g_j g_j^\star w_{ij} \right] \;=\; 0 \tag{7}$$

or

$$g_i \;=\; \frac{\sum\limits_{\substack{j \\ j \neq i}} \; X_{ij} g_j w_{ij}}{\sum\limits_{\substack{j \\ j \neq i}} \; |g_j|^2 \, w_{ij}} \tag{8}$$

This can also be derived by equating the partial derivatives of S with respect to real and imaginary parts of g_i as shown in the appendix.

[3] Complex derivatives can be evaluated by treating g_i and $g_i^\star$ as independent variables. See reference 1 and the appendix

Since the antenna dependent complex gains also appear on the right-hand side of Eq. 8, it has to be solved iteratively starting with some initial guess for g_js or initializing them all to (1,0).

Eq. 8 can be written in the iterative form as:

$$g_i^n = g_i^{n-1} + \alpha \left[\frac{\sum\limits_{\substack{j \\ j \neq i}} X_{ij} g_j^{n-1} w_{ij}}{\sum\limits_{\substack{j \\ j \neq i}} \left| g_j^{n-1} \right|^2 w_{ij}} - g_i^{n-1} \right] \tag{9}$$

where n is the iteration number and $0 < \alpha < 1$. Convergence would be defined by the constraint

$$|S_n - S_{n-1}| < \delta \tag{10}$$

(the change in S from one iteration to another) where δ is the tolerance limit.

4 Interpretation of the equation

Eq. 8 offers itself for some intuitive understanding in the following way.

X_{ij} is a product of two complex numbers, namely g_i and $g_j^\star$, which we wish to determine. X_{ij} is itself derived from the measured quantity V_{ij}^{Obs}. Numerically speaking, each term in the summation of the numerator of Eq. 8 will involve g_i (via X_{ij}) and the multiplication of X_{ij} with $g_j w_{ij}$ would give g_i an effective weight of $|g_j|^2 w_{ij}$. Since the denominator is the sum of this effective weight, the right-hand side of Eq. 8 can be interpreted as the weighted average of g_i over all correlations with antenna i.

In the very first iteration, when $g_j = (1, 0)$, the normalization would be incorrect since the numeric value of g_j as it appears inside X_{ij} would be different from that used in the denominator of Eq. 8. However, as the estimates of g_js improve with iterations, the equation would progressively approach a true weighted average equation. The speed of convergence will depend upon the value of α and the convergence would be defined by the constraint in Eq. 10. In the ideal case when the true value of all g_is is known, right hand side of Eq. 8 also reduces to g_i.

Estimating g_i for an antenna, by averaging over the measurements from all baselines in which it participates (for an unresolved source) makes sense since for an N element array, g_i would be present in N-1 measurements (all the $X_{ij}|_{j=1,N;j\neq i}$) and the best estimate of g_i would be the weighted average of all these measurements. Proper weight for g_i, buried in each of the products X_{ij}, can be found heuristically as follows. g_i, estimated from the measurements of a given baseline, must obviously be weighted by the signal-to-noise ratio on that baseline. This is w_{ij} in the above equations. It must also be weighted by the amplitude gain of the other antenna making the baseline, namely g_j, to account for variation in antenna sensitivities and T_{sys}. The total weight for g_i would then be $|g_j|^2 w_{ij}$, the sum of which appears in the denominator of Eq. 8. Knowing that ideally $X_{ij} = g_i g_j^\star$, each of the $X_{ij}|_{j=1,N}$ must be multiplied by $g_j w_{ij}$ (to apply the the above mentioned weights to g_i), before being summed for all values of j and normalized by the sum of weights to form the weighted average of g_i. One thus arrives at Eq. 8 using these heuristic arguments.

5 Estimating T_{sys}

For an unresolved source of known brightness I, in the limit $T_a \ll T_{sys}$, $\rho^\circ_{ij} = I$ and Eq. 1 can be written as

$$\rho_{ij}^{Obs} = I g_i g_j^\star \approx I\sqrt{\frac{\eta_i\eta_j}{T_{sys_i}T_{sys_j}}} \tag{11}$$

where $\eta_i = A_e/2k_b$, A_e is the effective area of the dish, k_b is the Boltzman's constant and

$$|g_i| = \sqrt{\frac{\eta_i}{T_{sys_i}}} \tag{12}$$

Hence, knowing η_i, T_{sys_i} can be estimated from the amplitude of the antenna dependent complex gains.

All contributions to ρ_{ij}^{Obs}, which cannot be factored into antenna dependent gains, will result in the reduction of $|g|$. η remaining constant, this will be indistinguishable from an increase in the effective system temperature. Since majority of later processing of interferometry data for mapping (primary calibration, bandpass calibration, SelfCal, etc.) is done by treating the visibility as a product of two antenna based numbers, this is the effective system temperature that will determine the noise in the final map (though, as a final step in the mapping process, baseline based calibration can possibly improve the noise in the map).

In the normal case of no significant baseline based terms (ϵ_{ij}) in X_{ij}, the system temperature as measured by the above method will be equivalent to any other determination of T_{sys_i}.

T_{sys} can also be determined by recording interferometric data for a strong point source with and without an independent noise source of known temperature at each antenna. In this case

$$T_{sys_i} = T_{n_i}\left(\frac{{g_i^{ON}}^2}{{g_i^{OFF}}^2 - {g_i^{ON}}^2}\right) \tag{13}$$

where g_i^{ON} and g_i^{OFF} are the antenna dependent gains with and without the noise source of temperature T_n. Note that η_i does not enter this equation. Also, T_n

should be such that $\sqrt{T_a/(T_n + T_{sys})} \geq 0.2$ to ensure that the correlated signal is measured with sufficient signal-to-noise ratio (in this case, ≥ 0.04).

References

1. Complex Functional Analysis; Palka, Bruce P., Springer-Verlag, 1990.

A Update direction in real and imaginary representation

g_is are complex functions. One can therefore write S in terms of g_i^I and g_i^R, the real and imaginary parts of g_i and minimize with respect to g_i^I and g_i^R separately. It is shown here that the complex arithmetic achieves exactly this and the results are same as that given by complex calculus. The superscripts I and R in the following are used to represent the real and imaginary parts of complex quantities.

Expanding Eq. 5, ignoring w_{ij}s and writing it in terms of real and imaginary parts we get

$$\begin{aligned}
\sum_{\substack{i,j\\i\neq j}} \left|X_{ij} - g_i g_j^\star\right|^2 &= \sum_{\substack{i,j\\i\neq j}} \left[X_{ij} - g_i g_j^\star\right] \left[X_{ij}^\star - g_i^\star g_j\right] \\
&= \sum_{\substack{i,j\\i\neq j}} \left[\left(X_{ij}^R + \iota X_{ij}^I\right) - \left(g_i^R + \iota g_i^I\right)\left(g_j^R - \iota g_j^I\right)\right] \\
&\qquad \left[\left(X_{ij}^R - \iota X_{ij}^I\right) - \left(g_i^R - \iota g_i^I\right)\left(g_j^R + \iota g_j^I\right)\right] \\
&= \sum_{\substack{i,j\\i\neq j}} \left[\left(X_{ij}^R - g_i^R g_j^R - g_i^I g_j^I\right) + \iota\left(X_{ij}^I + g_i^R g_j^I - g_i^I g_j^R\right)\right] \\
&\qquad \left[\left(X_{ij}^R - g_i^R g_j^R - g_i^I g_j^I\right) - \iota\left(X_{ij}^I + g_i^R g_j^I - g_i^I g_j^R\right)\right] \\
&= \sum_{\substack{i,j\\i\neq j}} S_0 S_0^\star
\end{aligned} \tag{14}$$

where

$$S_0 = \left[X_{ij}^R - g_i^R g_j^R - g_i^I g_j^I\right] + \iota\left[X_{ij}^I + g_i^R g_j^I - g_i^I g_j^R\right] \tag{15}$$

Taking partial derivative of S with respect to g_i^R and reintroducing w_{ij}, we get

$$\begin{aligned}
\frac{\partial S}{\partial g_i^R} &= \sum_{\substack{j \\ j\neq i}} \left\{ \left[-g_j^R + \iota g_j^I\right] S_0^\star - S_0 \left[g_j^R + \iota g_j^I\right] \right\} w_{ij} \\
&= -\sum_{\substack{j \\ j\neq i}} \left[S_0 g_j + g_j^\star S_0^\star\right] w_{ij} \\
&= -2\sum_{\substack{j \\ j\neq i}} Re\left(S_0 g_j w_{ij}\right) \\
&= -2\sum_{\substack{j \\ j\neq i}} \left[\left(X_{ij}^R - g_i^R g_j^R - g_i^I g_j^I\right) g_j^R + \left(X_{ij}^I + g_i^R g_j^I - g_i^I g_j^R\right) g_j^I\right] w_{ij} \\
&= -2\sum_{\substack{j \\ j\neq i}} \left[X_{ij}^R g_j^R - X_{ij}^I g_j^I - g_i^R {g_j^I}^2 - g_i^R {g_j^R}^2\right] w_{ij}
\end{aligned} \tag{16}$$

Therefore,

$$\frac{\partial S}{\partial g_i^R} = -2\sum_{\substack{j \\ j\neq i}} \left[Re(X_{ij} g_j) - \left|g_j\right|^2 g_i^R\right] w_{ij} \tag{17}$$

Equating $\frac{\partial S}{\partial g_i^R}$ to zero, we get

$$g_i^R = \frac{\sum\limits_{\substack{j \\ j\neq i}} Re(X_{ij} g_j w_{ij})}{\sum\limits_{\substack{j \\ j\neq i}} \left|g_j\right|^2 w_{ij}} \tag{18}$$

Similarly

$$\frac{\partial S}{\partial g_i^I} = -2\sum_{\substack{j \\ j\neq i}} \left[Im(X_{ij} g_j) - \left|g_j\right|^2 g_i^I\right] w_{ij} \tag{19}$$

Therefore the equivalent imaginary part of Eq. 18 is

$$g_i^I = \frac{\sum\limits_{\substack{j \\ j\neq i}} Im(X_{ij} g_j w_{ij})}{\sum\limits_{\substack{j \\ j\neq i}} \left|g_j\right|^2 w_{ij}} \tag{20}$$

writing $g_i = g_i^R + \iota g_i^I$ and substituting for g_i^R and g_i^I from Eq. 18 and 20 respectively, we get

$$g_i = \frac{\sum\limits_{\substack{j \\ j\neq i}} X_{ij} g_j w_{ij}}{\sum\limits_{\substack{j \\ j\neq i}} \left|g_j\right|^2 w_{ij}} \tag{21}$$

This is same as Eq. 8, which was arrived at by evaluating a complex derivative of Eq. 5 as $\partial S/\partial g_i^\star$, treating g_i and $g_I^\star$ as independent variables. Evaluating $\frac{\partial S}{\partial g_i} = 0$ would give the complex conjugate of Eq. 21. Hence, $\partial S/\partial g_i$ gives no independent information not present in $\partial S/\partial g_i^\star$.