VLB ARRAY MEMO No. 61

892

1. The VLBA Playback Correlator System

1/26/82
B. Clark

The basic correlator system specifications are given below. This is followed by a detailed discussion of one way to meet these specifications and then by a very brief discussion of alternatives. Of particular difficulty in setting the specifications is deciding on just what the requirements are for spectral line observations. The VLB observations of the maser sources can impose as great a load of data processing requirements as anything in radio astronomy. Indeed, if one were to try to get all the information possible out of an observation, the data processing requirements quickly rise into the realm of the absurd. The correlator specifications below are a compromise, meant to cover most cases of astrophysical interest without driving the correlator costs beyond reasonable bounds. In this connection, it should be noted that it would not be unreasonable for the correlator to spend one or two percent of its time replaying line observations, if this would halve the size of the required correlator.

Unfortunately, exactly how one states the specifications depends to some small extent on how one eventually ends up implementing the device. In the specification section following, the line system specifications are stated in the form most natural for the recirculating implementation discussed following. If a non-recirculating system eventually is chosen, a slight alteration in the form of the specifications as well as in their content would result (there is no way we would build individual hardware for 4096 line channels).

1.1. Correlator Specifications

The correlator interfaces to the tape recording system. Each station of the array will have a corresponding 'station' in the playback system. That is, the playback system will have a set of recorders of identical type to those of the record system at the station to play back that station's data. The correlator specifications therefore begin with the specifications of the tape recording system.

In addition, as discussed above, the playback processor should support the desireable feature of being able to add observations from other observatories to those of the dedicated array, for better (u,v) coverage or for more sensitivity. It is sufficiently desireable to do this that we believe that a 14 station processor should be built. This would permit real-time reduction of the dedicated array and four additional observatories. In addition, for normal operation of the dedicated array alone, it allows four spare playback stations, which can be switched in in case of failures, considerably smoothing out the maintenance requirements of the playback processor. This extra capability in the playback processor increases the cost only by the station recorders and electronics; as is discussed below, the extra correlator capacity is needed to handle the spectral line case.

1.1.1. Continuum correlator specifications

The requirements on the continuum processor are primarily set by the size of the field of view to be synthesized by the array. This sets two requirements--the number of delay channels needed and the frequency with which data must be recorded for further processing. We have adopted a specification of a field of view 1" radius, 2" diameter. The diameter of the earth is approximately 42 ms. The delay range necessary to synthesize a 2" field is thus 42 ms times 2" (expressed in radians) or 0.4 microseconds. An additional allowance must be made for a possible clock error. For a dedicated array with continuing coherence checks, it would seem quite feasible to keep track of clocks to an a priori accuracy of 0.3 microseconds. Thus the total delay range needed by the processor is about 1 microsecond.

The field of view also determines the required rate of recording of the correlation function. This relationship also involves the frequency of observation. The desired relationship is

$$BFST \ll 1$$

where B is the longest baseline length in nanoseconds, F the observing frequency in GigaHertz, S the field of view in radians, and T the integration time, also in radians. In more practical units, for a 2" diameter field of view, a 42 ms baseline, and avoiding only serious effects (of order ten percent) of too long integration gives

$$T < 30/F$$

where T is now in seconds. Since we want to work to frequencies of at least 25 GHz with this field of view, and preferably 50 GHz, we must be prepared to dump the correlator at intervals of 1, or even 0.5 seconds.

We may now summarize the specifications for the continuum operation of the playback correlator as follows:

Number of bit streams per station: 4

Bit rate per stream: 25 MBits/second (100MBits aggregate rate)

Polarization processing: All four cross products of two pairs of bit streams.

Simultaneous delay range: 1 microsecond (32 channels at 40 ns/channel)

Number of stations: 14

Integrator dump rate: Selectable from 0.5 to 30 seconds.

Derived quantities

Number of baselines: 91

Number of complex quantities to be calculated per baseline
and delay channel: 8

Number of complex quantities to be calculated per baseline: 256.

Length of each complex quantity: 4 bytes

Number of bytes output per correlator dump: 93184 bytes

Maximum output data rate: approx 200 kbytes/second

In the implementation discussed below we have

Number of physical complex correlators per baseline: 64

Number of simple real correlators: 11648

1.1.2. Line observation specifications

As Mentioned above, it is easily possible to think up astronomically reasonable requirements which place quite unreasonable demands on the correlator system. The specifications which are set down here are distinctly compromises--they are things which one might very often want to do, and carry no implication of being a "worst case".

Since the line system is a greater strain on the system than the continuum system, the first compromise to come to mind is to reduce the number of stations from 14 to the 10 stations of the dedicated array. Reducing the number of baselines from 91 to 45 immediately halves the number of correlations to be done.

It is necessary, or at least extremely desireable, to be able to process simultaneously full polarization information in two separate bands (for instance, two OH maser transitions).

It is necessary to process simultaneously a band at least as wide as 12.5 MHz (160 km/second for water masers).

Resolutions of 0.5 km/sec are necessary at both water vapor and at OH (40 kHz and 3 kHz respectively).

A single band (no polarization processing) should be processable with better than 1 km/sec resolution for a width exceeding 150 km/s at both water vapor and OH masers.

The correlator (and, of course, the recording systems) should be able to operate in the following modes:
A) Two bands, two IFs per band, full polarization processing.
B) One band, full polarization processing (two bit streams are idle).
C) Four bands, no polarization processing
D) Two bands, no polarization processing (two bit streams are idle).
E) One band (three bit streams are idle).

It is interesting to note that surplus recorder capability is available in modes B, D, and E above. An interesting use for this capability would be to allow three level sampling. This would give a higher signal-to-noise ratio for these cases by a factor of nearly 1.3 for minimal extra cost. Using the special purpose integrated circuits developed for the VLA automatically provides this on a chip basis; the extra cost in the correlator is only the wiring to carry the second bit to the chip. The extra cost at the antenna would be the provision of a three level sampler and the switching logic to turn it on when needed.

The observation of water masers can quickly lead to requirements beyond all reasonable bounds in total data output rate. It is commonly the case that the total extent of the masing region exceeds 10" on the sky, and that the ratio of feature width to the whole velocity width of the masing region exceeds 1:500. To maintain a 10" field of view at 23 GHz would require a dump time of 0.2 seconds. 500 complex correlations per baseline gives a total data output rate of 450 kBytes/second. This output rate could be handled (barely) by a high density tape drive (6250 bytes per inch) running at an average speed of 75 inches per second. It is doubtful that the correlator computer system could sustain this data rate, and it is clear that any reasonable data processing system would take several times slower than real time to process it (even super computers can't do all that much to a byte of data in the 2.2 microseconds before the next one arrives). For this and similar practical reasons, we arbitrarily specify that the data output rate should not exceed 400 kBytes/second, and that the maximum number of channels per baseline of interest is 4096 (4096 complex numbers for each of 45 baselines will amount to 0.75 to 3 MBytes of storage, depending on how it is managed; in any case, it is a practical and manageable amount with current technology).

These specifications are summarized below.

Number of stations: 10

Number of frequency channels: Variable, 32 to 4096

Number of IF channels per station: selectable--1, 2, or 4

Maximum data rate per channel: 25 Mbits/sec

Maximum data output rate: 400 kBytes/sec

In the following section a possible correlator design will be discussed, based on the VLA correlator technology. This is a recirulating design, which runs all multipliers at their maximum rate all the time. Therefore, the number of frequency channels that can be generated is dependent on the incoming data rate. Specifically, in the 5 possible observing modes discussed above, we have the following table for number of spectral channels for the various possible sample rates.

Sample Rate MHz	Band-width each band	A numb	A res kHz	B numb	B res kHz	C numb	C res kHz	D numb	D res kHz	E numb	E res kHz
25	12.5	32	391.	64	198.	64	198.	128	98.	256	49.
12.5	6.25	64	98.	128	49.	128	49.	256	24.	512	12.
6.25	3.12	128	24.	256	12.	256	12.	512	6.	1024	3.
3.12	1.56	256	6.	512	3.	512	3.	1024	1.5	2048	0.76
1.56	0.78	512	1.5	1024	0.76	1024	0.76	2048	0.38	4096	0.19
0.78	0.39	512	0.76	1024	0.38	1024	0.38	2048	0.19	4096	0.10
0.39	0.19	512	0.38	1024	0.19	1024	0.19	2048	0.10	4096	0.05
0.19	0.10	512	0.19	1024	0.10	1024	0.10	2048	0.05	4096	0.02

1.2. Implementation

The correlator specified above is a rather large one, at least in terms of correlators currently on line, if not in terms of correlators designed. The only correlator of this size in current production work is the VLA correlator. A larger correlator was prototyped for the NASA SETI project, but it is not in current production. It therefore seems most conservative to design the correlator using the VLA technology and philosophy. Since it has been done, and works well, one can set out to design a correlator of this size with reasonable confidence that it will work when complete. Other designs have not been attempted in this size, and may require more extensive design and prototype work.

Although the correlator is a relatively straightforward design and construction project, the tape recording system is much less so. It is enevitable that modifications will be made in it based on field experience. It seems very desireable to have the correlator itself independent of such changes. A well designed interface between the tape system and the correlator system does make them quite independent. The concept of having the playback system provide four bit streams at 25 Mbit/sec each, and having validity bits with well defined properties would lead to an interface sufficiently simple for both sides to work to that the independence could be maintained. The importance of this is that the correlator can be built in its entirety from the beginning, a substancial saving in money over having to build a prototype correlator system to get the initial field experience which will surely be necessary for the recording system.

An additional impetus toward the VLA correlator philosophy is given by simple cost considerations. A look at the total costs of recently constructed correlators seems to indicate that the cost is proportional to the number of channels, and nearly independent of speed. It is therefore reasonable, for correlators large enough to justify the rather large design effort of the controllers, to run all multipliers at the maximum possible rates. This rate is, for the VLA correlator technology, 100 Mbits/sec. Even for the continuum case, a recirculating correlator running at this rate requires four times fewer hardware multipliers than a simple correlator running at 25 MHz, and is therefore expected to be substancially less expensive. The contrast in the spectral line case is even greater.

The playback correlator system consists of the following component parts:
1) The tape playback systems (one per station)
2) The bit stream handlers (one per station)
3) The baseline handlers (one per baseline)
4) The channels
5) The hardware controllers
6) The computer system

1.2.1. The tape playback systems

The tape playback system is discussed above, together with the recording systemswhich it essentially duplicates. The additional equipment required at playback time is only a buffer, to remove the mechanical variation in the playback rate, and a delay control, to that the bits are removed from the various tapes in synchronism.

1.2.2. The bit stream handlers

The interface to the correlator proper is four bit steams at a 25 Mbit/sec rate, with one or more validity bits. It might be worthwhile having several validity bits which can be handled slightly differently down-stream. For instance, it might be worth having separate bits for recorder checksum error and an antenna generated error flag (eg, antenna still slewing). On the other hand, it is possible to conceive doing with only one flag.

The bit stream handlers as part of the correlator consist of the recirculators, phase extrapolators, and delay controls. Fringe rotation is normally done in VLB interferometers by multiplying one of the two bit streams going into the correlator by a three level (-1,0,+1) signal. Two three level signals displaced by 90 degrees of phase are used. This is the equivalent of an analog mixer. The desired function is a single sideband mixer, and the sideband is selected by performing the appropriate operations on the two correlator outputs to reject the unwanted sideband of the fringe frequency. Since the single sideband mixer is not completed until after correlation, it is not practical to perform the mixing operations on the bit streams from each antenna separately. The cross frequencies of the fringe rates are not completely rejected, and a cost in signal- to-noise ratio is incurred. Therefore, fringe rotation must be done on a baseline by baseline basis. However, This can be simplified as much as possible by having an antenna based calculation of phase, which is transmitted to each baseline. The baseline equipment would then difference the two antenna phase signals to give the baseline phase. This is worthwhile because the phase calculation is a rather more complicated affair than it would be for a simple correlator. The recirculators are playing the data back at several times real time, and the fringe rotators must be speeded up by the same factor, and must be reloaded with a new initial value each time the recirculators are reinitialized.

The fringe rates encountered in VLB interferometers are at moderate rates--an earth-diameter baseline at 44 GHz gives a maximum fringe rate of 130 kHz. Simple TTL circuitry can be used in the fringe function extrapolators. Appropriate buffers for their rates and initial values can be loaded by the control microprocessor. It is probably worthwile having a hardware delay extrapolator as well as a hardware fringe function extrapolator. The maximum delay rate on an earth baseline interferometer is about three microseconds per second, or one bit (25 Mbits/sec) in 12 microseconds. This would be doable by a fast microprocessor, but there is one additional complication. Since the delay is applied in unit steps of one bit, it is not possible for a multistation processor to delay the bit streams on an antenna basis; to do so means that the maximum correlator induced delay error is one bit time, rather than the desired half bit time. Therefore, the decision one the last bit of delay must be made on a baseline basis, in the same way that fringe rotation must be done on a baseline basis.

If one has a hardware delay extrapolator, it is reasonable to add its fractional bit outputs to the fringe extrapolator. This will mean that phases track the fringes at the band center frequency when the phase rotator is initialized with parameters suitable for the signed sum of local oscillators, and that nothing further up in the system need know about the "fractional bit" of delay. Since the system is large and complicated, doing this at the hardware level will remove a likely source of error.

The recirculators themselves are simply a memory, which are filled at the 25 MBit/sec basis record system rate, and dumped at the 100 MBit/sec correlator system rate, to give maximum utilization of the correlator hardware.

There are several ways to use recirculators, and it is not clear which will be most advantageous. The one suggested here is that a separate recirculator be provided for each of the four 25 MHz bit streams. In the continuum mode, the correlator can perform four opperations in the time taken for the recirculator to refill with new data. These could be to compute the polarization parameters of the radiation. That is if there are two left hand polarized IFs, LA and LB, and two right handed, RA and RB, a single section of the correlator could compute the four products from two antennas (1 and 2) RA1*RA2, LA1*LA2, RA1*LA2, and LA1*RA2. These four products carry the intensity and polarization information at the band ("A") from which these two IFs arise.

1.2.3. The baseline equipment

As mentioned in the preceeding section, there are several devices which must occur on a once-per-baseline basis. These are:

1) The phase rotator.
2) The plus or minus one bit extra delay.
3) A correlator channel to keep track of attempted correlations.
4) Logic to OR the flags from the two antennas involved.
5) A multiplexor to achieve the reconfiguration between the basic 14 station continuum configuration and the 10 station line configuration.

The phase rotator conventionally used in VLB correlators is a three level multiplication, in which the 0 level is present 1/4 of the time, and +1 and -1 are present 3/8 of the time each. This device has a loss in signal-to-noise ratio of about 4 percent, but its simplicity recommends it. The phase rotator then consists of a subtractor of the two phase values from the two antennas in question (four bits appears to be enough), and a one-of-eight decoder. Two of the eight lines suppress correlation and three invert one of the bit streams going into the correlator.

Although it should be possible, in principle, to calculate from the initial values just what the duty cycle of the correlator is, in practice, it is not easy to get right, and it would appear that an additional channel per baseline to get this number exactly (it is needed exactly because the correlator channels are up-counters, and therefore have a large bias which must be subtracted) would save a large amount of effort elsewhere.

1.2.4. The correlator channels

The VLA correlator channels are implemented in specially designed integrated circuits. One custom circuit incorporates multipliers, the other the fast integrator. The fast integrator holds the results of the integration for one recirculator full of data. At the end of a recirculator cycle, the fast integrator adds its contents to the main correlator memory (a rather conventional semiconductor memory) and is cleared for the next recirculator buffer.

As discussed above, there are several ways to use recirculators. In the implementation suggested here, the physical correlators are used to make a complex correlation (that is, with two orthogonal fringe functions applied) on a single pair of bit streams for 32 lags. Although the correlator is running with a 100 MHz clock, the bit streams were originally sampled with a 25 MHz clock, so the spacing of the lag channels is 40 nanoseconds. The four recirculations are used to calculate the four cross products of two pairs of bit streams (that is, two bit streams from each antenna). These four cross products carry the polarization information of the correlated radiation (they are linear combinations of the four stokes parameters).

There are actually four bit streams from each antenna, so the correlator channels and baseline equipment mentioned above must be duplicated in its entirity for the second pair of bit streams.

Let us now repeat the litany of numbers.
For each delay channel there is one complex physical correlator,
    or two simple real physical correlators.
For each baseline and bit stream pair there are 32 delay channels,
    32 complex physical correlators, 64 real physical correlators.
For each baseline there are two bit stream pairs, totalling 64
    complex correlators or 128 simple correlators.
There are 91 baselines, giving a grand total of 5824 complex
    correlators, or 11648 simple correlators.

In addition, there are the duty cycle correlators, two complex correlators per baseline, 182 complex or 364 simple correlators in all.

For spectral line use it is also desireable to have autocorrelation spectrometers on each of the incoming bit streams. This probably requires only 32 real correlators per antenna. Recirculation should take care of synthesising the number of needed channels and of processing all four bit streams. Thus 320 additional real correlators are necessary for the autocorrelation function for spectral processing of the 10 antenna spectral line array (in practice we would probably put them on all 14 stations to allow convenient spectral processing of the 14 station array with half the number of channels per baseline discussed above).

1.2.5. The correlator computer

The correlator proper is probably controlled most conveniently by a fast, bit slice microprocessor, as is done in the VLA correlator. This in turn must be controlled by, and pass data to, a more general purpose computer, which will do the necessary geometric calculations to provide fringe rates and starting phases, and which will format the data in convenient forms for post correlation processing.

The geometric calculations are not particularly ornerous, but must merely be done with a reasonable degree of care. To track fringes smoothly on a long baseline at 50 GHz requires an accuracy of about 34 bits. To attain this requires about 40 bits internal computation width. There is no problem in attaining this in any modern minicomputer. Again, to track fringes smoothly on a long baseline requires that the fringe phase be updated about every 0.1 seconds. Calculating the geometry of the fourteen station array will take only a few percent of this time on any modern minicomputer, if the calculation is properly organized.

The data handling is a bit more severe in its requirements. As mentioned above, Output data rates of up to 400 kBytes/second may be encountered. This carries the strong implication that input data rates will be at least as great (in a well ordered world computers perform data reduction, not data expansion). To handle this, and to resolve gracefully the inevitable bus contention problems, the computer should be specified with an agregate I/O bandwidth of at least 2 MBytes/sec.

Perhaps the most computation intensive chore of the correlator computer is to convert from lag spectrum to cross correlation spectrum in the spectral line case. The FPS AP120B array processor does longish fourier transforms at the rate of about 500,000 output points per second. However, since the fourier transform is being used to complete the quadrature mixer, half the output spectrum is empty, and thus a single FPS AP120B processor is fairly well matched to performing this chore with an output data rate of 400 kByte as discussed above.

1.3. Alternative approaches

There are many ways to design a correlator. The most obvious is to do without the complication of the recirculating design. This requires four times the number of physical correlators. It does do away with a complex part of the controller, as well as the recirculators themselves. Constructed in similar technology, one would expect the non-recirculating design for a correlator of this size to be two to four times more costly than the recirculating design discussed above. Its only chance of being economically competitive is if it can be more highly integrated. The difference between a 100 MHz clock rate and a 25 MHz clock rate, at which the non-recirculating design runs, may make the difference in being able to use Schottkey logic rather than the ECL logic, and thus enable another factor of two or four in density on a specially designed circuit. It does not seem like a profitable approach to implement with standard integrated circuits. One implementation of a three level correlator with standard circuits requires two j-k type flip-flops and three gates. Integration for 10 seconds at a 25 MHz clock requires an additional 27 stages of binary counter. With conventional MSI integrated circuits this is nine packages with a total power dissipation of approximately 0.6 watt (assuming conversion to CMOS after the first counter stage). Thus the entire correlator would amount to over 400,000 ICs drawing over 8000 amperes of power at 5 volts.

A second approach of rather more interest is that planned for the large correlator on the Nobeyama interferometer array. In this approach, the incoming bit stream is fourier transformed into individual frequency channels, and the channels are then multipied together and accumulated to give a correlated power spectrum, just what you want for the spectral line case. For use as a continuum processor, one would generate, say spectral channels of width 780 kHz which are narrow enough that the possible clock errors and delay range do not reduce the amplitude of the fringes. One can then do a fourier transform back into lag space, to be able to discard lags not of interest. A complete design of such a device has not been made, so the cost and operating properties are now well known, but some remarks on the subject are made below.

It is a property of the fourier transform that the number of points going into the transform is equal to that coming out. That is, the data rate is not changed by the fourier transform. There is a minor exception--for economy of tape bandwidth use, the incoming bit streams are two-level or three-level signals. The output cannot also be similarly truncated without the loss of significant information. It is probably sufficient to add one additional significant bit for every two levels of FFT butterfly, plus about two guard bits. Thus, an eight bit output would probably suffice for a 4096 channel spectrometer.

A second increase in the data to be handled comes from the fringe rotation. The most natural way to handle fringe rotation is to apply a phase to the incoming bit streams. This makes each sample a complex rather than a simple real number, doubling the amount of data to dled thereafter, essentially the same doubling encountered in the omain correlator in the application of two orthogonal three level ringe functions.

The advantage of the fourier transform correlator is that the one-per-baseline equipment becomes very simple. It is simply a few high speed multibit multipliers. For instance, eight eight bit complex multipliers per baseline that run at 25 MHz will extract the four polarization parameters from two pairs of orthogonally polarized bit streams of 25 MHz each.

Because of the general usefulness of the FFT algorithm, a considerable effort has been devoted to hardware implementations. Certainly, real-time transforms of 8 bit precision and 25 MHz rates have been implemented. The question is simply whether such devices are economical for the VLB correlator, which would contain 56 of them.

Thus, the hard part of the correlator would be implemented in 56 high speed fourier transform devices and 2912 eight bit high speed multipliers. This impressively low device count suggests that this approach may be economically effective. However, these devices are relatively complex, especially considering the speed with which they must operate.A detailed design and costing must be done on such an approach. Further investigation, while obviously necessary, may disclose serious problems. Therefore, the cost information in this proposal is based on the more conservative assumption that the correlator will be based on the VLA correlator philosophy and technology.