JIVE
Joint Institute for VLBI ERIC

# Comparison of AIPS and CASA VLBI processing using MeqSilhouette simulations

*BlackHoleCam deliverable*


| | |
|---|---|
| Revision | 1.1 |
| Author | Ilse van Bemmel, Des Small, Mark Kettenis |
| Date | 15 May 2017 |

## Scope

The MeqSilhouette team has provided several simulations to use for testing our new CASA fringe fitting task. The following questions will be addressed:

1. How does CASA perform compared to AIPS?
2. Can CASA and AIPS recover the actual values of fixed input delay offsets?
3. Do CASA and AIPS recover the same delay solutions as the South African (RSA) team?
4. What is the impact of noise in the simulation on the solutions?
5. Is the fixed input delay code robust or is there an off-by-1 error?

## Simulations provided by RSA team

| Name | Freq (GHz) | #chan | Array | Delay | Target | Date of simulation |
|---|---|---|---|---|---|---|
| Sim1 | 44 | 64 | VLBA | Fixed | 3C279 | Dec 2015 |
| Sim5 | 230 | 32 | EHT | Fixed + noise | SgrA* | Mar 2017 |
| Sim6 | 230 | 32 | EHT | Fixed | SgrA* | Mar 2017 |

The thermal noise properties in Sim5 are determined using the SEFD values from Lu et al. (2014). The first simulation is a full hour. I have split off 5 minutes and handled that as a single scan for the fringe fitter. The last two simulations contain only 3 minutes, which is handled as a single scan. For Sim1 and Sim5 I split off the CORRECTED_DATA column in CASA, making a new MS, and use *exportuvfits* to write a UVFITS file for AIPS to read. Sim6 was delivered as a uvfits file and I used *importuvfits* to read this into CASA.

## Processing in AIPS

The task FITLD is used to read the UVFITS file into AIPS. The default settings are used. I run INDXR to produce an NX table and a CL table. To ensure that AIPS uses the same frequency reference pixel, I run AXDEFINE on a copy of the data to set the frequency reference to channel 1, consistent with CASA. This impacts the phase solutions when the phase gradient over the band is large. For comparison with the RSA AIPS processing I use the original data. For comparison between my AIPS and CASA processing I use the re-referenced data.

All datasets contain only a single scan. Fringe fitting is done on the dataset as a whole with reference antenna 1 for Sim1, and 2 for Sim5 and Sim6. The solution interval is set to the scan length. The solutions are read from the SN table using a ParselTongue script. With the CLCAL task a CL2 table is produced to calibrate the data. Note that in Sim5 only the US-based antennas see the source, PV and PdbI (antennas 5 and 6 in AIPS based counting) are excluded from the analysis.

## Processing in CASA

The processing is done with the latest prototype fringefitter, version 62:b71bff5eae7c. The task *clearcal* is used to generate a CORRECTED_DATA column which is identical to the DATA column. The fringe fitter is run on the whole scan with the same reference antenna as AIPS. The solutions are written to a FringeJones calibration table and a text file for further analysis. The calibration table is applied with the *applycal* task, to produce a calibrated CORRECTED_DATA column.

## Comparing AIPS and CASA

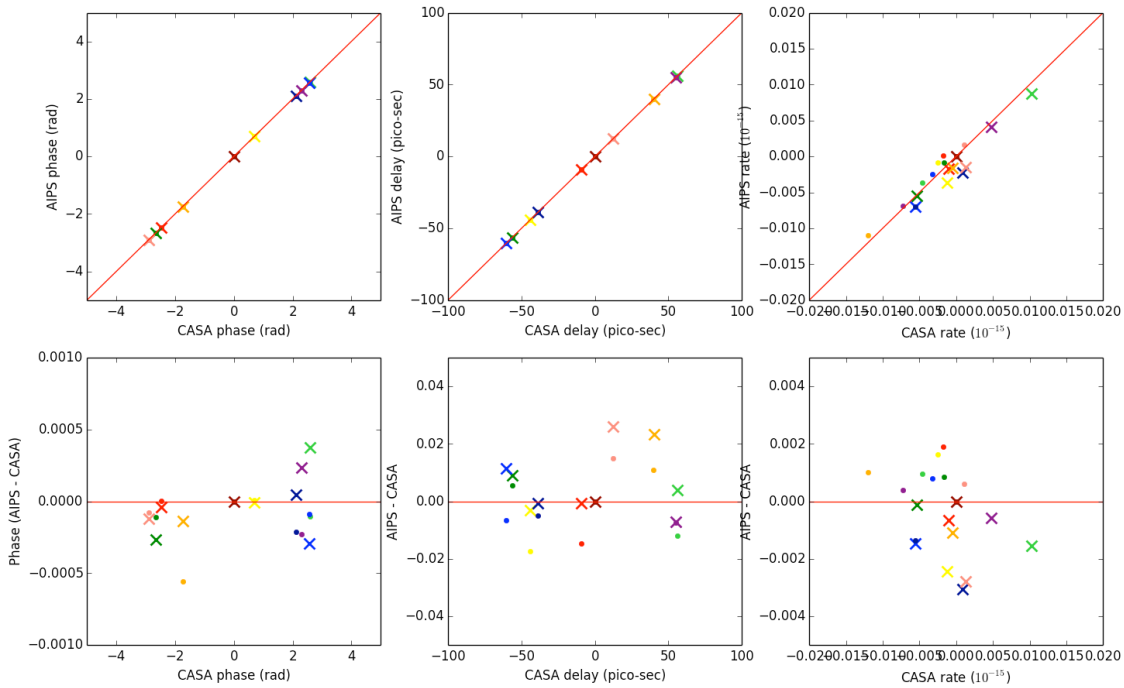


**Figure 1.** Sim1 comparison of phase, delay and rate solutions between CASA and AIPS fringe fitting. The colour coding is per station, the crosses and point represent the two orthogonal polarisations.

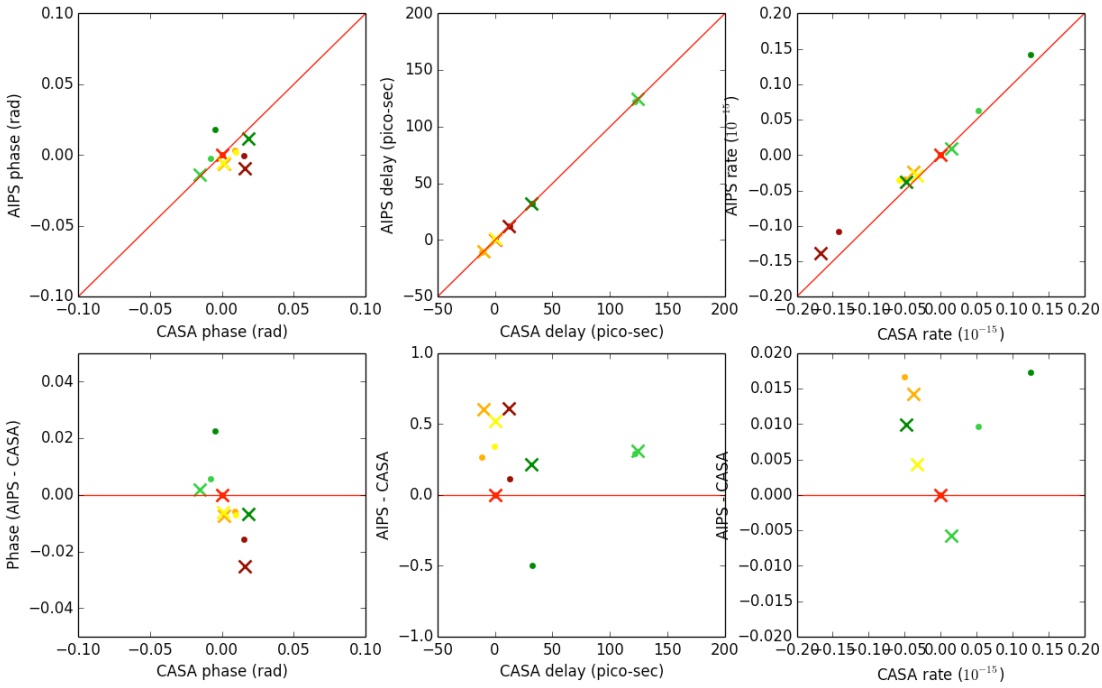


**Figure 2.** As figure 1 for Sim5, note the difference in scales.

For all simulations the solutions for phase and delay and rate are compared. For Sim6 I find that AIPS and CASA are identical down to 7$^{th}$ significant digit. For Sim1 and Sim5 I find there are minor differences between CASA and AIPS. For Sim1 the differences are significantly smaller than for Sim5.

Overall, the phase solutions from CASA are very small, which is consistent with the fact that CASA uses the first frequency channel as a reference. On all baselines the phase is nearly 0 degrees there.

For Sim5 the phase solutions in AIPS and CASA are consistent to within 0.02 radian (see Figure 2). The small variations seen on the solutions are at least one order of magnitude below the expected noise on the solutions of real data. The errors on the phases are relatively large due to the very small phase values.

## Recover input values & off-by-1 error

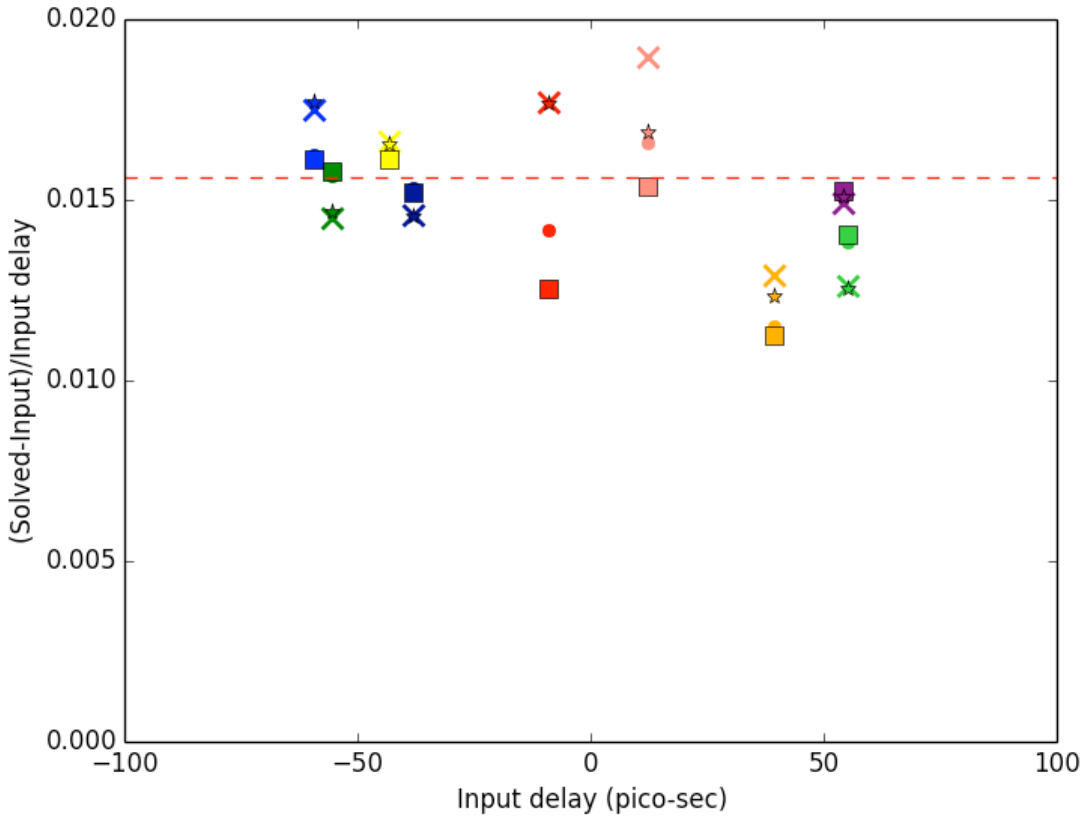


**Figure 3.** Relative error in the delay solutions for Sim1. The dashed line indicates the value expected if the code implementing the delay offset has an off-by-1 error in the channel numbering with 64 channels. The values are consistent with an off-by-1 error being present.

For all simulations I compare the delay solutions to the delay input value by calculating:

$$\frac{Solved\ delay\text{-}Input\ delay}{Input\ delay}$$

When the solutions are perfectly recovered, the values should be close to 0. For Sim1 there are offsets, with size and direction of the offset comparable for all antennas. This is attributed to an off-by-1 error in the code that implements the delay offset (see Figure 3). There are 64 channels in the data, and the offset is equal to 1/64, indicating that the problem is in the channel handling.

For Sim5 there are also large offsets, but here they are randomly distributed between the antennas. To verify if this is due to the thermal noise in the simulation, Sim6 is produced in a similar setup, but without the noise. Indeed, in Sim6 there is no longer any offset. The solutions are comparable to the

input delay down to 5 significant digits, which is several orders of magnitude above the precision ever required for real data. This implies that the off-by-1 bug which was present in Sim1 is fixed, and the noise on the solutions in Sim5 is indeed due to thermal noise in the simulation.

## Comparison of my AIPS and CASA results with RSA AIPS results

To assess the differences between the RSA AIPS results and a proto-type Bayesian fringe fitter, I compare the AIPS results I obtained with the results obtained by the RSA team.

For Sim5 the phase and delay solutions from a separate processing in AIPS by the South African team are available. The table provided with the dataset lists the solutions from AIPS processing done in RSA. There are two time intervals: 15h32m35.6s and 15h34m05.6s. The first interval is closest to the time stamp of my solutions, and has the same solution interval. I will use that for comparison. From my own processing I use the original data before frequency re-referencing

The phase and delay solutions are comparable to within 0.01 radian and 2 picoseconds (see Figure 4). Small differences like this can be caused by different settings in AIPS processing, and are not a major concern. The RSA team did not provide corrected phases, so these cannot be compared. The RSA phase and delay solutions are identical for both polarizations, which is due to setting APARM(3, averaging RR and LL. I did not set that for my AIPS processing.

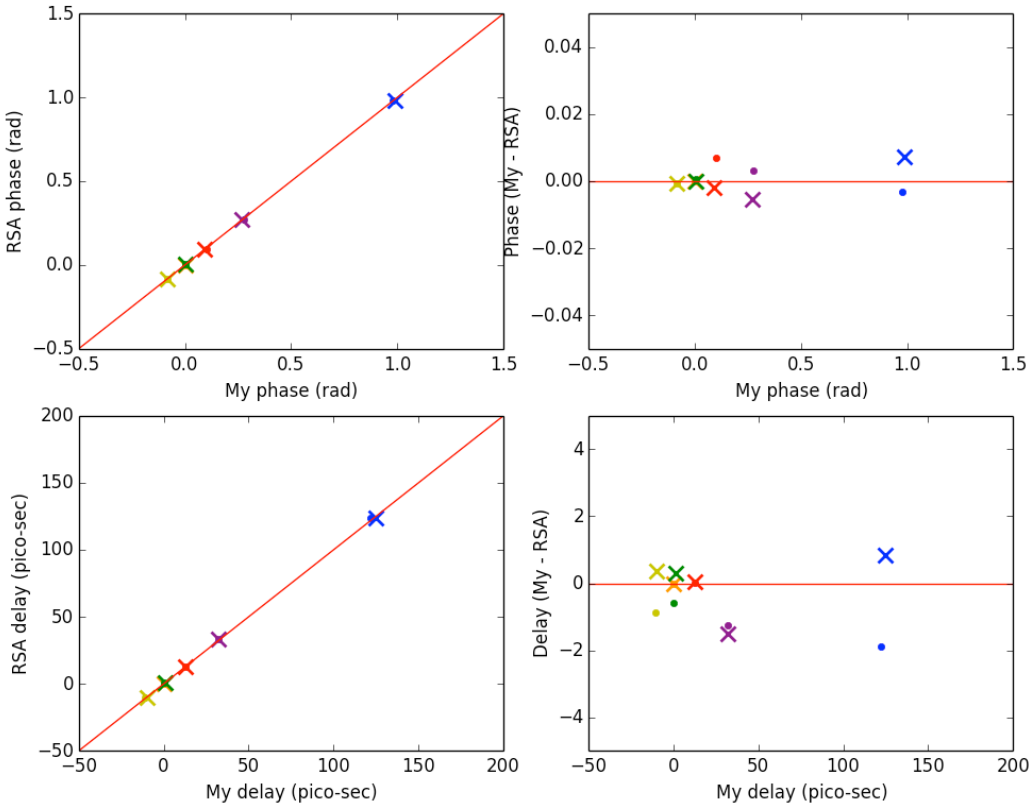


**Figure 4.** Comparison of my AIPS processing and the RSA AIPS processing. The colour coding in per station: red, orange, yellow, green, blue, purple from station 1 through 6. Dots are XX, crosses are YY polarization.

It is not possible to directly compare my CASA solutions with the RSA AIPS solutions, since CASA by default uses the first channel of the band as frequency reference. However, as discussed above, the re-referenced AIPS solutions are consistent with my CASA solutions, so by proxy the CASA solutions will also be consistent with the RSA AIPS solutions.

## Conclusions

The AIPS and CASA processing are fully consistent on both sides and recover the input values to high precision. The problems encountered in the simulations can be traced to an off-by-1 error in an earlier version of the code, and the thermal noise in Sim5 which causes significant scatter on the fringe fit solutions.

## Open issues

Since we are currently looking into this: why are the AIPS weights different from the CASA weights on the uncalibrated and untouched data? The higher the weight, the larger the difference.

## References

Lu R., Broderick A.E., et al., 2014, ApJ 788, 120