IVS Memorandum 2007-008v01

22 August 2007

“VLBI2010 Source Map Alignment Simulation”

Bill Petrachenko, Mario Bérubé

VLBI2010 Source Map Alignment Simulation

Bill Petrachenko, Mario Berube

August 22, 2007

Context: This study supports ongoing work to evaluate the feasibility of the broadband delay concept. Briefly, broadband delay involves using several RF bands spaced over a broad frequency range (e.g. 2-18 GHz) to determine the ambiguity-free RF phase. In previous studies, it was shown that, especially for sources with structure index 2 or 3, source structure could be a significant impediment to making the broadband delay concept work. Source structure corrections were proposed to reduce this problem, especially given the favorable UV coverage achievable with the larger networks and faster slew rates anticipated for VLBI2010. Correcting for structure includes three main aspects:

- Use amplitude and closure phase data to correct for structure in each band independently. If done perfectly, this would reduce structure to that of an apparent point source. However, due to the loss of absolute position information due to the nature of closure data, the position of the point source is, within limits, unknown. This is true in each of the RF bands, causing positional skews between bands.
- Remove the resulting inter-band positional skews.
- Align the images relative to some positionally invariant feature of the source, e.g. the centre of mass (which is unfortunately not usually visible).

Purpose: The purpose of this study is to see how well the inter-band positional skews can be removed. It is assumed that source structure has already been removed perfectly from each band. What remains is a series of apparent point sources (one in each band) whose positions are skewed relative to each other. The goal is to evaluate how well the relative positional skews can be determined.

Method: The study will be based on a Monte Carlo simulation, and will be carried out in the following steps:

1. Generate a test schedule
2. Select a test source
3. Generate random source position offsets in each band (excluding the highest frequency, where the position will be assumed to be known).
4. Generate random delay and ionosphere observables for each observation.
5. Generate random noise for each observation.
6. Combine the random elements from step 3-5 to generate a fake data set.
7. Invert the data to produce estimates (and their sigma's) of the delay and ionosphere observables for each observation and the positional offsets of the source maps in all of the bands except the one with the highest frequency.
8. Repeat steps 3 to 7 several times to determine bias and scatter of all estimated parameters and then compare these to the computed sigma's.

The study will be executed in two phases. In the first, it will be assumed that the RF phase ambiguities are known a priori. In the second, the ambiguities will be assumed to be unknown. In the latter case, step 7 will also include procedures to resolve the RF phase ambiguities. In this way, difficulties can be evaluated with respect to resolving phase ambiguities in the presence of inter-band position offsets, and strategies can be developed for doing this at the lowest possible SNR.

Theory: The schedule will produce observations for a number of different baselines, sources and times. Each observation, itself, will include delay and a phase observables, each at a number of different frequencies. The delay and phase observables can be written respectively as follows,

$$\hat{\tau}_{ijkl} = \tau_{ijk} + \frac{K_{ijk}}{f_l^2} + \frac{\partial \tau_{ijk}}{\partial \alpha} \cdot \Delta\alpha_{il} + \frac{\partial \tau_{ijk}}{\partial \delta} \cdot \Delta\delta_{il} + \tau_{jl}^I + \tau_{ijkl}^{\varepsilon} \tag{1}$$

$$\hat{\phi}_{ijkl} = f_l \cdot \left(\tau_{ijk} - \frac{K_{ijk}}{f_l^2} + \frac{\partial \tau_{ijk}}{\partial \alpha} \cdot \Delta\alpha_{il} + \frac{\partial \tau_{ijk}}{\partial \delta} \cdot \Delta\delta_{il} + \tau_{jl}^I \right) + \phi_{jl}^I + \phi_{ijlk}^{\varepsilon} + n_{ijkl} \tag{2}$$

Where

i ~ The source index
j ~ The baseline index
k ~ The time index
l ~ The frequency band index
f_l ~ The frequency of the l^{th} band
τ_{ijk} ~ The non-dispersive part of the delay
K_{ijk} ~ An ionosphere term proportional to line-of-sight electron content
$\Delta\alpha_{il}$ ~ The right ascension offset
$\Delta\delta_{il}$ ~ The declination offset
τ_{jl}^I ~ The uncalibrated instrument delay
τ_{jl}^{ε} ~ The random noise error of the delay
ϕ_{jl}^I ~ The uncalibrated instrument phase
ϕ_{jl}^{ε} ~ The random noise error of the phase
n_{ijkl} ~ The ambiguous number of cycles of the phase

For this study, it will be assumed that the instrumental terms, τ_{jl}^I and ϕ_{jl}^I can be determined in a preceding self-cal step and removed. Since it is possible to process the data on a source by source basis, the source index, i, will be removed to make the equations somewhat more readable,

$$\hat{\tau}_{jkl} = \tau_{jk} + \frac{K_{jk}}{f_l^2} + \frac{\partial \tau_{jk}}{\partial \alpha} \cdot \Delta\alpha_l + \frac{\partial \tau_{jk}}{\partial \delta} \cdot \Delta\delta_l + \tau_{jkl}^{\varepsilon} \tag{3}$$

$$\hat{\phi}_{jkl} = f_l \cdot \left(\tau_{jk} - \frac{K_{jk}}{f_l^2} + \frac{\partial \tau_{jk}}{\partial \alpha} \cdot \Delta\alpha_l + \frac{\partial \tau_{jk}}{\partial \delta} \cdot \Delta\delta_l + \tau_{jl}^I \right) + \phi_{jlk}^{\varepsilon} + n_{jkl} \tag{4}$$

Estimating Map Offsets Assuming Phase Ambiguities are known a Priori

Let us begin by considering the easier case where the phase ambiguities, n_{jkl} in equation (4) are perfectly known so they don't need to be estimated. This will let us know the precision with which the map offsets can be determine, but will not yet inform us whether or not the parameters can be determined in practice. [Note: Determining the phase ambiguities in the presence of map offsets will be the subject of the next section.]

As it turns out, if either the phase or delay observable is used alone, the inversion is degenerate. However, since the ionosphere affects phase and delay with the opposite sign, the degeneracy can be removed by using both observables together. Unfortunately, the accuracy of the map offset estimates is then limited by the lower precision of the group delay observable. This can be offset somewhat through averaging if, over the course of a session, a large number of observations are taken on each source. Fortunately, this is well aligned with the requirement for good uv coverage imposed by the mapping phase of source structure corrections and will be feasible with the new VLBI2010 observing modes where larger networks and faster slewing antennas will be used. All analysis we be made assuming an overall SNR of 14 (SNR=7 in each of the 4 bands) which is roughly the minimum SNR where the broadband delay approach can be applied reliably. As a result, the precisions reported for the map offsets are the minimums for a schedule and network similar to the one used.

The system assumed for this analysis has the following parameters:

Band centre frequencies (GHz)	2.5, 5.25, 7.0 and 11.25 GHz
Bandwidths (GHz)	1.0 GHz
Reference band for map offsets (GHz)	11.25 GHz
SNR total	14
SNR/band	7

The schedule assumed for the analysis is refered to as stat16_6_2p1D0 and has the following parameters:

Network size (no. of antennas)	16
Antenna azimuth slew rate (deg/s)	6.0
Antenna elevation slew rate (deg/s)	2.1
Source list	best 56 of the current operational list
Total no. of observations	74,613
Scans/station/hr	59

The analysis involves a standard linearized parameter inversion using all the delay and phase observations for a source and using equations (3) and (4) to solve for τ_{jk} and K_{jk} for each value of j and k, and $\Delta\alpha_l$ and $\Delta\delta_l$ for all values of l except the largest, i.e. the highest frequency band where the offset is assumed to be zero. The formal errors appear on the diagonal of the variance covariance matrix.

The formal errors for the delay measurements, assuming an overall SNR of 14, were all near 1.92 ps. The formal errors for the source map offsets are summarized in Figures 1 and 2 for α and δ respectively. As can be seen, the precisions are dependent on both the band being considered and the number of observations in the schedule for a particular source.

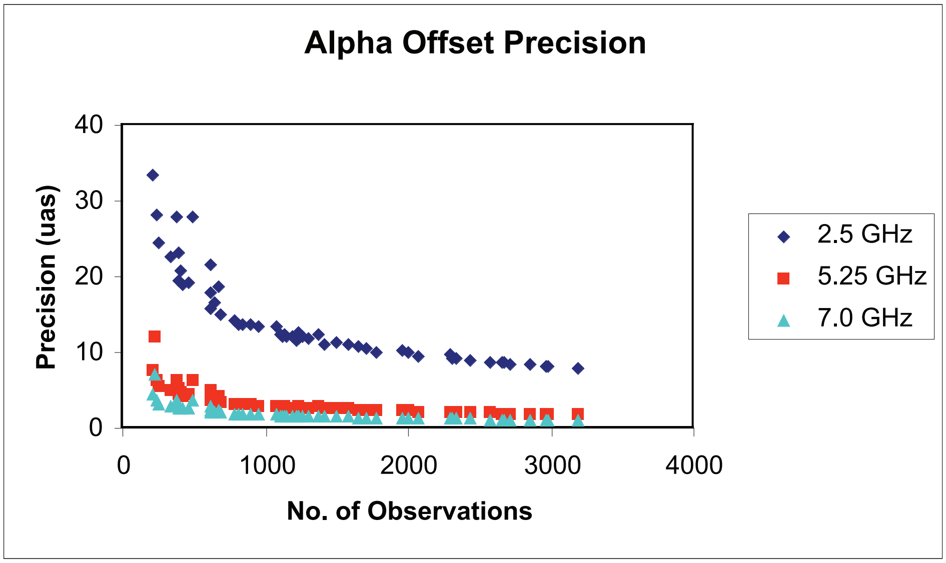


Figure 1. This plot represents the precision of the alpha map offsets vs. the number of observations for a particular source.

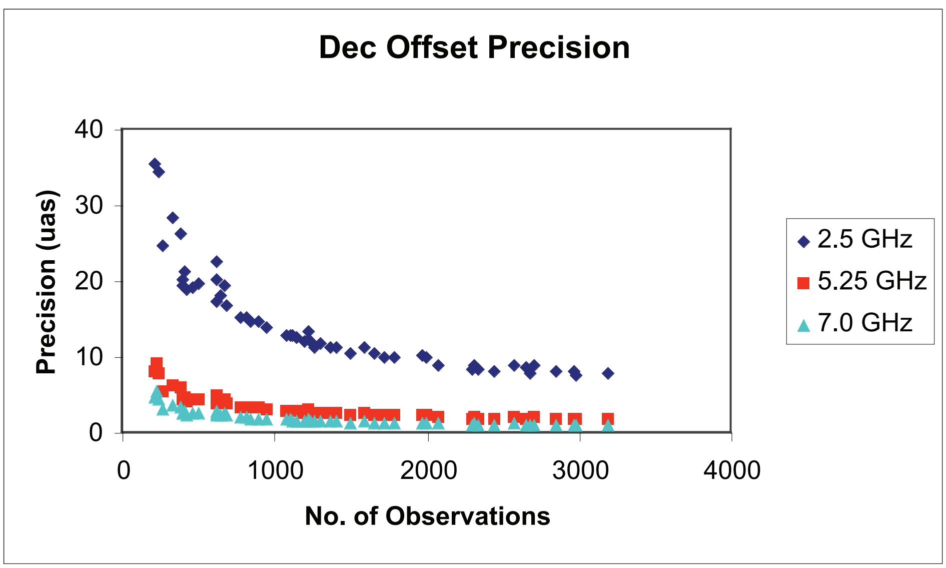


Figure 2. This plot represents the precision of the dec map offset determination vs. the number of observations for a particular source.

In addition, in order to verify the validity of the formal error estimates, a Monte Carlo simulation was performed to compare the scatter of the estimated parameters with the formal error. In the Monte Carlo simulation, random values of α_{il}, δ_{il}, τ_{ijk}, K_{ijk}, and the noise n_{ijkl} (appropriate for overall SNR=14) were generated to synthesize a fake data set. An inversion was performed to recover α_{il}, δ_{il}, τ_{ijk} and K_{ijk}. This was repeated 1000 times for each source and the scatter of the estimates was compared to the formal error. The results are summarized in figure 3. Although a statistical analysis was not performed to determine if the relationship is as expected, it is clear that there is good agreement between the formal error and the scatter of the Monte Carlo estimates.

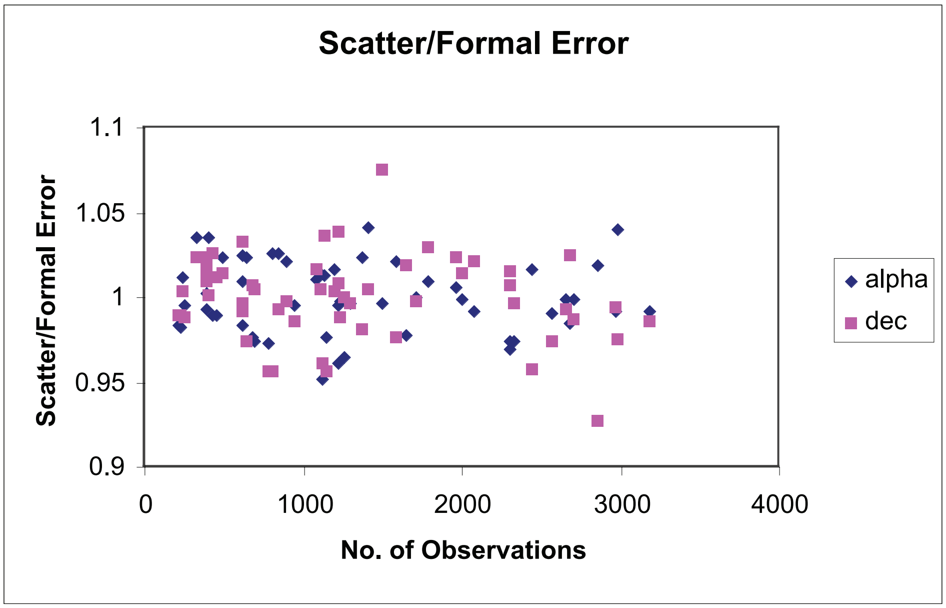


Figure 3. This is the ratio of the Monte Carlo scatter to the formal error plotted with respect to number of observations.

Estimating Map Offsets Assuming that Ambiguities are NOT known a Priori

In this section we will focus on the problem of resolving RF phase in the presence of map offsets. In VLBI group delay estimation, the normal procedure for resolving phase within an IF band is to do a search through all reasonable values of τ and $\dot{\tau}$. The values that produce the largest coherent average are the maximum likelihood estimates of the parameters. It has been shown [refer to presentations by myself and Alan Rogers at the Haystack f2f meeting] that the same procedure can be applied to resolving RF phase in the broadband delay technique if the search is extended to also include terms to account for phase curvature due to the ionosphere. Although the computational load increases, it is not a problem. In both group and broadband delay cases, phase resolution can be accomplished on an observation-by-observation basis.

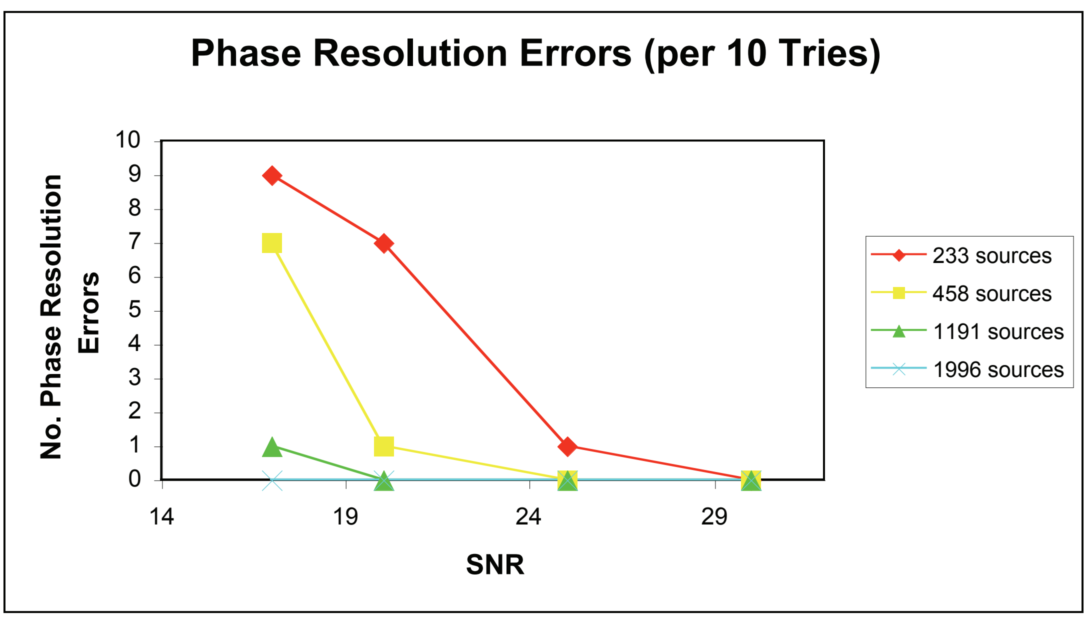
The problem is considerably more difficult in the presence of map offsets. Since each observation contains information about the map offsets and many observations are required to get the required precision, it becomes necessary, while solving for the map offsets, to simultaneously resolve phase for many observations. If a straightforward search algorithm is applied, the search space rapidly becomes astronomically large.

To avoid this, a multi-step process has been developed and tested. In principle, the process is applied as follows. An initial estimate of the map offsets is determined using the group delays only. This initial estimate is used to correct the observables for the map offsets and then the phases are resolved as if they had no map offsets. The resolve phases are then used in a final inversion to improve the estimates of the map offsets. There is a problem though, in that the group delay solution is degenerate if all map offsets are estimated simultaneously. The solution used here was to estimate, in the initial group delay solution, all but one set of the map offsets. To determine the final initial estimates an alpha/dec search is then performed based on minimizing the rms phase residuals after each attempt at phase resolution. The process is described in more detail below:

1. Prior to resolving the phases, the data in each band is processed on its own. In this way, a large number of raw correlation coefficients in each band are reduced to two outputs, those being the group delay and phase. This data compression greatly reduces the number of calculations in following steps. [Note that in order to do this reliably, SNR=7 is required in each band.]
2. Since the group delays are free of ambiguity problems, the first estimate of the map offsets can be achieved using group delay data only. Unfortunately, the group delay inversion is degerate if all relative map offsets are used. It is thus necessary to fix the map offsets in a second band and only estimate the offsets in the remaining two bands. Although the resulting offsets are in general incorrect, they form a starting point for the search in step 3.
3. Since the offsets in the band that was fixed in step 2 are in general incorrect, this step involves searching, in the fixed band only, through all reasonable sets of map offsets (both alpha and dec) to try to find the best set. The process involves resolving phase for all sets of sample offsets and choosing the set that minimizes the post-resolution rms phase residual.
4. Finally, used resolved phase offsets from step 3 to do a full data inversion.

This process was tested with the same simulated data set as for figures 1-3. Four sources were selected each having 233 (0003-066), 458 (0208-512), 1191 (0607-157) and 1996 (1300+580) scheduled observations respectively. The results for 10 tries of phase resolution are displayed in Fig. 4.

Fig. 4. Phase resolution errors (for 10 tries) vs. SNR for four sources, each having a



different number of scheduled observations.

As can be seen, performance depends strongly on the number of scheduled observations. Although these results indicate that the process is workable, it places higher demands on system sensitivity, especially for sources with fewer than about 1000 scheduled observations, e.g. in the case of the source with 233 observations, an SNR of 30 or more is required for reliable phase resolution.

To avoid this sensitivity related degradation, a further strategy was devised, and tested. It is an iterative process in which the least sensitive sources are resolved first. Step 1 above was performed exactly as before. However, steps 2, 3 and 4 were executed repetitively, beginning with only the fraction of sources that are least sensitive to errors in the map offsets and gradually adding, in each iteration, sources that are more and more sensitive until all sources are included. Each time through the loop the values of the map offsets become more precise since more data are used.

As hoped, when tested with simulated data, this addition to the algorithm improved the ability to resolve phase at lower SNR. Since source 0003-066, with only 233 observations in the simulated schedule, was the most problematic in Fig.4, it was used for the tests. The phase resolution test was repeated 366 times to simulate a full year of observations on that source. Assuming an SNR of 14, it was found that two phase resolve errors occurred out of a total number of attempts of 85278, which is an error rate of about 2 parts in 10**5. It is likely, with such a small error rate, that the phase resolution errors will appear as outliers and can be edited out. To see if the process fails gracefully, the test was repeated for an SNR of 11. The number of errors was 63 out of 85278, which is an error rate of 1 part in 10**3, which indicates that failure to meet an SNR target of 14 would not be catastrophic.

Conclusions

This study has shown that relative map offsets can be reliably and accurately determined directly from “broadband” VLBI data. At an overall SNR of 14 (which is minimum for reliable detection) and for an example 4-band system, the map offset precision ranged from about 35 to 1 uas depending on the frequency of the band and the number of observations taken of the source over the course of the 24-hour session. Since the process depends on using the resolved RF phase observable, it was also shown that phase could reliably be resolved down to an SNR of about 14 providing that at least about 200 well-spaced observations of the source were made over the course of the 24-hour session.

[Note that Thomas Hobiger has proposed and analyzed a different, simpler and more rigorous process for resolving the integer phase ambiguities based on the LAMBDA technique that is typically used with GNSS data. His analysis further supports the conclusion that phase ambiguities can reliably be resolved (even in the presence of map offsets) down to an overall SNR of at least 14. This analysis will be reported separately.]