VLB ARRAY MEMO No. 388

National Radio Astronomy Observatory
Charlottesville, Virginia

16 Sept. 1984

To: VLBA Project

From: R. C. Walker

Subject: Bit Error Rates

During a recent meeting of the Recorder Group, it became clear that the specification of the allowed bit error rate was not well established. In this memo, I discuss some of the considerations that should determine the specification and suggest some allowed rates.

I hope that the dynamic range of VLBA maps will not be limited by easily controlled parts of the hardware such as the tape recording system. As an example of what might be possible, dynamic ranges (peak to off-source rms) of 30,000 have been achieved recently at the VLA and another factor of 3 or more should be possible before the thermal limits are reached. I suggest that an ultimate criterion for specifications that can affect the dynamic range be that the errors introduced in the map plane for full track observations (about 12 hours duration) be smaller than 1e-5 of the peak in the map (1). I am not as prepared to give a general specification for the geodetic/astrometric observations. However, where one is needed in the discussion of sync errors below, I suggest that errors should not cause a shift in measured path length of more than 1 mm.

Errors that are introduced by the imperfect transmission of the clipper output to the correlator, generically refered to here as bit errors, can have a variety of effects depending on their nature. Any error that is not corrected by the processing will show up in the map plane at a level about equal to the magnitude of the error multiplied by the fraction of the data affected and reduced by the square root of number of independent occurances of the error. For example, if all baselines are affected by non-closing errors of similar magnitude and the errors are constant in time on each baseline, the errors in the map plane will be about equal to the magnitude of the errors divided by the square root of the number of baselines (about the number of stations). This result is about what is seen on the VLA.

Errors will generally produce a combination of closing and non-closing offsets in the data. Closing offsets are ones that can be removed by correcting all data to a given station using a single gain adjustment - they are antenna, not baseline, dependent. Modern mapping methods are capable of removing closing errors because such errors mimic the instrumental offsets due to weather, pointing (point source case), system temperature fluctuations etc. that the self-cal

----

(1) Throughout this document, NeM means N times 10 to the M power (FORTRAN notation) for ease of typing.

proceedures correct. Non-closing errors are far more serious because they are not corrected by self-cal and they cannot be calibrated unless they are constant over long enough periods for calibration using calibration sources (at least several hours). Non-closing errors introduced by errors in the recording system are unlikely to be constant and so must be kept very small. Note that the ability of self-cal to correct closing errors will be limited by the noise and by the non-closing errors.

To emphasize the importance of non-closing errors, I would like to point out that, when not corrected, they currently limit the dynamic range (peak to rms) of VLA observations to about 13,000 and of VLBI observations (Mark II) to about 1000. The 30,000 peak to rms obtained for some VLA observations (Walker, VLA Scientific Memorandum 152) required correction of the constant component of non-closing errors. The far more serious limits encountered in current VLBI experiments are thought to be mostly the result of mismatched bandpasses (probably a filter, not recording system, problem) and we are exploring the possibility of making corrections. However, I hope that the VLBA will be designed so that the non-closing errors are much smaller than in either current VLBI or the VLA.

The specification of allowed error rates on the recording system depends on the types of errors involved. Below is a breakdown of several types and suggested specifications for those types. I encourage discussion of these specifications and, especially, I encourage others to try to identify aspects of the entire VLBA system that might limit the dynamic range of maps or the accuracy of geometric results.

I. Detected and flagged errors.

Errors that are detected and only cause a loss of integration time result in a decrease in signal-to-noise. Since the SNR is proportional to the square root of the integration time, relatively high error rates of this type can be tolerated. A 10% data rejection rate would not cause serious problems although, in the interests of generating a robust system, a specification of a maximum 1% data rejection rate is probably more reasonable. The actual number of bad bits allowed in order to meet this rejection rate spec will depend on the mimimum number of bits rejected when an error is detected.

II. Undetected errors.

Undetected errors are correlated as if they are good data and cause degradation of the results. They can be separated into random and systematic errors, the later of which can cause more serious problems.

A. Random errors.

Errors that do not result in a systematic shift of one type of bit to another, or a systematic shift in any of the processing parameters such as delay, are refered to here as random errors. Their sole effect is to reduce the correlated amplitude of the affected data. To the extent that the errors close, they effect the a-priori calibration but are eventually removed by self-cal. It always helps to have the a-priori calibration as good as possible. Therefore errors introduced by system problems should not degrade that calibration. The a-priori calibration should be good to somewhat better than 1% so a specification that the rate of random errors be smaller than 1e-3 is reasonable from this point of view.

There are ways that random errors can introduce closure errors:

1. Bandpass and integration effects.

If the effective bandpass is altered in different ways on more than one antenna, or if the signal is degraded on two different antennas at different times within the same integration period, closure errors with a magnitude of about the square of the error rate are introduced. If the error rate is less than 1e-3, the closure errors will be less than a 1e-6 which will not limit the dynamic range of maps.

2. Multiple pass processing

Multiple pass processing should be rare on the VLBA (or the correlator will have trouble keeping up) but will occur occasionally when very large experiments are done. These very large experiments are just those that are attempting to get the highest possible dynamic range so their needs should not be dismissed. When multiple pass processing is done, closure errors on the order of the difference in error rates on different passes are introduced. Hence, for a 25 station experiment, errors that are relatively constant on each station during a pass, but different between passes (eg some sort of alignment problem - this is fairly typical in current Mark II operations) by 1e-3 would appear at the level of 4e-5 in the map (divide by the square root of the number of baselines as discussed above). This is somewhat high and suggests that the specification on the allowed rate of random errors be reduced. A rate of 1e-4 would be comfortable although a few times that would be ok.

B. Biased errors

1. Bit decoding bias

Biased errors occur when the errors tend to be of one type, eg. 1's that are read as 0's. They can also occur if non-data bits (eg sync patterns) are correlated as data. Such errors have effects similar to the effects of clipper biases and can be a problem if not properly attenuated. If the biased errors occur at uncorrelated times on different stations, they cause spurious correlation coefficients at zero fringe rate of about the product of the magnitudes of the errors on the two stations involved in a baseline (roughly the square of the error rate). If the times of the errors are correlated, (eg. they occur at the same time after geometric delays are removed) they can cause much larger spurious correlations, up to the error rate itself.

For current VLBI observations, biased errors are highly attenuated because of the high fringe rates applied by the processor. The VLBA processor will be required to process low frequency data and baselines as short as a few kilometers (internal to the VLA) so the fringe rates can be very slow and cannot be relied upon to attenuate offsets. The data offsets introduced will be fringe rate dependent and, hence, non-closing. In fact, they are likely to mimic large scale, linear structures in the map plane that have integrated amplitudes of about spurious correlation coefficient times the system temperature in Jy. They also will hinder the self-cal process. The offsets can be attenuated either by phase switching or by offsetting local oscillators. One or the other scheme must be used to remove clipper biases (note that the VLA works poorly without phase switching). Offset LO's will be difficult because of the 10 kHz minimum step and the need to keep 64 kHz bandpasses matched. Therefore phase switching will be needed at the telescope and could be carried through the recording system.

The system temperatures on the VLBA will be about 500 Jy and map noise levels in the 10's of micro Jy should be possible, especially with experiments using many extra stations (such noise levels, or better, are reached with the VLA in some observations). To be sure to avoid problems, spurious correlation coefficients of greater than 1e-7 should be avoided. The specification of a bit error rate of less than 1e-4 suggested above will insure that such offsets are avoided in the absence of biased errors that are correlated in time at different stations.

2. Missed or added bits (undetected sync errors).

Sync errors have the effect of altering the effective time of the data for time between the occurance of the error and the next time the data is resynced. The bit error rate due to sync errors must count all bits correlated with the wrong time, not just the missed or added bits. Sync errors degrade mapping data by somewhat less than the amplitude times the bit error rate. Delays are offset by about the bit error rate times the delay per bit. Sync errors are likely to produce systematic offsets (they certainly do in the amplitudes) so their effects are not attenuated by multiple, independent occurances. However, for mapping observations, they will at least partially close in the absence of multiple pass processing. The most serious effects of sync errors will be on geodetic/astrometric observations where closure does not help. A specification good for geodesy will easily suffice for astronomy.

A slip of one bit at the 2 MHz bandwidth (probably a common choice for geodesy) causes a delay error of 0.25 microsecond or 7500 cm of path length. We hope to measure baselines to 1 cm or less with the VLBA so, in the interests of not being limited by instrumental problems, a limit to bit error induced offsets of 1 mm of path would be reasonable. This suggests a maximum error rate 1e-5.

C. Other Baseline dependent errors.

Any errors that are baseline dependent are likely to be sources of non-closing errors and therefore must be kept small. The requirements are very much like those discussed above in the section on multiple pass processing. I'm not sure what sources of baseline dependent errors might arise. They could occur if the decoding is baseline dependent (as in the Haystack Mark III processor). They could also occur when the fringe rate is very slow on a baseline and the three level fringe rotator is not a good sin wave approximation over an integration time (at slow fringe rates, the fringe rotation should probably be done in the fringe processor).

D. Loss of a recorder track.

If each recorder track corresponds to a specific part of the total recorded bandwidth (as in Mark III), loss of a track alters the effective total bandpass and hence introduces closure errors. Therefore, such correspondence should be avoided. If the channels are spread across the tracks in an appropriate way, loss of a track will simply cause a loss of integration time which is not a serious problem.

In summary, loss of integration time and purely antenna dependent errors are much less serious than biased and non-closing errors. This suggests that an effective error detection scheme should be used, but that error correction is not especially important. I hope that others with more information on error detection will suggest methods, but I understand that, with less overhead than currently used by the parity bits in Mark III, much better error detection can be obtained. This is especially true if pairs of errors, which are not detected by parity, are especially likely as suggested by Benno Rayhrer (private communication).

The following summarizes my suggested maximum error rates:

1. Discarded data < 1e-2

2. Undetected, bit errors < 1e-4

3. Undetected sync errors (count all bits correlated at the wrong time) < 1e-5

4. Spurious correlation coefficient due to biased errors < 1e-7

In addition:

1. The data from all channels should be spread across all recorder tracks.

2. Phase switching should be used at the clippers. Extra insurance against biased errors would be obtained if the phase switching is not undone until after the tapes are decoded.

3. A good error detection scheme should probably be used.

I concur with Larry's desire to set a specification that prevents concentration of the allowed errors into short periods of very high error rates. I will leave it to others to devise this specification.